Aid Effectiveness and Intervention Outcomes in a Fragile Situation: Comparative Evaluation of Three Aid Management Modalities in Timor-Leste’s Health Sector

Hasibul Haque

A thesis submitted for the degree of
Doctor of Philosophy
at the University of Otago, Dunedin, New Zealand

2015
1 Abstract

1.1 Background

In the context of long and inconclusive discourse on the development effect of foreign aid on the recipient country, the Paris Declaration of Aid Effectiveness (2005) espoused five principles of aid effectiveness: recipient country ownership, alignment, harmonization, managing for results, and mutual accountability. However, in fragile states or situations, where the government lacks capacity, or willingness to ensure basic services for its people, donors often employ multiple aid management modalities ranging from state-building and state partnerships to state-avoidance for externally funded health interventions. Very few studies have been conducted so far to evaluate to what extent different aid management modalities in a fragile state’s health sector comply with the Paris principles of aid effectiveness and if such compliance is associated with better outcomes of the externally funded interventions for achieving programmatic objectives as well as strengthening the health systems.

1.2 Research question

This research attempted to conduct a range of pilot studies that critically evaluate potential analytical tools used for comparing three different aid management modalities in Timor-Leste’s health sector during the period from 2007 to 2013. These aid management modalities were: the state-building multi-donor trust fund (MDTF) approach of the AusAID and World Bank funded Health Sector Strategic Plan Strengthening Project (HSSP-SP), the state-partnership hybrid approach of the Global Fund funded National HIV/AIDS and STI Programme (NAP), and the state-avoiding NGO contracting approach of the USAID funded Immunizasaun Proteje Labarik (IPL). The pilot evaluation approaches used in this research wanted to investigate to what extent each modality adhered to the Paris principles of aid effectiveness, their programmatic results in terms of achieving intervention objectives as well as health systems strengthening, and if adherence to the Paris principles were linked to the programmatic outcomes.
1.3 Methods

This research employed a theory driven ‘realist’ evaluation approach using mixed methods. Realist synthesis and qualitative interpretative methods were used for analyzing context and aid mechanisms, collecting stakeholders’ views, and triangulating and interpreting findings. A range of quantitative methods and tools were used for analyzing and presenting stakeholder interviews in the form of balanced scorecards; conducting cost-effectiveness and technical efficiency analysis; comparing performance efficiency through data envelopment (DEA) technique; and looking for correlations among different variables. In order to measure the contribution of an intervention to the health outcomes, a quasi-experimental design was used for cost-effectiveness analysis by mathematically modeling the counterfactual scenarios of likely effect of the absence of an intervention on outcomes. Results were validated by comparing the findings with each other and by triangulating the empirical evidence from the qualitative and quantitative analysis with their underlying programme theories.

1.4 Results

Stakeholders viewed the NAP as more compliant with the Paris principles of aid effectiveness than the HSSP-SP and IPL. While the NAP demonstrated good economy, efficiency, and effectiveness, the chance of sustainability of its results beyond the funding period was perceived to be lowest among three projects. IPL’s compliance with the Paris principles was viewed as lowest among three projects. But IPL’s performance on economy, efficiency, and effectiveness was believed to be closely matching with those aspects of the NAP and IPL’s results were believed to be more sustainable than the other two projects in terms of willingness and capacity of the government to continue the intervention on termination of the donor funding. HSSP-SP’s performance on economy, efficiency, and effectiveness was believed to be lowest among three projects. However, HSSP-SP scored slightly higher than IPL for compliance with the Paris principles. For sustainability of results, HSSP-SP scored higher than the NAP.

The cost effectiveness analysis showed that HSSP-SP saved a total of 44,628 disability adjusted life years (DALYs) from June 2008 to December 2012 with a $363.90 cost per DALY saved. The NAP saved a total of 33,325 DALYs from June 2007 to December 2013
with a $325.25 cost per DALY saved. The IPL, on the other hand, saved a total of 246,233 DALYs from April 2011 to September 2013 with a $10.72 cost per DALY saved. Comparing these results with their respective targets gives a technical efficiency of 53.62% for HSSP-SP, 67.77% for the NAP, and 53.04% for IPL.

The DEA results comparing the three projects showed that HSSP-SP’s performance could be improved in the areas of economy of procuring goods and equipment, and achieving more project outputs by using the same level of inputs. The NAP, on the other hand, could improve the effectiveness of its performance by increasing its investments on health systems strengthening and achieving more health system strengthening related outcomes. IPL could also improve its performance by achieving more project outputs with the same level of inputs.

An attempt to run correlation analysis among various variables indicated a plausible correlation between the compliance with the Paris principles and technical efficiency for cost per DALY saved by three projects. However, due to the small sample size, significance of this finding could not be concluded.

### 1.5 Discussion and conclusion

Contrary to the belief that a sector wide approach (SWAp) through the MDTF modality more readily complies with the Paris principles of aid effectiveness, the stakeholders in Timor-Leste’s health sector believed that the vertical programme such as the NAP funded by the Global Fund rather complied with the Paris principles more than the SWAp. While the contracted service delivery model through the NGOs is believed by some experts as weakening the necessary state-building efforts in a fragile situation, the stakeholders in this research viewed the IPL approach with limited budget, time-bound exit plan, and working closely with the government very handy for taking over and owning by the government for greater likelihood of continuation and sustainability of intervention results beyond the funding period.

The research indicates that aid management modality matters in achieving their programmatic outcomes and having an effect on the health systems in a fragile situation. However, aid effectiveness of a management modality is not solely determined by its
design; rather, it is much determined by how the modality is implemented and interacts with its particular context. Increased compliance with the Paris principles of aid effectiveness shows good indication of delivering better efficiency and intervention outcomes even in fragile situations.

The limited attempts so far to evaluate the impact of the implementation of the Paris principles in the health sector claimed some clear pathways between the adherence to the Paris principles and better outcomes for health results. However, none of those studies could establish a clear causal link between the compliance with the Paris principles and development effectiveness. The pilot studies attempted in this research were constrained by small sample size (n=3), comparability of the outcomes of very different interventions, and limitations of the Likert scaling method for analyzing stakeholder views. Nevertheless, the analytical tools and techniques used in this research showed potential as an approach for evaluating aid funded projects. It appears that the approach can be more useful with increased sample size and more sophisticated statistical analysis. Further research is recommended to strengthen the empirical evidence base and before any generalized conclusion is drawn about correlation between the adherence to the Paris principles and better health results.
2 Acknowledgements

I am indebted to the following people and institutions for their invaluable support and contribution in completing this thesis:

• The University of Otago for offering me a doctoral research place, scholarship, and performance based research funding for undertaking this research;
• The Technical and Ethical Committee of the Office of the Health Research and Development of the MoH of Timor-Leste and His Excellency the Minister of Health of Timor-Leste for providing ethical approval and permission for accessing required information and data for this research;
• Key informants and interview participants in Timor-Leste for their time and support they kindly extended to this research;
• My supervisors Professor Robin Gauld and Professor Philip Hill for all their valuable support, review, and guidance;
• Associate Professor Dr. Peter Hill of the University of Queensland for reviewing the primary research plan as an external adviser
• Dr. Trudy Sullivan for reviewing and commenting on the research outlines and the chapter on cost effectiveness analysis;
• Dr. Nisha Nair and Dr. Giorgi Kvizhinadze of the Burden of Disease Epidemiology, Equity and Cost-Effectiveness Programme (BODE³) for reviewing and commenting on the chapter on cost effectiveness and technical efficiency analysis;
• Professor Rob McGee and the Research Advisory Committee of the Department of Preventive and Social Medicine of the University of Otago for reviewing the research proposal and periodical progress of this research;
• And, of course, families and friends were always there to provide essential psychological nutrients and support when I needed them most!

I acknowledge with gratitude all their contributions to help this work progress to this stage.
# 3 Table of contents

1  Abstract ........................................................................................................... iii
   1.1  Background ................................................................................................. iii
   1.2  Research question ...................................................................................... iii
   1.3  Methods ...................................................................................................... iv
   1.4  Results ........................................................................................................ iv
   1.5  Discussion and conclusion ....................................................................... v

2  Acknowledgements ......................................................................................... vii

3  Table of contents .......................................................................................... viii

4  List of tables ................................................................................................ xiv

5  List of figures ................................................................................................ xviii

6  Abbreviations ................................................................................................ xx

7  List of annexes ............................................................................................... xxiii

1  Introduction ...................................................................................................... 1
   1.1  Background ................................................................................................. 1
   1.2  Rationale .................................................................................................... 2
   1.3  Research question ...................................................................................... 4
   1.4  Overview of research process .................................................................. 5
   1.5  Key definitions .......................................................................................... 7
   1.6  How this thesis is organized .................................................................... 10
   1.7  Expected contribution of this thesis ......................................................... 12

2  Review of related literature .......................................................................... 14
   2.1  Introduction ................................................................................................. 14
   2.2  Methods ..................................................................................................... 14
   2.3  Changing concepts of aid effectiveness .................................................. 17
       2.3.1  Macro-economic focus of aid effectiveness ........................................... 17
       2.3.2  Aid delivery principles and practices .................................................. 18
       2.3.3  Development impact of aid interventions .......................................... 20
   2.4  Systematic review of literature related to the Paris principles in the health sector 21
       2.4.1  Evaluation of implementation of the Paris principles in the health sector .... 22
2.4.2 Other factors influencing aid effectiveness..............................................................27
2.4.3 The Paris principles in fragile situations..................................................................29
2.4.4 The Paris principles in different aid modalities.........................................................30
2.4.5 Evaluation methods for the Paris principles..............................................................32

2.5 Selected literature for scoping .......................................................................................33
2.5.1 State fragility .............................................................................................................34
2.5.2 Health systems .........................................................................................................36
2.5.3 Aid instruments in a fragility context ........................................................................37
2.5.4 Evaluation theories and techniques .........................................................................40
2.5.5 Aid effectiveness evaluation .....................................................................................42
2.5.6 Methods and tools of comparative analysis ..............................................................45

2.6 Critical appraisal of studies dealing with impact of the Paris principles on health outcomes..................................................................................................................49

2.7 Conclusion ..................................................................................................................54

3 Country context..............................................................................................................55
3.1 Introduction ..................................................................................................................55
3.2 Geography ...................................................................................................................55
3.3 People and culture .......................................................................................................57
3.4 History ........................................................................................................................58
3.5 Economy ......................................................................................................................61
3.6 Politics ..........................................................................................................................63
3.7 Millennium Development Goals (MDGs)..................................................................64
3.7.1 MDG 1: Eradicate extreme poverty and hunger .........................................................65
3.7.2 MDG 2: Universal enrollment in primary education ..................................................66
3.7.3 MDG 3: Promote gender equality and empower women .........................................66
3.7.4 MDG 4: Reduce child mortality ..............................................................................67
3.7.5 MDG 5: Improve maternal health ..........................................................................67
3.7.6 MDG 6: Combat HIV/AIDS, malaria and TB ............................................................68
3.7.7 MDG 7: Ensure environmental sustainability ..........................................................68
3.7.8 MDG 8: Develop a global partnership for development ..........................................69

3.8 Discussion and conclusion ..........................................................................................70

4 Aid instruments and results framework: analysis of project backgrounds ............72
4.1 Introduction ..................................................................................................................72
4.2 Methods .......................................................................................................................73
4.3 Health Sector Strategic Plan- Strengthening Project (HSSP-SP) .............................75
5.3 Theoretical framework for aid effectiveness evaluation

5.4 Research design and methods

5.5 Quality control of research

5.6 Conclusion

Conclusion

Theoretical framework and research methods

Aim

Objectives

Setting

Interventions to evaluate

Research questions

Research methods

Quality control of research

Ethical approvals

Triangulation

Checking against ‘triple-A’ requirements

Analysis and discussion

Conclusion

Conclusion

Immunizasaun Proteje Labarik (IPL)

Background

Goals and objectives

Activities

Management modality

Results framework

The Global Fund funded National HIV/AIDS and STI Programme (NAP)

Background

Goals and objectives

Activities

Management modality

Results framework

Activities

Goals and objectives

Management modality

Results framework

Aim

Objectives

Setting

Interventions to evaluate

Research questions

Research methods

Quality control of research

Ethical approvals

Triangulation

Checking against ‘triple-A’ requirements

Conclusion
6 State fragility and health systems: a case study of Timor-Leste .......................... 128

6.1 Introduction ........................................................................................................... 128

6.2 State fragility analysis ......................................................................................... 128

6.2.1 Methods of state fragility assessment ............................................................... 129
6.2.2 State fragility indices for Timor-Leste .............................................................. 131
6.2.3 Timor-Leste’s self assessment of fragility ......................................................... 135
6.2.4 Factors of fragility in Timor-Leste ................................................................... 136

6.3 Health system analysis ......................................................................................... 143

6.3.1 Methods of health systems analysis ................................................................. 144
6.3.2 Health governance ............................................................................................ 145
6.3.3 Health service delivery ...................................................................................... 147
6.3.4 Medicine and technologies .............................................................................. 149
6.3.5 Health information management system ......................................................... 150
6.3.6 Health workforce ............................................................................................. 151
6.3.7 Healthcare financing ......................................................................................... 152

6.4 Discussion ............................................................................................................ 156

6.5 Conclusion .......................................................................................................... 160

7 Aid effectiveness and stakeholders’ perspectives: balanced scorecards ...... 162

7.1 Introduction ........................................................................................................... 162

7.2 Balanced scorecards and the research framework ............................................. 162

7.3 Methods used ...................................................................................................... 163

7.3.1 Sources of information .................................................................................... 163
7.3.2 Fieldwork and stakeholder interviews .............................................................. 163
7.3.3 Selection of stakeholders ................................................................................. 164
7.3.4 Focus group discussions ................................................................................. 165
7.3.5 Questionnaire .................................................................................................. 166
7.3.6 Importance weighting ...................................................................................... 167
7.3.7 Data processing and analysis ......................................................................... 167
7.3.8 Presentation and interpretation of scores ......................................................... 168

7.4 Results ................................................................................................................ 169

7.4.1 Summary of overall response and importance weight .................................... 169
7.4.2 Compliance with the Paris principles .............................................................. 173
7.4.3 Economy .......................................................................................................... 179
7.4.4 Efficiency .......................................................................................................... 181
7.4.5 Effectiveness ..................................................................................................... 184
8 Cost-Effectiveness and technical efficiency ........................................................................ 204

8.1 Introduction ...................................................................................................................... 204
8.2 Key concepts and cost-effectiveness analysis framework .............................................. 204
8.3 Cost effectiveness and technical efficiency of HSSP-SP ................................................. 208
  8.3.1 Methods ......................................................................................................................... 208
  8.3.2 Results for HSSP-SP ..................................................................................................... 212
8.4 Cost effectiveness and technical efficiency of the NAP ................................................. 216
  8.4.1 Methods ......................................................................................................................... 216
  8.4.2 Results for the NAP ....................................................................................................... 224
8.5 Cost effectiveness and technical efficiency of IPL ......................................................... 240
  8.5.1 Method ......................................................................................................................... 240
  8.5.2 Results for IPL ............................................................................................................. 244
8.6 Comparison of the Cost-effectiveness and technical efficiency of the three projects 252
  8.7 Sensitivity analysis ........................................................................................................... 252
    8.7.1 Uncertainty range for HSSP-SP ................................................................................... 252
    8.7.2 Uncertainty range for the NAP .................................................................................. 253
    8.7.3 Uncertainty range for IPL .......................................................................................... 254
    8.7.4 Results of sensitivity analysis .................................................................................... 255
8.8 Discussion .......................................................................................................................... 255
8.9 Conclusion ........................................................................................................................ 265

9 The Paris principles and comparative economy, efficiency, and effectiveness .................. 266

9.1 Introduction ....................................................................................................................... 266
9.2 Data Envelopment Analysis: concept and techniques .................................................... 266
9.3 Methods ............................................................................................................................ 270
  9.3.1 Data sources .................................................................................................................. 270
  9.3.2 Calculating economy ..................................................................................................... 271
  9.3.3 Calculating performance efficiency ............................................................................. 271
  9.3.4 Calculating effectiveness .............................................................................................. 273
  9.3.5 DEA model used ......................................................................................................... 273
9.4 Results ........................................................................................................................................274
  9.4.1 Economy ..................................................................................................................................274
  9.4.2 Performance efficiency ..............................................................................................................277
  9.4.3 Effectiveness .............................................................................................................................284
  9.4.4 Summary of DEA results ..........................................................................................................287
9.5 Looking for correlations ..................................................................................................................287
  9.5.1 Data sources ..............................................................................................................................288
  9.5.2 Method and results ....................................................................................................................289
9.6 Discussion ......................................................................................................................................295
9.7 Conclusion .....................................................................................................................................300
10 Lessons learned ...............................................................................................................................301
  10.1 Introduction .................................................................................................................................301
  10.2 Balanced scorecards ......................................................................................................................301
  10.3 Cost effectiveness analysis ..........................................................................................................305
  10.4 Data Envelopment Analysis (DEA) ..............................................................................................308
  10.5 Correlation analysis .....................................................................................................................311
  10.6 Conclusion ..................................................................................................................................313
11 Summative discussion and conclusion ..............................................................................................315
  11.1 Introduction ..................................................................................................................................315
  11.2 Summary of research findings .....................................................................................................317
    11.2.1 Contextual analysis ..................................................................................................................317
    11.2.2 Project backgrounds and instruments ......................................................................................318
    11.2.3 Stakeholders’ views, cost effectiveness analysis and DEA ......................................................318
    11.2.4 Correlation ................................................................................................................................321
  11.3 Discussion ....................................................................................................................................321
  11.4 Rigour and limitations of this research .........................................................................................326
  11.5 Conclusion ..................................................................................................................................331
    11.5.1 Recommendations for aid projects ..........................................................................................332
    11.5.2 Implications for further research .............................................................................................335
12 References ..........................................................................................................................................337
4 List of tables

Table 2.1: Search and primary screening of related literature ..................................................15
Table 2.2: Year of publication of sample literatures on Paris principles in the health sector ...........................................................................................................................................21
Table 2.3: Literature selected for systematic review ................................................................22
Table 2.4: Paris Declaration monitoring indicators and progress, 2010..................................43
Table 2.5: Critical summary of studies dealing with impact of the Paris principles .............50
Table 3.1: Population distribution in Timor-Leste (2004 and 2010) .....................................57
Table 3.2: Estimated population distribution of Timor-Leste in 2013 .................................57
Table 3.3: Timor-Leste’s annual budgets and sources of financing: 2008-2014 (in Million US Dollars) ........................................................................................................................................62
Table 3.4: Timor-Leste’s Human Development Indicators ...................................................64
Table 4.1: Monitoring indicators of Paris Declaration for intervention assessment ..........74
Table 4.2: Results framework of HSSP-SP .............................................................................79
Table 4.3: Results framework for Timor-Leste’s NAP .........................................................83
Table 4.4: Results framework for IPL ...................................................................................86
Table 4.5: Timor-Leste’s selected health interventions in light of monitoring indicators of the Paris Declaration ..................................................................................................................88
Table 4.6: Advantages and disadvantages of donor financing options for health services in fragile states ........................................................................................................................................93
Table 5.1: Summary of research design and methods .........................................................108
Table 5.2: Sampling of stakeholder interviews ....................................................................113
Table 5.3: Steps in contribution analysis ..............................................................................126
Table 6.1: Common factors of state fragility .........................................................................131
Table 6.2: Fragile State Index for Timor-Leste 2007-2013 ................................................133
Table 6.3: Summary of Fragility Assessment of Timor-Leste, 2012 .................................135
Table 6.4: Health service delivery systems in Timor-Leste ..............................................147
Table 6.5: Timor-Leste’s health workforce 2001, 2008, and 2011 .....................................151
Table 6.6: Healthcare financing in Timor-Leste 2008-2013 ................................................153
Table 6.7: State budget by cost category for MoH of Timor-Leste (in thousand US dollars) ........................................................................................................................................154
Table 6.8: Main causes of morbidity and mortality in Timor-Leste, 2010 .......................158
Table 7.1: Sampling of stakeholder interviews ................................................................. 165
Table 7.2: Balanced scorecard domains and indicators used in the questionnaire ........... 166
Table 7.3: Summary of stakeholder responses .................................................................. 169
Table 7.4: Importance weight of evaluation indicators (1= least important, 10= most important) ........................................................................................................... 171
Table 7.5: Weighted scores on compliance with the Paris principles ............................. 173
Table 7.6: Score converted to equal scale for compliance with the Paris principles .......... 177
Table 7.7: Weighted scores on economy of three interventions ..................................... 179
Table 7.8: Perceived efficiency of three interventions ...................................................... 182
Table 7.9: Perceived effectiveness of three interventions .................................................. 185
Table 7.10: Perceived sustainability of three interventions .............................................. 189
Table 7.11: Summary of balanced scorecards for three interventions .............................. 191
Table 7.12: Implementation of Paris Declaration in Timor-Leste ................................... 196
Table 8.1: Overview of HSSP-SP ..................................................................................... 212
Table 8.2: Estimated DALYs per 100,000 in Timor-Leste, 2008-2012 ............................ 213
Table 8.3: Total DALYs in different scenarios and DALYs saved by HSSP-SP in Timor-Leste, 2008-2012 ........................................................................................................... 215
Table 8.4: Cost effectiveness ratio and technical efficiency of HSSP-SP ......................... 216
Table 8.5: Key parameters of epidemiological modeling used in Spectrum for Timor-Leste ...................................................................................................................... 220
Table 8.6: Impact matrix for programme coverage .......................................................... 220
Table 8.7: Overview of design, outputs, and outcomes of the NAP ................................. 224
Table 8.8: HIV prevalence estimates in Timor-Leste, 2005-2007 ..................................... 225
Table 8.9: Estimated size of risk sub-populations in Timor-Leste, 2007-2013 ................ 225
Table 8.10: STI prevalence estimates used in the Spectrum exercise ............................... 228
Table 8.11: Estimated number of sexual partners for HIV risk groups in Timor-Leste, 2007-2013 .................................................................................................................. 229
Table 8.12: Sex acts per partner and condom use by HIV risk groups in Timor-Leste, 2007-2013 .................................................................................................................. 230
Table 8.13: Estimated coverage by the NAP ..................................................................... 231
Table 8.14: ART and PMTCT services coverage in Timor-Leste, 2007-2013 .................. 231
Table 8.15: Impact of PMTCT and ART programme components in Timor-Leste, 2007-2013 .................................................................................................................. 232
Table 8.16: DALYs saved by ART and PMTCT programme components in Timor-Leste
...........................................................................................................................................234
Table 8.17: New infections and deaths averted by NAP, 2007-2013 .................................236
Table 8.18: DALYs saved by prevention components of NAP..............................................236
Table 8.19: DALYs saved by STI programme components .....................................................238
Table 8.20: DALYs saved attributable to Global Fund funded NAP activities .....................239
Table 8.21: Cost effectiveness ratio and technical efficiency of the NAP .........................239
Table 8.22: Parameters for calculating number of cases and deaths prevented by vaccination ..........................................................................................................................241
Table 8.23: Parameters for calculating DALYs averted by vaccination ................................242
Table 8.24: Overview of Immunizasaun Proteje Labarik (IPL) .............................................244
Table 8.25: Summary of DPT3 and Measles vaccination coverage in IPL-focus and IPL-
non-focus districts in Timor-Leste, 2010-2013 ...................................................................245
Table 8.26: Infections and deaths averted by vaccination in IPL-focus districts from
2011-2013 ................................................................................................................................246
Table 8.27: Target infections and deaths expected to be averted in IPL-focus districts
from 2011-2013 .........................................................................................................................246
Table 8.28: DALYs saved by vaccination in ILP-focus districts, 2011-2013 .......................247
Table 8.29: DALYs expected to be saved in optimum scenario IPL-focus districts, 2011-
2013 ........................................................................................................................................248
Table 8.30: Programmatic outcomes attributable to IPL intervention, 2011-2013 ..........249
Table 8.31: Cost per DALY saved by IPL ..................................................................................250
Table 8.32: Expected programmatic outcomes attributable to IPL in the optimum scenario.................................................................................................................................250
Table 8.33: Cost per DALY saved by IPL in optimum scenario ............................................251
Table 8.34: Comparison of cost effectiveness and technical efficiency of three projects
.................................................................................................................................................252
Table 8.35: Uncertainty ranges for HSSP-SP .........................................................................253
Table 8.36: Uncertainty analysis ranges for the NAP ............................................................253
Table 8.37: Uncertainty ranges for IPL ....................................................................................254
Table 8.38: Uncertainty ranges of cost per DALY saved and technical efficiency of three
projects ........................................................................................................................................255
Table 8.39: "Best buy" health interventions in South Asia ....................................................262
Table 9.1: Illustrating DEA: efficiency measurement using linear programming ..............269
Table 9.2: DEA report of illustrative case using Microsoft Excel Solver.................................270
Table 9.3: DEA result: Economy of procurement of goods and equipment ..................274
Table 9.4: Stakeholders’ views on economy of procurement of three projects ..............275
Table 9.5: Unit costs of local staff and international technical assistance inputs ............276
Table 9.6: DEA results for economy of procuring local staff and international technical assistance inputs ..........................................................................................................................277
Table 9.7: Performance indicators and achievement of targets by HSSP-SP ..................278
Table 9.8: Performance indicators and achievement of targets by NAP .......................279
Table 9.9: Attributable vaccination coverage and achievement of targets by IPL ............281
Table 9.10: Extent of health system outputs achieved by each project .......................281
Table 9.11: Use of allocated resources by three projects .................................................282
Table 9.12: DEA result for performance efficiency of three projects .............................283
Table 9.13: Stakeholders’ views on efficiency of productivity measure of three projects ..........................................................................................................................283
Table 9.14: Stakeholders’ views on health system related outcomes of three projects 284
Table 9.15: Health system strengthening investments and health system outcome points ..........................................................................................................................285
Table 9.16: DEA results for effectiveness of three projects ..........................................285
Table 9.17: Stakeholders’ views on project effectiveness ...............................................286
Table 9.18: Summary of DEA results for three projects ...............................................287
Table 9.19: Variables considered for correlations analysis ..........................................289
Table 9.20: Spearman correlation coefficients for the Paris principles and project performance ..........................................................................................................................290
Table 9.21: Pearson correlation coefficients for the Paris principles and project performance ..........................................................................................................................293
Table 10.1: Comparison of cost effectiveness and technical efficiency of three projects ..........................................................................................................................307
Table 10.2: Values considered for correlations analysis between compliance with the Paris principles and technical efficiency of cost per DALY saved .................311
Table 11.1: Summary of cost effectiveness analysis .......................................................320
Table 11.2: Summary of results from DEA .................................................................320
5 List of figures

Figure 2.1: The aid effectiveness pyramid ................................................................. 19
Figure 3.1: Map of Timor-Leste .................................................................................. 56
Figure 3.2: Timor-Leste's annual budget and sources of financing 2008-2014 (in Million US Dollars) .............................................................................................................. 62
Figure 4.1: Aid instruments in fragile situations ............................................................ 74
Figure 5.1: A comprehensive view of aid effectiveness and development effectiveness 98
Figure 5.2: Corporate scorecard model using results-chain ......................................... 100
Figure 5.3: WHO health systems framework .................................................................. 102
Figure 5.4: Theoretical framework for aid effectiveness evaluation ................................ 104
Figure 5.5: Illustrating the triangulation technique ........................................................ 124
Figure 6.1: Timor-Leste's scores on Fragile State Index 2007-2013 .............................. 133
Figure 6.2: Health systems assessment framework ....................................................... 144
Figure 6.3: Organisation of Timor-Leste's health system governance .............................. 146
Figure 6.4: State budget by cost category for MoH of Timor-Leste (in thousand US dollars) ................................................................................................................................. 154
Figure 6.5: Total health expenditure, government and donor contribution to health financing in Timor-Leste 2008-2013 .................................................................................................................. 155
Figure 6.6: Proportion of government contribution to health budget in Timor-Leste 2008-2013 ......................................................................................................................................... 156
Figure 7.1: Compliance with the Paris principles by three projects ................................. 178
Figure 7.2: Perceived economy of three interventions .................................................... 181
Figure 7.3: Perceived efficiency of three interventions .................................................. 184
Figure 7.4: Perceived effectiveness of three interventions .............................................. 188
Figure 7.5: Perceived sustainability of three interventions ............................................. 190
Figure 7.6: Overall scores of three interventions on individual domains ....................... 192
Figure 7.7: QuODA assessment of the World Bank, Global Fund, and US Government aid ..................................................................................................................................... 201
Figure 8.1: DALYs saved by HSSP-SP against pessimistic and optimistic target scenarios ..................................................................................................................................... 215
Figure 8.2: Life years gained by PMTCT and ART programme components in Timor-Leste, 2007-2014 .......................................................................................................................... 233
Figure 8.3: Number of new HIV infections in counterfactual, actual, and optimum scenario ........................................................................................................................................................................234
Figure 8.4: Number of AIDS related deaths in counterfactual, actual, and optimum scenario ........................................................................................................................................................................................................235
Figure 8.5: Vaccination coverage targets and achievements in IPL-focus and IPL-non-focus districts in Timor-Leste, 2010-2013 ........................................................................................................................................................................................245
Figure 9.1: Graphical presentation of comparative performance. Source: Author's drawing ........................................................................................................................................................................................................268
Figure 9.2: Logic model of the Paris principles and improved health outcomes ........299
Figure 11.1: Stakeholders’ views on three projects ..........................................................319
### 6 Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>AEM</td>
<td>Asian Epidemic Model</td>
</tr>
<tr>
<td>AIDS</td>
<td>Acquired Immune Deficiency Syndrome</td>
</tr>
<tr>
<td>ART</td>
<td>Anti Retroviral Treatment</td>
</tr>
<tr>
<td>ASEAN</td>
<td>Association of South East Asian Nations</td>
</tr>
<tr>
<td>AUD</td>
<td>Australian Dollar</td>
</tr>
<tr>
<td>AusAID</td>
<td>Australian Agency for International Development</td>
</tr>
<tr>
<td>BCC</td>
<td>Behaviour Change Communication</td>
</tr>
<tr>
<td>BSP</td>
<td>Basic Service Package</td>
</tr>
<tr>
<td>BSS</td>
<td>Behaviour Surveillance Survey</td>
</tr>
<tr>
<td>CAVR</td>
<td>Commission for Reception, Truth and Reconciliation (Portuguese acronym)</td>
</tr>
<tr>
<td>CCM</td>
<td>Country Coordinating Mechanism</td>
</tr>
<tr>
<td>CEA</td>
<td>Cost Effectiveness Analysis</td>
</tr>
<tr>
<td>CER</td>
<td>cost effectiveness ratio</td>
</tr>
<tr>
<td>CHC</td>
<td>Community Health Centre</td>
</tr>
<tr>
<td>CIA</td>
<td>Central Intelligence Agency</td>
</tr>
<tr>
<td>CPIA</td>
<td>Country Policy and Institutional Assessment</td>
</tr>
<tr>
<td>CPLP</td>
<td>Community of Portuguese-Speaking Countries</td>
</tr>
<tr>
<td>CSW</td>
<td>Client of Sex Worker</td>
</tr>
<tr>
<td>DAC</td>
<td>Development Assistance Committee</td>
</tr>
<tr>
<td>DALY</td>
<td>Disability Adjusted Life Year</td>
</tr>
<tr>
<td>DEA</td>
<td>Data Envelopment Analysis</td>
</tr>
<tr>
<td>DFID</td>
<td>Department for International Development</td>
</tr>
<tr>
<td>DMU</td>
<td>Decision Making Unit</td>
</tr>
<tr>
<td>DP</td>
<td>Development Partner</td>
</tr>
<tr>
<td>DPMU</td>
<td>Development Partnership Management Unit</td>
</tr>
<tr>
<td>DPT</td>
<td>Diphtheria, Pertussis, Tetanus</td>
</tr>
<tr>
<td>DW</td>
<td>Disability Weight</td>
</tr>
<tr>
<td>EC</td>
<td>European Community</td>
</tr>
<tr>
<td>EPI</td>
<td>Expanded Programme on Immunisation</td>
</tr>
<tr>
<td>FSW</td>
<td>Female Sex Worker</td>
</tr>
<tr>
<td>GAVI</td>
<td>Global Vaccine Alliance</td>
</tr>
<tr>
<td>GBD</td>
<td>Global Burden of Diseases and Injury</td>
</tr>
<tr>
<td>GDP</td>
<td>Gross Domestic Product</td>
</tr>
<tr>
<td>Global Fund/GF</td>
<td>The Global Fund to fight AIDS, Tuberculosis and Malaria</td>
</tr>
<tr>
<td>GGE</td>
<td>General Government Expenditure</td>
</tr>
<tr>
<td>GHE</td>
<td>Global Health Estimates</td>
</tr>
<tr>
<td>GHP</td>
<td>Global Health Partnership</td>
</tr>
<tr>
<td>GNI PPP</td>
<td>Gross National Income Purchasing Power Parity</td>
</tr>
<tr>
<td>HDI</td>
<td>Human Development Index</td>
</tr>
<tr>
<td>HEMI</td>
<td>HIV Epidemiological Modeling and Impact</td>
</tr>
<tr>
<td>HIV</td>
<td>Human Immune Deficiency Virus</td>
</tr>
<tr>
<td>HMIS</td>
<td>Health Management Information System</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>HMIS-SE</td>
<td>Health Management Information System and Surveillance</td>
</tr>
<tr>
<td>HSS</td>
<td>Health Systems Strengthening</td>
</tr>
<tr>
<td>HSSP-SP</td>
<td>Health Sector Strategic Plan Support Project</td>
</tr>
<tr>
<td>IBBS</td>
<td>Integrated Biological and Behavioural Survey</td>
</tr>
<tr>
<td>ICER</td>
<td>Incremental Cost Effectiveness Ratio</td>
</tr>
<tr>
<td>IDA</td>
<td>International Development Association</td>
</tr>
<tr>
<td>IDP</td>
<td>Internally Displaced People</td>
</tr>
<tr>
<td>IDU</td>
<td>Injection Drug User</td>
</tr>
<tr>
<td>IHP+</td>
<td>International Health Partnerships</td>
</tr>
<tr>
<td>IPL</td>
<td>Immunizasun Proteje Labarik (Children Immunisation Project)</td>
</tr>
<tr>
<td>ISF</td>
<td>International Stabilizing Force</td>
</tr>
<tr>
<td>JMP</td>
<td>Joint Monitoring Programme</td>
</tr>
<tr>
<td>JSI</td>
<td>John Snow Inc.</td>
</tr>
<tr>
<td>LFA</td>
<td>Local Fund Agent</td>
</tr>
<tr>
<td>M&amp;E</td>
<td>Monitoring and Evaluation</td>
</tr>
<tr>
<td>MCC</td>
<td>Millennium Challenge Corporation</td>
</tr>
<tr>
<td>MCHIP</td>
<td>Mother and Child Health Integrated Programme</td>
</tr>
<tr>
<td>MDG</td>
<td>Millennium Development Goal</td>
</tr>
<tr>
<td>MDTF</td>
<td>Multi-Donor Trust Fund</td>
</tr>
<tr>
<td>MNCH</td>
<td>Maternal, Neonatal and Child Health</td>
</tr>
<tr>
<td>MoH</td>
<td>Ministry of Health</td>
</tr>
<tr>
<td>MoT</td>
<td>Modes of Transmission</td>
</tr>
<tr>
<td>MSM</td>
<td>Men who have Sex with Men</td>
</tr>
<tr>
<td>MTEF</td>
<td>Mid Term Expenditure Framework</td>
</tr>
<tr>
<td>NAP</td>
<td>National AIDS Programme (National HIV/AIDS and STI Control Programme)</td>
</tr>
<tr>
<td>NDAE</td>
<td>National Directorate of Aid Effectiveness</td>
</tr>
<tr>
<td>NGO</td>
<td>Non Governmental Organisation</td>
</tr>
<tr>
<td>NICE</td>
<td>(UK) National Institute of Health and Care Excellence</td>
</tr>
<tr>
<td>NPV</td>
<td>Net Present Value</td>
</tr>
<tr>
<td>ODA</td>
<td>Official Development Assistance</td>
</tr>
<tr>
<td>OECD</td>
<td>Organization for Economic Cooperation and Development</td>
</tr>
<tr>
<td>PDSS</td>
<td>Committee of the Suco Council (portuguese acronym)</td>
</tr>
<tr>
<td>PEPFAR</td>
<td>President’s Emergency Plan For AIDS Relief</td>
</tr>
<tr>
<td>PEST</td>
<td>Political, Economic, Social, and Technological</td>
</tr>
<tr>
<td>PFM</td>
<td>Public Financial Management</td>
</tr>
<tr>
<td>PIU</td>
<td>Project Implementation Unit</td>
</tr>
<tr>
<td>PLHIV</td>
<td>People Living with HIV/AIDS</td>
</tr>
<tr>
<td>PMP</td>
<td>Project Monitoring Plan</td>
</tr>
<tr>
<td>PMTCT</td>
<td>Prevention of Mother to Child Transmission</td>
</tr>
<tr>
<td>PMU</td>
<td>Project Management Unit</td>
</tr>
<tr>
<td>PSF</td>
<td>Community Health Worker (portuguese acronym)</td>
</tr>
<tr>
<td>PSG</td>
<td>Peace and State-building Goal</td>
</tr>
<tr>
<td>QALY</td>
<td>Quality Adjusted Life Years</td>
</tr>
<tr>
<td>QuODA</td>
<td>Quality of Official Development Assistance</td>
</tr>
<tr>
<td>RCT</td>
<td>Randomized Controlled Trial</td>
</tr>
</tbody>
</table>
SAMES  Serviço Autónomo de Medicamentos e Equipamentos de Saúde
       (Central Medical Store Warehouse)
SDA    Service Delivery Area
SDP    Sector Development Plan
SFA    Stochastic Frontier Analysis
SISCa  Servisu Integrado du Saude Comunidade (Integrated Community Health Services)
STD    Sexually Transmitted Disease
STI    Sexually Transmitted Infection
SWAp   Sector Wide Approach
SWOT   Strength, Weakness, Opportunities, Threats
TA     Technical Assistance
TAIS   Timor-Leste Asisténsia Integraddu Saudé
TB     Tuberculosis
UN     United Nations
UNAET  United Nation’s Administration for East Timor
UNAIDS Joint United Nations Programme on HIV/AIDS
UNDP   United Nations Development Programme
UNESCO United Nations Educational, Scientific and Cultural Organization
UNFPA  United Nations Population Fund
UNICEF United Nations Children’s Fund
UNMIT  United Nations Integrated Mission in Timor-Leste
US     United States
USA    United States of America
USAID  United States Agency for International Development
USD    United States Dollar
VCT    Voluntary Counseling and Testing
WHO    World Health Organization
YLD    Year Lived in Disability
YLL    Year of Life Lost
7 List of annexes

Annex 1: Questionnaire for stakeholder interview
Annex 2: Informed consent form
Annex 3: Ethical approval from Timor-Leste
Annex 4: Ethical approvals from the University of Otago
Annex 5: List of project documents used as data sources
1 Introduction

1.1 Background

As an international health and development professional I provided technical support to the Ministry of Health (MoH) of Timor-Leste from mid 2007 to early 2012 with respect to managing the National HIV/AIDS, TB and Malaria programmes of the country funded by the Global Fund. For Timor-Leste, this was a period of reconstruction and transition towards development from the second widespread conflict that the country experienced in 2006, after having the first one in 1999.

In my professional role I had to work closely with the government and donors and interacted with NGOs, civil society and various bilateral and multilateral development partners. In a challenging working environment in a fragile situation, I came across multiple capacity building approaches, aid instruments and health aid management modalities. The range of aid management approaches was quite diverse—sometimes complementary to each other but often disjointed and sometimes even conflicting with each other. With insufficient information, a volatile situation, and seemingly mismatches between donor requirements and the country’s assessment of priorities, we, the international advisers, often wondered what would be the best approach to support the country in achieving its development objectives.

Timor-Leste (also known as East Timor), a post-conflict country of around fifteen thousand square kilometers, with just over one million population, is an important candidate for a global health aid effectiveness study. The country emerged as the world’s youngest nation in 2002 after several hundred years of Portuguese occupation until 1975, almost a quarter century of occupation by Indonesia until 1999, and UN Interim Administration until 2002: “Violence had left the country and families torn apart, nearly 70 per cent of all buildings, homes and schools destroyed, and an estimated 75 per cent of the population was displaced” (World Bank, 2012d).
Timor-Leste received much attention from the international development community. In the health sector alone, there were 16 development partners active with a variety of project support approaches (AusAID, World Bank, & Ministry of Health of Timor-Leste, 2011). In this ‘donor-heavy’ environment, the government of Timor-Leste was trying to strengthen its leadership over aid coordination. The government chaired annual Development Partners Meetings for Timor-Leste, and created an Aid Effectiveness Directorate within the Ministry of Finance. Timor-Leste also volunteered to pilot the OECD-DAC Principles for Good International Engagement in Fragile States and Situations (OECD, 2007d) and undertook periodic reviews of its implementation (OECD, 2010).

In fact, this was a complex environment in post-conflict transition, with associated peace-building and state-building goals, contesting development priorities, multiple aid management modalities, and prevailing challenges from the growing institutions and health system's capacity. It occurred to me that the presence of different aid modalities within the same context of transition and fragility in Timor-Leste could be an important subject for an in-depth analysis of the role of institutional arrangements and health systems strengthening in the effectiveness of health aid interventions.

1.2 Rationale

For many years researchers tried to understand aid effectiveness in terms of the effect of aid on national economic growth and saw mixed results (Burnside & Dollar, 2000; Chauvet & Collier, 2004; Collier, 2007; Dalgaard & Hansen, 2001; Dalgaard, Hansen, & Tarp, 2004; Doucouliagos & Paldam, 2009; Easterly, 2006; Fielding & Knowles, 2007; Howes, Otor, & Rogers, 2011; Mosley, 1986; Rawlings & Rubio, 2005; Tarp, 2010). Such studies often lack a particular focus of aid effectiveness in the health sector and are influenced either by the effect of aggregation of aid in all sectors or by the generalization of different country contexts in cross-country regression models. It was increasingly felt that the broader notion of human development cannot be solely explained by the economic growth and studying aid effectiveness in the health sector requires a new approach (Victora et al., 2011).
With new ideas on aid effectiveness, new policy directions, and an attempt at “reversal of negatives” to alter what had not worked so far (Stern et al., 2008), aid donors, recipient countries and civil society came to a consensus in 2005 in the form of the Paris Declaration setting the principles of country ownership, harmonization, alignment, mutual accountability and managing for development results (Paris Declaration, 2005). These principles of the Paris Declaration are popularly known as the Paris principles of aid effectiveness.

Although the Paris Declaration advocated for a set of processes and their applicability in fragile and conflict affected situations (OPM/IDL, 2008), actual effectiveness of adhering to these principles in fragile situations in terms of achieving intended health outcomes and development results have not been adequately studied.

In a fragile situation, development partners are often divided into a ‘relief/development’ dichotomy in their aid modalities ranging from state-avoiding mechanism such as contracting NGOs for service delivery to state building mechanisms such as a sector-wide approach under the Government’s leadership. There are other approaches such as global health partnerships that fall in-between the state-avoidance and state building ends of the continuum (Vergeer, Canavan, & Rothmann, 2009).

In Timor-Leste’s health sector, different aid modalities were working at the same time. At one extreme, for example, the Millennium Challenge Fund (MCC) and USAID supported projects tended to avoid state mechanisms and tried to implement interventions through NGOs (MCC, 2010). At the other extreme there was an AusAID and World Bank supported health sector wide approach (SWAp) that tried to work through the government to strengthen implementation of a health sector strategic plan (World Bank, 2007a). A third modality evolved for implementing the Global Fund grants for the national HIV/AIDS, TB and Malaria programmes by channeling the funds through the government but by creating a “facilitation” unit within the MoH to enable implementation of these ‘vertical’ programmes (Chan et al., 2013).

These ‘state-avoiding’, ‘state-building’ and 'state-partnership’ mechanisms in Timor-Leste’s health sector created an opportunity for piloting an approach to evaluate the
comparative effectiveness of each modality in terms of compliance with the Paris principles and actual achievement of health and development results.

As presented in Chapter 2, a systematic review of literature related to the evaluation of the Paris principles in the health sector shows that most such studies focused on the process and progress of implementation of the Paris principles in the health sector. Although a few studies attempted to draw a link between the effect of implementation of the Paris principles and the health results (Dalil et al., 2014; Dickinson, 2011; OECD Working Party on Aid Effectiveness Task Team on Health as a Tracer Sector, 2011; Paul et al., 2013; Wood & Betts, 2013), they mostly focused on the effect of the implementation of the Paris principles at the sectoral level. The shortage of research on compliance with the Paris principles and programmatic outcomes at the intervention level encouraged me to explore the feasibility of such a study in Timor-Leste's health sector.

**1.3 Research question**

With this backdrop, I was primarily interested in developing and piloting a comparative evaluation approach that seeks answers to the following key questions:

- In a fragile situation what management modality of external funding in the health sector is more effective in terms of achieving programmatic objectives and strengthening the country’s health systems?

- What management modality of external funding in the health sector is more compliant with the Paris Principles of aid effectiveness and does compliance have any effect on the achievement of programmatic objectives and strengthening of the country's health systems?

More specifically, I was interested in testing if we can combine and use some analytical techniques such as balanced scorecard, cost effectiveness analysis, data envelopment analysis, and correlation analysis to evaluate some aid funded projects in answering these questions. However, as described in Chapter 5 (Research Methods), the list of questions were further refined and specified as the research unfolded.
1.4 Overview of research process

It appeared from the review of literature that both the health systems in a fragility context and the aid mechanisms contribute to the outcomes of an aid intervention. Following the tradition of “realist evaluation” I was interested in investigating “what works for whom under which conditions” (Marchal et al., 2012). The underlying assumption of a realist evaluation is-- in a given context, different mechanisms (in this case aid management modalities) would produce different outcomes, which can be fairly compared with each other to see what worked, why they worked, under which conditions they produced such outcomes and for whom (Pawson & Tilley, 2004).

Based on this assumption I selected three externally funded projects from Timor-Leste’s health sector representing three different aid modalities for comparative evaluation. These three interventions were:

• The Health Sector Strategic Plan Support Project (HSSP-SP) funded by AusAID and World Bank that used a ‘sector-wide’ approach with funding channelled through the government (World Bank, 2007g);
• The Global Fund funded National HIV/AIDS and STI Programme (NAP) that channelled the funding through the MoH but not fully using the government systems (Timor-Leste Country Coordinating Mechanism for the Global Fund, 2010); and
• The Children Immunization Project or ‘Immunizasaun Proteje Labarik’ (IPL) funded by USAID with funding channelled outside the government (MCC, 2010).

The HSSP-SP was implemented from June 2008 to December 2012 with a total budget of $21 million US dollars. The NAP was also implemented from June 2007 to December 2013 with a total budget of $14.4 million US dollars, whereas, IPL was implemented from April 2011 to September 2013 with a total budget of $2.6 million US dollars.

The reasons for selecting these three specific interventions were—they represented three different aid management modalities in terms of the country government’s involvement; they all operated within the same context of Timor-Leste's health sector; and they were more or less implemented during the same period. As these projects shared the same context, I believed their outcomes could be fairly compared to their
management modalities and inputs without the need for further controlling the context. This was also influenced by the fact that, from my experience of working in the health sector of Timor-Leste, I knew where I could find required data for evaluating these projects and I was confident that I could access the data through my professional network.

I left Timor-Leste and started this research project in April 2012 at the University of Otago, New Zealand with occasional visits and fieldwork in Timor-Leste. After finalizing the detailed proposal, methodology and work-plan and obtaining research advisory approval from the University of Otago, I conducted the first field-visit to Timor-Leste in July 2012. During this visit I obtained full approval for my research from the MoH’s Technical and Ethical Committee for Health Research in Timor-Leste and collected necessary background information related to a formative assessment of three projects.

As will be explained in Chapter 5, an evaluation framework was developed by combining the different aspects of intervention logic models, the Paris principles of aid effectiveness, and health systems thinking to see the interactions and pathway between the Paris principles, project design, inputs, outputs, and outcomes of each project to allow a comparative analysis.

After receiving a full approval on the research design from the University of Otago’s Research Ethics Advisory Committee, I conducted an extensive fieldwork in Timor-Leste from April to November 2013. In this process, I obtained further approval from the Minister of Health of Timor-Leste to access necessary data and to interview MoH officials.

During the fieldwork I visited all the projects under study, talked to the project implementation teams, collected relevant background information, and accessed relevant project data as listed in Annex 5. Based on a structured questionnaire I interviewed 22 key informants from various levels of stakeholders including project personnel, MoH officials, and donors. I also conducted three focus group discussions with the three project implementation teams.
In order to be able to capture the diverse issues of fragility and health system context, aid instruments, and project results, this research employed a mixed methods approach combining both qualitative and quantitative techniques. The qualitative techniques included contextual, fragility, and health systems analysis, and taking stakeholder interviews. The quantitative techniques included balanced scorecards, cost effectiveness analysis, data envelopment analysis (DEA), and statistical analysis for correlations.

The underlying programme theory and empirical evidence from both qualitative and quantitative analysis allowed for triangulation of research results for increased validity, reliability, and robustness of interpretation (Ostlund et al., 2011).

1.5 **Key definitions**

The key concepts used in this research are explained in their respective chapters where they appear. However, a few definitions need to be clarified from the onset.

**Project, programme, and intervention:**

Although programme usually refers to a broader concept with a set of ongoing planned activities contributing to higher-level objectives than the time-bound activities and objectives of a project or an intervention (Cordella & Dell’Ariccia, 2007), this research used all three terms interchangeably. While interventions are designed under a time and objective bound project, they actually contribute to a national level programme.

**Fragile state and situation:**

The terms fragile state, failed state, and fragile situation often refer to the weak capacity, unresponsiveness, or lack of legitimacy of a state to provide services to their people (DFID, 2005). However, recognizing the possible sensitivity associated with labelling a country with a negatively perceived term, this research uses the term ‘fragile situation’ to refer to the post-conflict transitional situation in Timor-Leste.

**Health sector and health systems strengthening:**

This research uses the terms health sector and health systems strengthening as defined by the World Health Organization (WHO). According to WHO, “a health system consists of all organizations, people and actions whose primary intent is to promote, restore or
maintain health” (WHO, 2007, p. 2). The health sector includes all players including government ministries and departments, hospitals and other health service providers, health insurance schemes, voluntary and private organizations in health, as well as the pharmaceutical industry and drug wholesale companies.

Health systems strengthening is defined as improving the six health system building blocks such as health service delivery, health work-force, health information, medical products, vaccines and technologies, health financing, and health governance and leadership, and managing their interactions in ways that achieve more equitable and sustained improvements across health services and health outcomes (Savigny & Adams, 2009).

**Health systems thinking:**
System thinking is a problem solving approach that views problems as part of a wider, dynamic system. A system is composed of parts working together in a dynamic interactive way to produce the end result. In this research, health systems thinking is referred to see how the health system interacts with an intervention as well as to the external factors to produce the outcomes (Adams & Savigny, 2012).

**Aid:**
Aid refers to any form of financial assistance. However, in this research the term aid is used in the sense of overseas development assistance (ODA) as defined by OECD-DAC. According to OECD-DAC (2008c), ODA has three elements: i) it is provided by an official agency including state or local government or their executive agency; ii) the primary objective of such assistance is economic development or welfare; and iii) there is a concessional or grant element in such assistance.

**Aid effectiveness and development effectiveness:**
The term aid effectiveness is often taken as synonymous to the Paris principles of aid effectiveness, which focus on the aid management principles and processes such as country ownership, alignment, harmonization, managing for results, and mutual accountability (OECD, 2015). According to OECD, “Aid effectiveness is about value for money – it is about managing aid in a way that maximizes its impact on development”
(Working Party on Aid Effectiveness, 2009, p. 1). However, by the term aid effectiveness, AusAID Office of Development Effectiveness (2008) focuses on the results by defining effectiveness as the extent to which a development intervention’s objectives were achieved, or are expected to be achieved.

To distinguish the process from the results, this research uses the term “the Paris principles” to mean the Paris principles of aid effectiveness and uses the terms “programmatic effectiveness” or “development effectiveness” to mean the effectiveness of programmatic outcomes in achieving their objectives.

**Budget Support:**
As defined by OECD/DAC (2006), budget support is a method of financing a recipient country’s budget by the aid donors through a transfer of resources from an external financing agency to the recipient country’s national treasury. The funds thus transferred are managed in accordance with the recipient country’s budgetary procedures.

**Result, output, outcome, and impact:**
Outputs are immediate results of an activity, which can be measured by a process indicator. Outcomes are long-term results from the use of the outputs contributing towards achievement of an objective. Impact, on the other hand, is the longer-term result towards achievement of a programme’s goals (OECD-DAC, 2002). The term “result” denotes the effect or an intervention and is inclusive of outputs, outcomes and impact. This research considers the results by considering outcomes and impact together and these terms are used interchangeably.

**Impact pathway:**
An impact pathway illustrates a programme logic or theory of change behind the planned activities of an intervention. Such logic tries to understand how the allocated resources or inputs are used in organizing an activity and how the outputs or results of that activity contribute to the stated objectives and goal of that intervention. A health system’s impact pathway also tries to understand how the intervention or an activity contributes to the changes and interactions of the health system’s building blocks (Savigny & Adams, 2009).
1.6 How this thesis is organized

This thesis is organized in ten chapters.

**Chapter 1: Introduction**—Chapter 1 provides a broad context and background, rationale for the study, key research questions, overview of research process, clarifications of key definition, structure, and contribution of this thesis.

**Chapter 2: Review of related literature**—Chapter 2 provides a review of related literature on aid effectiveness, fragile situation, health systems strengthening, aid instruments, programme evaluation, and the qualitative and quantitative techniques used in this research. Through a mapping of the changing concept of aid effectiveness, systematic review of literature related to the evaluation of the Paris principles in the health sector, and scoping of related concepts, this chapter attempts to establish the theoretical framework, concepts, methods and tools used in this research.

**Chapter 3: Country context**—In order to understand how the selected interventions interacted with their context, Chapter 3 presents an analysis of the basic country profile including the geographic, demographic, cultural, historical, economic, and political contexts as well as the country’s performance in relation with the achievement of Millennium Development Goals (MDGs).

**Chapter 4: Project backgrounds**—Chapter 4 provides brief background, organization, objectives and performance framework of the three projects selected for evaluation. This chapter attempts to evaluate the mechanisms, relevance and appropriateness of the selected interventions in the context of country’s background.

**Chapter 5: Evaluation framework and research methods**—Chapter 5 presents and explains the overall theoretical framework used for comparative evaluation of the three different aid management modalities in Timor-Leste’s health sector. The framework draws on the Paris principles of aid effectiveness, intervention logic model, and health systems thinking. This chapter also provides an overview of the mixed methods and
analytical tools used in this research including balanced scorecard, cost-effectiveness analysis, and DEA.

Chapter 6: State fragility and health systems: a case of Timor-Leste—Chapter 6 provides an analysis of Timor-Leste’s health systems in the context of state fragility to assess their possible effects on the outcomes of the three projects selected for evaluation. This chapter is organized in three parts-- Part A provides a fragility analysis by using the findings from Timor-Leste’s self assessment of fragility, changes in the Fragile State Index for Timor-Leste over the period from 2007 to 2013 and by identifying different factors of fragility remaining in the country’s situation. Part B provides a health system analysis in the context of fragility by using WHO’s health systems building block concepts. Part C employs some planning and management tools such as strengths, weaknesses, opportunities, and threats (SWOT) analysis and risk analysis and attempts to synthesize the key considerations from the fragility and health system analysis to identify their possible implications for the health intervention outcomes.

Chapter 7: Stakeholder perspectives: balanced scorecards—Chapter 7 presents the findings from the key informant interviews on various aspects of the three projects such as compliance with the Paris principles, economy, efficiency, effectiveness, and likely sustainability. This chapter presents the findings from respondent views in the form of Balanced Scorecards and comparison graphs and validates the findings by comparing them with findings from other sources.

Chapter 8: Cost-effectiveness and technical efficiency—Chapter 8 presents a cost-effectiveness analysis for the three projects in terms of disability adjusted life year (DALY) saved and cost per DALY saved by each project by using the DALY calculation method used in the Global Health Estimates (WHO, 2013a). For calculating DALY saved by each intervention, three scenarios were considered: a counterfactual scenario of respective disease burden (new infections and deaths) without the intervention, the intervention results scenario of actual disease burden with the intervention, and a possible optimum scenario of disease burden if the intervention had achieved all the targets fully. This chapter provides detailed findings from the process of calculations.
and projections for different scenarios by using the Spectrum (Futures Institute, 2014a) and Modes of Transmission (UNAIDS, 2012) software and mathematical formulas. By comparing the actual cost per DALY saved by the intervention with the likely cost per DALY saved in the optimum scenario, this chapter also calculates the technical efficiency of each project in terms of the cost per DALY saved.

**Chapter 9: The Paris principles and comparative economy, efficiency, and effectiveness**—Chapter 9 employs a non-parametric linear programming (DEA) to compare the relative economy, efficiency, and effectiveness of each project in comparison to each other. These findings are also compared to the findings from the stakeholder interviews in Chapter 7 for further analysis and triangulation of findings. This chapter then employs a statistical analysis to see if the degree of compliance with the Paris principles by three projects were correlated with any aspects of their performance.

**Chapter 10: Lessons learned**—Chapter 10 captures the key lessons learned from the series of pilot studies in critically evaluating the evaluation approach used in this research. This chapter provides an overview of the analytical tools used in this research, their strengths, limitations, data requirements, and how the use of the tools could be further improved to produce meaning results.

**Chapter 11: Summative discussion and conclusion**—Chapter 11 attempts to summarize, triangulate, synthesize and discuss the findings from all other chapters. This chapter also discusses robustness, reliability, validity, limitations, recommendations, and implications for further research.

**1.7 Expected contribution of this thesis**

In line with continuous efforts to improve the outcomes of development assistance, internationally agreed principles such as the Paris principles of aid effectiveness present ‘theory’ that following them will result in development effectiveness. However, the theory of aid effectiveness in relation to the Paris principles is yet to be proved. Macro-economic studies on the effect of aid on growth often provide a "spotlight effect" with focus on growth in an "input-black box-output" model and miss the critical part of what
was happening within the "black box" of health systems (Grap-Pa Sante, 2011). The limited number of studies that have been conducted so far evaluating the Paris principles, on the other hand, attempted to synthesize the findings from a number of case studies in different countries. The country context in these generalized studies was often missing. Although a few studies commented on a pathway between the implementation of the Paris principles and better health results, they could not establish a causal relationship between these two variables at the intervention level.

Employing a systems theory, intervention logic models, contextual analysis and various qualitative and quantitative techniques through a theory-driven ‘realist evaluation’ approach, this research specifically attempted to see what was happening in a single country context and if compliance with the Paris principles had any effect on the programmatic outcomes. This research also attempted to pilot the evaluation approach for the feasibility of finding any significant correlation between compliance with the Paris principles and outcomes of an intervention.

Since contexts are always dynamic and no two contexts are identical, this in-depth study of different aid management modalities and processes within a single fragile state context provides a useful framework for analysis and exposes useful findings for policy makers and researchers. This approach also creates opportunities for further research on health aid effectiveness studies with other samples and in other contexts. This can help improve the understanding of global public health policy for more effective policymaking and management of development cooperation, especially in fragile situations.
2 Review of related literature

2.1 Introduction

As introduced in Chapter 1, this thesis attempts a comparative evaluation of the extent of compliance with the Paris principles of aid effectiveness (2005) by three different aid management modalities in Timor-Leste’s health sector and their intervention outcomes. Accordingly the research draws upon a range of diverse concepts and tools from multiple disciplines including development economics, programme evaluation, epidemiology, health economics, management, and operational research. The purpose of this chapter is three-fold. First, it provides a mapping of the changing discourse of aid effectiveness to introduce the Paris principles within the context of evolving research and practice of international development assistance. Second, it presents a systematic review of the literature related to the evaluation of the Paris principles. And, third, it offers a scoping study for this research by highlighting the relevant concepts and additional works cited in different chapters of this thesis.

Accordingly, this chapter is organized in three sections: 1) changing concept of aid effectiveness; 2) systematic review of Paris declaration evaluations; and 3) related concepts and tools employed in this thesis. Through the mapping of the changing discourse, a systematic review of literature related to the key research question, and scoping of the thesis, this chapter attempts to capture and synthesize the diverse concepts employed in this thesis to see how they provide the theoretical background and support the conceptual framework of this research.

2.2 Methods

As suggested by Denyer and Tranfield (2009), the literature review in this chapter employed a ‘divergent/convergent approach’ to be able to capture and balance the breadth and depth of the research. This approach involved three steps of identifying and reviewing related literature: mapping of the broader context, systematic review of the key research question, and then scoping of the research methods.
The mapping of the context, as presented in Section 1 of this chapter, involved a broad scan and critical review of the key issues and changing concepts in the discourse of aid effectiveness. This section introduces the Paris principles of aid effectiveness and places the research question of this thesis in its broader theoretical context.

The systematic review of related literature, as presented in Section 2, involved a protocol driven search strategy focusing on the evaluation of the Paris Declaration in the health sector. Four databases such as PubMed, Ovid, Web of Science, and Academic Search Complete were searched for this purpose using the key words and additional criteria described in Table 2.1:

Table 2.1: Search and primary screening of related literature

<table>
<thead>
<tr>
<th>Database</th>
<th>Key words searched</th>
<th>Additional criteria</th>
<th>Returned results</th>
<th>Not relevant</th>
<th>Relevant articles</th>
<th>Already included</th>
<th>Added</th>
</tr>
</thead>
<tbody>
<tr>
<td>PubMed</td>
<td>Paris declaration evaluation</td>
<td>Last 10 years</td>
<td>27</td>
<td>Global burden of disease studies: 2; Bio-medical research: 8; Research ethics: 10</td>
<td>7</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>PubMed</td>
<td>Paris principles evaluation</td>
<td>Last 10 years</td>
<td>79</td>
<td>Bio-medical research: 34; Research methods/ethics: 33; Animal health: 1; Medical education: 2; GLP principles: 1</td>
<td>8</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Ovid</td>
<td>Paris declaration evaluation</td>
<td>All resources</td>
<td>0</td>
<td>ACADEMIC PRACTICE &amp; ETHICS: 1; BAROLO: 6; BURNS: 3; BURNS: 3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Ovid</td>
<td>Paris principles evaluation</td>
<td>All resources</td>
<td>0</td>
<td>ACADEMIC PRACTICE &amp; ETHICS: 1; BAROLO: 6; BURNS: 3; BURNS: 3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Ovid</td>
<td>Paris declaration evaluation</td>
<td>All resources; 2005 to 2015</td>
<td>76</td>
<td>Related to non-health sectors: 10; Bio-medical research: 1; non-research interview: 1; GIPO: 1; duplicate: 22</td>
<td>41</td>
<td>6</td>
<td>35</td>
</tr>
<tr>
<td>Ovid</td>
<td>Paris principles evaluation</td>
<td>All resources; 2005 to 2015</td>
<td>17</td>
<td>Non-health sector: 8; Bio-medical: 1; Duplicate: 3</td>
<td>5</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Web of Science</td>
<td>Paris declaration evaluation</td>
<td>2005-2015</td>
<td>12</td>
<td>Non-health sector: 2</td>
<td>10</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Web of Science</td>
<td>Paris principles evaluation</td>
<td>2005-2015</td>
<td>17</td>
<td>Non-health sector: 8; Bio-medical: 2</td>
<td>7</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>Academic Search Complete</td>
<td>Paris declaration 'and' evaluation</td>
<td>2005-2015</td>
<td>11</td>
<td>Non-English: 2; Non-health: 1; Duplicate: 1</td>
<td>7</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>Academic Search Complete</td>
<td>Paris principles 'and' evaluation</td>
<td>2005-2015, English</td>
<td>5</td>
<td>Other sector: 1</td>
<td>4</td>
<td>4</td>
<td>0</td>
</tr>
</tbody>
</table>
As shown, each database was extensively searched several times with the key words: “Paris declaration evaluation”, “Paris principles evaluation”, “Paris declaration”, “Paris principles”, and “Paris principles ‘and’ aid effectiveness”. The search included all resources listed in the databases. However, since the Paris Declaration was adopted in 2005, some additional criteria such as publication date (last 10 years or January 2005 to March 2015), and language (English) were used to search for relevant articles.

The search results were primarily screened from the title and abstracts of the results to separate the items not relevant to the Paris declaration of aid effectiveness (2005) and to the health sector research. Duplicate entries were also identified and excluded from the list. The list of relevant articles derived from this process was then added to the sample of related literature in the EndNote reference library database for further review.

The list of articles added to the sample were then closely scrutinized with the inclusion criteria of selecting published research articles related to the health sector and dealing with one or more principles of the Paris declaration of aid effectiveness. Out of 73 primarily selected items a total of 55 articles met the inclusion criteria.

While the systematic review concentrated on the core research theme, a search through the OECD website on aid effectiveness and through the snowballing technique of looking for additional work cited in the already selected literature generated a wide range of additional literature on different aspects of the planned research. Such additional literatures are reviewed in Section 3 of this chapter for a scoping study of this research.
2.3 Changing concepts of aid effectiveness

A mapping of aid effectiveness literature shows that there has been a paradigm shift from a focus on the effect of aid on poverty alleviation towards a process evaluation for efficiency and transparency of aid management and further towards impact evaluation of aid funded programmes (Booth & Fritz, 2008; Bourguignon & Sundberg, 2007; McGillivray, 2005; Prowse, 2007; Victora et al., 2011).

This paradigm shift can be broadly discussed under three different headings: macro-economic focus on aid and growth; aid delivery principles and practices; and development impact of aid intervention.

2.3.1 Macro-economic focus of aid effectiveness

The early focus on econometric models in terms of the effect of aid on national growth and poverty alleviation raised two competing positions for and against aid in international development policy discourses. Supporters of one position argued that doubling of worldwide aid flows would send forth “mighty currents of hope” and lead to the “end of poverty” (Sachs, 2005), while others argued that “aid is not a solution, but rather a part of the poor countries’ problems” (Easterly, 2006). Those in the middle ground suggested that aid has made positive contributions but needs some changes towards more effective delivery (Kenny, 2008; Whitfield, 2009).

Although it is evident that health has direct correlation with economic growth and economic growth would reduce poverty (Sachs, 2005; Spence & Lewis, 2009), a challenge often faced by the econometric studies is the “micro-macro paradox” (Mosley, 1986). Mosley pointed out that while micro-economic data from aid funded project evaluations were showing that the majority of projects were successful, macro-economic data from regressions examining the impact of aid on economic growth were not so encouraging (Mosley, 1986). This raised the question whether this paradox is due to the ineffectiveness of aid, due to the selection biases, or due to the deception of data (Howes et al., 2011; White & Bamberger, 2008).
Research on aid effectiveness, following econometric models, tried to examine aid as a contributing factor to economic growth. The studies were often engaged in cross-country data analysis that tried to generalize findings from a wide variety of contexts (Dalgaard & Hansen, 2001; Dalgaard et al., 2004; Easterly, 2002; Fielding & Knowles, 2007). The studies lacked particular focus on the effectiveness of an aid intervention in the health sector.

Hjertholm and White (1998) identified that although a development objective is common, aid often comes with diversified motives such as “colonial relationships” (client-patron relations), foreign policy, and commerce. Therefore, the macro-economic effect of aid may be due to the cumulative impact of all different aid activities with different purposes and time, not just because of a particular aid activity in a particular period.

Another issue related to the appropriateness of macro-economic study is whether the amount of aid can create a 'critical mass' to have any influence at the macro level. Tarp (2010) shows that “...the allocation of aid from 1996 to 2005, as measured by the aid to GNI ratio in aid receiving countries, is highly skewed. Most recipient countries received aid to the order of 1.8 per cent of their GNI per year with a median of 3.2 per cent. This corresponds to a distribution of aid per capita with a mean of US$17.9 per year and a median of US$31.5” (p. 23). Tarp suggests that “With this background, modest expectations are advisable when analysing the overall impact of past aid on growth” (p. 24).

2.3.2 Aid delivery principles and practices

In order to improve the practice of aid disbursement and management, the international community of donors and recipient countries met at the 1st High Level Forum on Harmonization that took place in Rome in February 2003. The Rome Declaration adopted the principles of country ownership (partner countries), alignment (donor-partner), and harmonization (donor-donor), which were believed to be “of strategic importance to the enhancement of aid effectiveness” (OECD, 2006, p. 3). This was followed by the 2nd High Level Forum on Aid Effectiveness that took place in Paris in 2005.
With new ideas, new policy directions, and attempts at a “reversal of negatives” (Stern et al., 2008) to reverse what had not worked so far, donors, recipient countries and civil society came to a consensus at the 2nd High Level Forum with the “Paris Declaration”, which set the principles of country ownership, harmonization, alignment, managing for results, and mutual accountability with an expectation of greater effectiveness of aid (Paris Declaration, 2005). These five principles of Paris Declaration forms the “aid effectiveness pyramid” as shown in Figure 2.1:

**Figure 2.1: The aid effectiveness pyramid**

*Source: (OECD, 2007a, p. 16)*

According to the aid effectiveness pyramid, development projects focus on managing for results. However, ownership of the development project is placed on the top of the pyramid. The base of such partnership approach lies in the practice of donors harmonizing their efforts with each other by establishing common agreements, simplifying procedures, and sharing information with each other. Donors are then expected to align their procedures with recipient country systems.

At the 3rd High Level Forum on aid effectiveness, held in Accra in 2008, the principles of aid effectiveness adopted in the Paris Declaration were reconfirmed with recommendations for actions for the international community (Accra Agenda for Action, 2008).
However, none of these high level forums specifically defined ‘aid effectiveness’ and, in fact, so far there have been only a few attempts to settle on a definition. For example, Bourguignon and Sundberg (2006) tried to narrow down the concept of aid effectiveness from the broader concept of the effect of aid on poverty reduction and human development and suggested that aid effectiveness represents “what aid can buy” in the recipient economy at a given level of aid and given existing physical and institutional constraints. This clearly draws the focus of the aid effectiveness concept from the impact of aid on economic growth to the question how the effectiveness of aid can be improved through better management of aid in a given context.

The focus on aid management was reinforced when three years after the Paris Declaration, different elements of aid effectiveness principles were put together in an attempt to define aid effectiveness. The thematic review of the Paris Declaration defined aid effectiveness as an ‘arrangement for the planning, management and deployment of aid that is efficient, reduces transaction costs and is targeted towards development outcomes including poverty reduction’ (Stern et al., 2008). In the same line the Working Party on Aid Effectiveness confirmed that “Aid effectiveness is about value for money – it is about managing aid in a way that maximises its impact on development” (Working Party on Aid Effectiveness, 2009, p. 1). The AusAID Office of Development Effectiveness (2008) also noted that the effectiveness of aid is influenced by the way that it is delivered and suggested “Effectiveness can be defined simply as the extent to which a development intervention’s objectives were achieved, or are expected to be achieved”.

### 2.3.3 Development impact of aid interventions

The Fourth High Level Forum on Aid Effectiveness held in Busan from 29 November to 1 December 2011 contributed significantly to the evolution of the concept of aid effectiveness as development effectiveness. The Busan Outcome Document (4th High Level Forum on Aid Effectiveness, 2011) shifted the idea of aid effectiveness from following the set of Paris Declaration principles to the effective development cooperation between donors, recipient country governments, private sector, NGOs and the population for achieving sustainable developmental impacts.
Unlike the declarations from the previously held high level forums, the Busan Outcome Document sees development cooperation as a more inclusive and holistic process aiming for leveraging maximum impact in terms of sustainable development results and graduation from aid dependence to self-reliance. Clearly the shift is from seeing aid as a ‘buyer’ of development results to a ‘catalyst’ of effective development. As the Busan Outcome document says:

“Aid is only part of the solution to development. It is now time to broaden our focus and attention from aid effectiveness to the challenges of effective development...To this effect, we will rethink what aid should be spent on and how, in ways that are consistent with agreed international rights, norms and standards, so that aid catalyses development” (4th High Level Forum on Aid Effectiveness, 2011, p. 9).

Recognizing the catalytic role of aid in development cooperation, the Busan Outcome Document, in fact, changed the challenges of measuring and evaluating aid effectiveness from identification of “attribution” (that is, what development results can be attributable to the aid) to the analysis of “contribution” (that is, what is the contribution of aid funded activities in the overall development results).

Manning and Trzeciak-Duval (2010) noted that, “It is the underlying assumption of the Paris Declaration that improved aid effectiveness [i.e. adherence to the set of aid effectiveness principles] will in turn improve development outcomes as measured by progress towards the MDGs” (p.125). However, research has been sparse so far with respect to the comparative outcomes of adhering to the Paris principles by an aid intervention in the health sector. With this backdrop, Section 2.4 presents a systematic review of literature related to the evaluation of the use of the Paris principles in the health sector.

2.4 Systematic review of literature related to the Paris principles in the health sector

As mentioned in the methods section of this chapter, a total of 55 articles met inclusion criteria for systematic review. These 55 articles were published from 2007 to 2014 as shown in the following table:

<table>
<thead>
<tr>
<th>Year of publication</th>
<th>Number of articles</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005</td>
<td>0</td>
</tr>
<tr>
<td>2006</td>
<td>0</td>
</tr>
</tbody>
</table>
It can be seen that as the Paris principles were only adopted in 2005, articles related to the Paris principles in the health sector only started to come out from 2007 with highest number of articles of this sample published in 2011. This was probably linked to the 4th High Level Forum on Aid Effectiveness held in Busan in 2011.

A thematic review of these 55 articles, however, places them in the following broad categories:

**Table 2.3: Literature selected for systematic review**

<table>
<thead>
<tr>
<th>Category</th>
<th>Number of articles</th>
<th>Percentage of sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evaluation of the Paris declaration</td>
<td>17</td>
<td>30.9%</td>
</tr>
<tr>
<td>Other factors influencing aid effectiveness</td>
<td>22</td>
<td>40.0%</td>
</tr>
<tr>
<td>Paris declaration in fragile situations</td>
<td>3</td>
<td>5.5%</td>
</tr>
<tr>
<td>Aid modalities</td>
<td>10</td>
<td>18.2%</td>
</tr>
<tr>
<td>Evaluation methods for the Paris Declaration</td>
<td>3</td>
<td>5.5%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>55</td>
<td><strong>100.0%</strong></td>
</tr>
</tbody>
</table>

### 2.4.1 Evaluation of implementation of the Paris principles in the health sector

Out of the 17 articles related to the implementation of the Paris principles in the health sector, 11 presented the progress and challenges of the process of implementing one or more than one principle of aid effectiveness in the health sector. The other 6 articles actually tried to comment on the outcomes of implementing the Paris principles in the health sector.

Studies in different countries identified challenges and gaps in implementing the Paris principles in the health sectors. Assessing the case of implementation of the Paris principles in Rwanda’s health sector, Hyman (2011) claimed that the progress on the
progressive ownership by the recipient country, a concept implicit within the Paris principles of aid effectiveness, was not so encouraging in Rwanda. This was because the donors used to assert joint ownership and control over the aid funded interventions in most cases. Similarly, Monye et al. (2010) pointed out a disconnect between the intent of the Paris principles and actual donor practice in Nigerian health sector. Sossouhounto et al. (2011) also found a gap between the policy and practice in Benin. Their study found that although Benin’s government issued a health policy, most actors at the operational level of the health sector were unaware of the contents of the policy. The study by Sundewall et al. (2010) in Zambia’s health sector implied the difficulties in harmonizing and aligning different organizational requirements with the Zambian health-sector plans.

Earlier another study in Zambia (Sundewall et al., 2009) noted non-compliance with the Paris principles at the district level as donor funding was not included in the local plans, use of the programme based support mechanism was not increasing, and predictability of external funding remained low. However, despite vertical programming this study noted that the funding for HIV/AIDS showed better integration in planning and implementation at the district level than the funding mechanisms for other projects. But Uganda’s experience with the HIV/AIDS funding was far from complicit with the Paris principles. Stierman et al. (2013) found that in Uganda’s health sector although there was a sector-wide approach formed, funding for disease specific activities such as for HIV/AIDS was increasingly provided through short term, project based approaches.

McGee and Heredia (2012), however, reported a different experience in Columbia. Assessing the implementation of the Paris principles of aid effectiveness in Columbia from the perspective of aid-relationships, McGee and Heredia argued that as a result of the Paris principles of aid effectiveness, aid donors in Columbia were subjected to an attempt to be pushed into a technocratic corner, and advocacy outcomes of the social actors were put under threat. This indicates a limited role of both donors and civil society in the recipient country as an unintended effect of implementing the Paris principles.
In relation to the principle of country ownership, Sridhar (2009) identified three structural challenges for developing countries: the proliferation of donor funded initiatives, donor influence on priority setting and donors’ lack of accountability, and the sustainability of current levels of external financing. Few other studies on implementation of the Paris principles in the health sector also reported poor progress. Muvudi et al. (2011) reported that progress of adhering to the principles of aid effectiveness in five years after the Paris Declaration had been very limited in the Democratic Republic of Congo. Paradoxical to the Paris principles, Muvudi et al. argued, resources provided through the bilateral and multilateral aid mechanisms contributed to weakening the health system and compliance with the principles of alignment, harmonization, and country ownership were still rudimentary in Congo’s health sector.

Shorten et al. (2012) reported mixed results from a self-reported survey by IHP+, an international partnership of international organizations, bilateral agencies and country governments committed to putting the principles of effective aid and development cooperation into practice in the health sector. Their results from the surveys in Burundi, Mozambique, Cambodia, Nepal, Ethiopia, Nigeria, Kenya, Mali, Madagascar, and Zambia found incremental improvements in implementing the principles of country ownership, mutual accountability, and alignment, but noted a lack of progress in the use of countries’ financial management, procurement, and monitoring systems.

While all these studies focused on the process of implementing the Paris principles of aid effectiveness in the health sector, they did not attempt to evaluate the impact of such implementation. Based on the Paris Declaration evaluations conducted by OECD, Martini et al. (2011) attempted a critical review of the progress of the implementation of the Paris principles in the health sector. Like the self reported survey results of IHP+ (Shorten et al., 2012), Martini et al. also noted mixed results in the health sector with good progress in terms of country ownership and coordination, but lack of expected progress in the use of country systems, managing for results, and mutual accountability. However, due to the limitations of surveys and data availability, these writers could not make generalizations or clearly link the progress in aid effectiveness with improvements in health outcomes.
Attempts to link progress in following the principles of aid effectiveness with the health results can be traced in a report for the Spanish aid (Camacho & Batista, 2008). This report noted:

“Aiming at these principles [of aid effectiveness], the use of new instruments-programmatic instruments- and new approaches - the SWAp [sector-wide approach]- are being promoted to increase the effectiveness of aid. Despite some weaknesses, the assessment conducted on their utilization seems to show better performance than that achieved with traditional instruments” (p.237).

Capturing lessons learned from the US government’s PEPFAR funding mechanism, Mark Dybul (2009) noted that PEPFAR followed the principles of country ownership, results-based accountable approach, engagement of all sectors, and good governance. Dybul claimed, “With resources and a strong intellectual base, PEPFAR saved lives and provided lessons learned for effective development” (p. S12).

Studying the effect of the Paris principles on the human resources for health in Laos, Dodd et al. (2009) found a number of pathways through which aid effectiveness was promoting an integrated, holistic response to a range of human resources for health challenges. They identified further opportunities for using the Paris principles for health workforce planning, regulation, and capacity development.

Proposing an analytical framework based on the logical chain between implementation process, health system strengthening, and health outcomes in the context of monitoring results of aid effectiveness in the health sector, Paul et al. (2011) conducted a three-level assessment in Mali for the process, system’s effect, and health outcomes. Their assessment found some progress and positive changes attributable to the implementation of aid effectiveness principles as well as improved outcome and impact indicators in the health sector. However, Paul et al. could not claim a clear link between the adherence to the Paris principles and improved health outcomes and concluded that despite the changes observed in behaviour and practices, adherence to the principles of aid effectiveness was far from complete in Mali. Thus expectations of better health outcomes from the adherence to the Paris principles, Paul et al. argued, should be realistic.
In an attempt to see if there was a difference in outcomes of aid delivered in line with the Paris principles of aid effectiveness from the aid in general, Taylor et al. (2013) conducted a systematic review of studies on aid-funded interventions designed to improve maternal and reproductive health outcomes. While their review found some association between aid interventions and small improvements in maternal and reproductive health outcomes, data available to them based on search of electronic databases did not allow for a meaningful comparison between the general aid and aid delivered in line with the Paris principles.

Dalil et al. (2014) reviewed the use of aid effectiveness principles in rebuilding Afghanistan’s health systems and identified a number of factors including the Paris principles that contributed to the apparent success of the country-donor partnerships. These factors included the ownership by the country with the Ministry of Public Health playing a stewardship role, donor-country alignment, joint decision making by government and donors, managing for results, reliable aid flows, and human factors such as deployment of a “critical mass of individuals with the right experience and expertise” (p. S124).

While findings from these studies cannot be generalized, a few points stand out in these attempts to evaluate adherence to the Paris principles in the health sector:

- Implementation of the Paris principles of aid effectiveness faced challenges from the gap between policy and practice and an apparent tension between the subset of principles of country ownership, alignment, and harmonization; and the subset of the principles of managing for results, and mutual accountability;
- Countries with adequate management systems and capacity can more readily benefit from the implementation of the Paris principles;
- There are indications that adherence to the Paris principles improves health outcomes. However, since the Paris principles have not been fully adhered to, a causal link between the adherence to the Paris principles and improved outcomes of health interventions cannot be established without some experimental data.

It can, however, be noted that all these studies focused on either the process or the process and effect of adherence to the Paris principles at the sector level. While this
sector focus at the country level is consistent with the indicators of monitoring of progress of the Paris principles (OECD, 2007a), the high level focus makes it difficult to measure the outcomes of compliance with the Paris principles at an individual aid funded intervention level. It appears that none of the studies mentioned successfully attempted to measure the health outcomes at the intervention level and link them to the extent of compliance by that intervention with the Paris principles.

2.4.2 Other factors influencing aid effectiveness

A relatively large percentage (40%) of sampled literature was devoted to the factors influencing the implementation of the Paris principles or influenced by such implementation in the health sector. Studies by Armon (2007), de Koster and Holvoet (2012), Engberg-Pedersen (2014), and Holvoet and Rombouts (2008) highlighted the importance of political context, political governance, and domestic accountability in aid management for greater effectiveness of aid. Other studies by Hyden (2008) and Poku and Whitman (2012) argued for the need for an understanding of transformation of power relations between the donors and recipient governments as well as between the governments and civil society organisations for and; as a result of, implementation of the Paris principles of aid effectiveness.

Apart from the narrowly defined technocratic view of the Paris principles lacking an understanding of the political context and power relations, the Paris Declaration is also criticized for its lack of attention to the role of civil society and gender. Gruber (2011) and Kelly and Birdsall (2010) noted that the new aid modalities encouraged by the Paris principles might limit the scope of unique contributions that civil society can make, especially to the national responses to HIV/AIDS. Drawing on recently conducted research on women and development in Kyrgyzstan, Campbell and Teghtsoonian (2010) pointed out that the Paris principles of aid effectiveness raised new questions about the place of gender in development practice. By employing a gender lens, they noted that while the engagement by the development professionals and activists with the aid effectiveness discourse increased visibility of the governance of aid, it did not necessarily benefit women.
A handful of studies emphasized the factors related to the donors—donor institutions, donor commitment, and donor quality—for the success of aid effectiveness. Analyzing Brazilian international cooperation in health, Buss (2011) identified a need for legislation on international cooperation and harmonization of many of its institutions in order to improve its effectiveness. Based on an empirical examination of donor dynamics in Norway, the UK, and Canada, Gulrajani (2014) also showed the importance of the relationships between donor organisational factors and aid effectiveness goals. Analyzing a local governance case in Indonesia, Winters (2012) explored the political and bureaucratic obstacles faced by bilateral and multilateral aid organizations in trying to harmonize aid at the country level. Winters believed that the lack of such harmonization can be “linked to the dominance of strategic interests of some donors and the structure of bureaucratic incentives within aid agencies” (p.316). In a similar way, White (2011) argued that donor commitment is a critical success factor for the effectiveness of aid in the health sector while Knack et al. (2011) proposed use of a donor index based on the Paris Declaration (2005) and the Accra Agenda of Action (2008) to measure and rank donor quality.

A number of other studies, however, pointed at factors related to the recipient country such as country capacity, governance, and the quality of country systems required for successful implementation of aid effectiveness principles. Analyzing the implementation challenges for a SWAp in Nepal related to the issue of funding through the NGOs in the health sector, Giri et al. (2013) raised the question of the capacity of the government’s MoH and Population for the success of such an approach in line with the Paris principles. Goldberg and Bryant (2012) also highlighted the need for country-owned capacity building but maintained that although the Paris Declaration detailed the principles of country ownership to which partner and donor countries should commit, it did not identify the specific mechanisms to carry out these principles. Through an analysis of the experience of results based management by several African countries, Wone et al. (2012) also identified the need for improved governance and capacity building for successful implementation of such an approach in the health sector. In another study using the indicators endorsed by the Paris Declaration, Knack (2014) found a positive, significant, and robust relationship between the quality of the country systems such as the public financial management system, procurement system, and monitoring and
evaluation system of the recipient country and use of these systems by donors as part of the Paris principle of ‘alignment to country systems’.

Some studies have drawn attention to special areas of interest that might have an effect on health intervention outcomes along with the implementation of aid effectiveness principles, whereas others have pointed to the complex landscape for aid effectiveness. For example, Borda-Rodriguez and Huzair (2011) advocated for the use of pharmacogenomics and personalized medicine to achieve greater aid effectiveness.

Drawing on case studies in four countries, Dodd et al. (2009) analyzed that planning process in the health sector could be improved with the contribution of sexual and reproductive health policies at the high level. Based on a structured review of literature, Pridmore and Carr-Hill (2011) advocated addressing the drivers of child under-nutrition for improved effectiveness of interventions. Their recommendations included improved implementation of effective interventions on a large scale, uninterrupted donor commitment toward adherence to the Paris principles, enabling governance structures at the country level, capacity building, and domestic accountability. Analysing the transitions in the conceptualization and practice of international development cooperation, Hill et al. (2012) appreciated the fact that the Fourth High Level Forum in Busan recognized “a much more complex landscape for development than previously conceived”. But they believed that the challenges to coordination still remain in this complex landscape.

It seems the political context, donor organisational factors, as well as recipient country’s capacity and readiness of its systems play important roles in the interplay of aid effectiveness principles. Although strongly advocated for in several studies, this review did not find any empirical results on gender and civil society related to the implementation of the Paris principles in the health sector.

2.4.3 The Paris principles in fragile situations

Only 5.5% of the sampled literature discussed the application of the Paris principles in fragile states and situations. Discussing a case in Afghanistan, Goodhand and Sedra (2010) commented that it is not clear how international donors’ commitment to country ownership and partnership are translated in practice in fragile state or ‘post-conflict’
settings. “The developmentalist principles outlined in the 2005 Paris Declaration”, as Goodhand and Sedra (2010) think, “may carry little meaning in such contexts and their application can have paradoxical effects that impede the emergence of broad-based ownership” (p. S78).

Reviewing the Asian Development Bank’s approach to engaging with the fragile states, Rosser (2009) criticized the paradoxical practice. According to Rosser, while the Paris principles of aid effectiveness encouraged allowing fragile states to “exercise leadership in determining their own development strategies”, the Asian Development Bank rather undertook a risk avoidance strategy that prioritized neo-liberal policies, and used coercive means to promote such policies and interests.

Reviewing the aid modalities used in Zimbabwe during the period of crisis, 2008-2012, Salama et al. (2014) appreciated a “Transition Fund Model” employed by UNICEF and other development partners in Zimbabwe. Their study found that in line with the Paris principles, the Transition Fund Model, worked with the government Ministries to determine local priorities, pooled donor funds to achieve scale, built on national systems and capacities, and avoided delivering aid through fragmented projects and setting up parallel mechanisms. The study claimed that in this process the Transition Fund Model “was able to achieve important results in health, education, social support and water services in a challenging setting” (Salama et al., 2014, p. 1).

These studies indicate that although the principle of country ownership is often contested in fragile states, complying with the Paris principles can produce good results even in fragile and post-conflict situations.

2.4.4 The Paris principles in different aid modalities

Around one-fifth (18.2%) of the sample literature considered compliance with the Paris principles by different aid disbursement and aid management modalities. Seven out of the ten selected studies in this category discussed the global health partnerships including the Global Fund. Articles by Dujardin (2009) and Kerouedan (2010) indicate that the performance of vertically targeted health programmes such as the ones funded by the Global Fund can be affected by the weaknesses of the health systems. They
believe that expansion of such vertically funded programme services and their impact can be constrained by the Health systems’ weaknesses, such as the shortage of human resources, laboratory commodities, and medicines.

Acknowledging the importance of the health system’s effect on programme outcomes, Galichet et al. (2010) and Sambo et al. (2011) encouraged the use of health systems strengthening initiatives of the global health partnerships by the recipient country. Grepin (2011), however, cautioned that “integrating additional health services into existing vertical HIV infrastructure may not mitigate some of the challenges these programs have introduced in implementing countries” (p. S77).

For reshaping the global health initiatives, Dybul et al. (2012) recommended the use of the principles of results based management and mutual accountability in line with the aid effectiveness declarations. Kvale and Sommerfelt (2009), however, expressed doubt that “results based funding may have short term beneficial effects on specific areas, but the total and long-term effects [of this approach] are not known and adverse effects have been described” (p.6).

As opposed to a disease specific vertical funding modality, budget support and the SWAp are believed to be more compliant with the aid effectiveness principles. However, Molenaers (2012) identified that:

“Budget Support (BS) has been considered the aid modality that best realises the Paris Declaration principles of alignment, harmonization and respect for recipient ownership. In design the modality has a very strong technocratic focus, and the OECD/DAC has endorsed the idea that BS should be delinked from broader political concerns. In reality, however, donors do use BS to leverage more and better democratic governance” (Molenaers, 2012, p. 791).

This finding, in fact, resonates the influence of political factors in the implementation of aid effectiveness principles and raises the question what aid or implementation of aid effectiveness principles would primarily like to leverage: better health outcomes or better democratic governance.

Paul et al. (2011), on the other hand, identified a positive impact of a SWAp in preparation and implementation of “country Compact”. As introduced by IHP+, a country Compact is “a negotiated and signed agreement in which partners commit to
implement and uphold the priorities outlined in the national health strategy” (Paul, Berthe, & Samake, 2011, p. 42). Comparing two different approaches to putting a “country Compact” into practice in Mali and Benin, Paul et al. (2011) identified that Mali benefited from its experience in leading a SWAp in the health sector that resulted in improved donor coordination, government leadership in piloting national programme, trust between partners, and broad ownership of the health plan. Benin’s process, on the contrary, suffered from a lack of SWAp experience and being less inclusive, resulting in narrow ownership and vague commitments to the country Compact.

In another study, Molenaers et al. (2014) analyzed a Belgian case of the use of NGOs for channeling development assistance. As opposed to the experience of ‘country Compact’, their research found that channeling development assistance through NGOs in the Belgian case made it difficult to fight aid fragmentation by contributing to multiple channels and incentive structures and, thus, created a burden for the recipient country.

Regardless of aid modalities, all these studies point out the importance of health systems and compliance with the Paris principles for better results in the health sector. However, like other studies, they predominantly focus on the process of following the aid effectiveness principles, rather than gathering substantial evidence of outcomes from such a process.

2.4.5 Evaluation methods for the Paris principles

Very few studies found in the literature detailed methods for evaluating the outcomes of following the Paris principles. Armytage (2011) identified that the practice of evaluating aid effectiveness is constrained by an apparent “paradigm war” between a positivist approach of gathering empirical evidence through randomized controlled trials (RCTs) and use of counterfactuals; and a constructivist approach of participatory methods, case studies, and qualitative observations. According to Armytage, “Despite substantial refinement in evaluation approaches, evaluation remains without any orthodoxy about how to assess effectiveness” (p.261), and there is an ‘evaluation gap’ between the rhetoric of donors for results based management and accountability, and their evaluation practice.
Caspari (2009) also believed that as reinforced by the principles of aid effectiveness, the task of evaluation is to provide reliable findings on the impact of aid interventions for the policy makers to choose from evidence-based options in designing future interventions.

In the searches for adequate evaluation methods for the Paris principles, the study by Mensa-Bonsu and Andersen (2010) provides documentation and lessons learned from a joint evaluation of development cooperation between Denmark and Ghana, that is, an evaluation jointly conducted by the donor and recipient country. They noted that such “joint evaluation of development cooperation is becoming more imperative within the framework of the Paris Declaration” (p.295) but engagement in a joint evaluation takes time, demands resources, and requires strong commitments both from the donor and the recipient.

Although highlighted the demand for evaluations and discussed about some broad approaches to evaluation, these three articles did not present any practical framework through which the impact of the implementation of the Paris principles could be evaluated.

In summary, the systematic review of sample literature reveals a dominance of process evaluation over the impact evaluation of the Paris principles; a list of contextual factors that influence the process and outcomes of the implementation of the Paris principles in the health sector; and the need for adequate methods for evaluating the impact of different aid modalities and impact of aid effectiveness principles. The following section reviews a specific set of literature on different aspects of these issues for defining the scope of this thesis.

2.5 Selected literature for scoping

While the systematic review based on a protocol driven search presented in Section 2.4 provided a good sample of related literature, as the study unfolded, an additional set of literature was identified through snowballing, from the donor websites including the
website of OECD/DAC, and additional search based on other thematic topics as discussed below.

The wide range of additional literature reviewed in this section can be grouped under the following thematic categories:

- State fragility
- Health systems
- Aid instruments in fragility context
- Evaluation theories and techniques
- Aid effectiveness evaluation
- Methods and tools of comparative analysis

### 2.5.1 State fragility

A number of donor documents tried to define characteristics of state fragility. DFID (2005) defines fragile states as “those where the government cannot or will not deliver core functions to the majority of its people, including the poor”. The OECD/DAC (OECD, 2007d; OECD-DAC International Network on Conflict and Fragility, 2009) provides a number of classifications for fragile states: (i) post-conflict/crisis or political transition situations; (ii) deteriorating governance environments; (iii) gradual improvement; and (iv) prolonged crisis or impasse. According to the World Development Report on Conflict, Security and Development (World Bank, 2011c), building institutions in a fragile state to help the country get out of fragility may take 20 or 30 years.

Fragile states often lack effectiveness in delivering the core functions of government due to their low capacity or willingness or they lack legitimacy to be a representative and accountable government (Newbrander, 2007). Barakat (2009) also pointed out that fragile situations are often marked with insecurity, administrative weakness of the government, excessive centralisation, ethnic hostilities, and high competition among stakeholders. Fragility, in both donor documents and scholarly literature, is accepted as “multi-causal and multi-dimensional” in any given context (Chabal & Daloz, 1999; Clapham, 2002; Collier, 2007; Douma, 2006; Engberg-Pedersen et al., 2008; Kaplan, 2008). Mcloughlin (2010) produced a topic guide on fragile states and summarized the causes and characteristics of fragile states that can be found among the more prevalent
in the related literature. Mcloughlin broadly categorized such factors of fragility as: structural and economic factors, political and institutional factors, social factors, and international factors.

There are attempts at measuring state fragility. A few institutions including the World Bank, Center for Systemic Peace, and Fund for Peace regularly publish indices to measure and rank state fragility. The World Bank assesses state fragility through a “harmonized list of fragile situations” with scores on ‘Country Policy and Institutional Assessment’ (CPIA) and the presence of a United Nations and/or regional peace-keeping or peace-building mission during the past three years (World Bank, 2015).

The Center for Systemic Peace, on the other hand, uses an index based on a country's scores on the effectiveness and legitimacy in security, political, economic and social areas (Marshall & Cole, 2012). The Fragile States Index of the Fund for Peace also considers a number of social, economic, political and military indicators to measure a country’s fragility or degree of failure every year (Fund for Peace, 2014). Another tool for fragility assessment is the in-country fragility assessment based on the peace building and state building goals as agreed in the “New Deal for Engagement in Fragile States” (International Dialogue on Peacebuilding and Statebuilding, 2011).

While these documents suggest the characteristics, causes and measurement of state fragility, there are studies suggesting that measuring aid effectiveness in fragile situations can be a challenge (Chauvet & Collier, 2004; Levin & Dollar, 2005; McGillivray, 2005). Although a set of indicators was suggested for monitoring the progress of good international engagement in fragile states and situations (OPM/IDL, 2008), actual outcomes and development results of implementing the Paris principles in fragile situations have not been adequately studied so far. This is probably because of the volatility of the fragile situation, lack of data, presence of multiple aid management modalities, and multiple primary and secondary objectives tied to the aid such as peace building, state building and development objectives in fragile situations.
2.5.2 Health systems

Berman and Bitran (2011) proposed health systems analysis as a ‘distinct methodology’ for designing policies and programmes for health systems strengthening. According to Berman and Bitran,

“Health systems analysis seeks to understand the determinants of health system performance and to develop better policies and strategies to improve that performance. Health systems analysis involves data collection on health system inputs, processes, and outputs. It then analyzes how these combine to produce outcomes— that is, their effects on individual and population well-being” (p.5).

As highlighted in the systematic review, outcomes of aid interventions are also influenced by the health systems of a country. WHO provides a health systems framework that analyzes the process, organization and goals of a health system in terms of six “system building blocks” such as health governance and leadership, health service delivery, health workforce, health information, health commodities and supplies (medical products, vaccines, and technologies), health financing, and leadership and governance (WHO, 2000).

There has been growing emphasis (Adams & Savigny, 2012; Coker, Atun, & McKee, 2008; Leischow et al., 2008; Naaldenberg et al., 2009; Savigny & Adams, 2009; WHO, 2007) on considering the effect of health systems on a project (as well as effect of the project on the health systems) to bring into the analysis of contributions and impacts of a health intervention. A number of other writers also confirmed that conflict affected fragile states are often worse off than the non-fragile states in terms of key health indicators and social determinants of health due to the weak and incomplete health system building blocks and the health system’s inability to provide quality health services to a large proportion of population (Eldon, Waddington, & Hadi, 2008; Newbrander, Waldman, & Shepherd-Banigan, 2011; Ranson et al., 2007). In this line, Kruk et al. (2010) pointed out that the national health system, in a fragile state, is often a victim of conflict, experiencing destruction of clinic and hospital infrastructure, the flight of health professionals, and the interruption of drugs and other medical supplies to those in need.
These studies indicate that state fragility is one of the significant factors for the performance of the health systems, which ultimately influence the outcomes of a health intervention.

2.5.3 Aid instruments in a fragility context

Health systems in fragile states often have to go through the phases of relief, rehabilitation, reconstruction and development (Newbrander, Ickx, & Leitch, 2003). Different aid instruments and aid management modalities in fragile situations often fall within a ‘relief/development’ dichotomy (Vergeer et al., 2009) in a state-avoiding (i.e., channeling aid directly to the target people by avoiding state mechanisms of the recipient country as often in case of relief) and state-building (i.e., channelling aid through state mechanisms of the recipient country in an effort to strengthen the state and address longer term development needs) continuum. Therefore, finding an effective aid instrument in fragile situations has been subject to policy research and experimentation (Leader & Colenso, 2005). As the Management Sciences for Health (2007) points out, the aid disbursement options in fragile situations generally include: direct budget support to the government, a SWAp with donor pool funding, contracting NGOs for service delivery, and global health partnerships.

Budget support, Sector wide approach (SWAp) and pool funding:

A SWAp espouses increased reliance on budget support based on a government-owned comprehensive sector policy; greater ownership of the policy by key stakeholders; enhanced co-ordination and harmonization among major donors; and use of government systems (Hill, 2002). Walford (2003) believes that a SWAp in health reduces transaction costs of aid and has a systemic effect to increase efficiency of the government’s own resource management. However, a number of published works on health SWAp evaluation show mixed results. White (2007), Chansa, et al. (2008), Watt (2005), and McNee (2012) argue that SWAs are often donor-led, and have not moved away from the project finance approach to the extent suggested by the rhetoric of proponents of the sector-wide approach. Moreover, as these writers believe, harmonization in implementing health SWAPs has been weak and the claimed savings in transaction costs are non-existent. Brown et al. (2013) also point out a similar gap between the rhetoric of country ownership and World Bank’s actual policies and
practices in taking part in a pool funding mechanism such as the health system funding platform.

For fragile situations, Chapman and Vaillant (2010) noted that with some notable exceptions, the experience of a SWAp and pooled funding mechanism at the country level is that such mechanisms were slow to set up, costly to manage, and had achieved modest development impact. They also identify some typical reasons of such weak performance of a pooled fund in a fragile situation:

“Operating in a fragile state is typically more labour-intensive and expensive than elsewhere for a variety of reasons including the weakness of host governments, a risky operating environment, difficult communications and ill-adapted internal procedures and regulations” (p.xiii).

An investigation of lessons learned from a successful implementation of health SWAp in Uganda, however, noted that:

“SWAp s take time, and even when performance indicators do not improve after 2-3 years implementation, donors should recognize the scale of changes going on, and that recipient governments need donors’ continued commitments to the SWAp” (Hutton, 2004, p. 11).

Negin and Martiniuk (2012) also gave account of a similar experience in the Solomon Islands. There were initial tensions between donors and the government that undermined the country ownership of the SWAp but over time it made strong progress with greater government ownership and with more focus on partnership and harmonization rather than on funding mechanisms.

These studies indicate that the success of a SWAp varies in different contexts and depends on how the SWAp is implemented in a particular context.

**Global health partnerships:**

The review of additional literature related to the global health partnerships such as the Global Fund also shows both positive and negative effects. According to Biesma et al. (2009), positive effects of global health partnerships include:

“a rapid scale-up in HIV/AIDS service delivery, greater stakeholder participation, and channeling of funds to non-governmental stakeholders, mainly NGOs and faith-based bodies. Negative effects include distortion of recipient countries’ national policies, notably through distracting governments from coordinated efforts to strengthen health systems and re-verticalization of planning, management and monitoring and evaluation systems” (p.239).

Similarly, in fragile states’ context, Newbrander et al. (2011) identified that:
“GHPs [Global Health Partnerships such as the Global Fund] can aid fragile states in filling gaps, such as restarting a national tuberculosis (TB) programme with a grant from the Global Fund to Fight AIDS, Tuberculosis and Malaria. Potential disadvantages are that such programmes may not be integrated into basic health services, may not be sustainable, or may not provide support for health system development” (p.652).

Atun and Kazatchkine (2009) argued that the Global Fund had been promoting country ownership and stewardship in health programmes. However, Cruz and McPake (2011) identified that in Uganda the Global Fund did not contribute to the health SWAp, rather opted to create a parallel system of management by using a separate project management unit within the MoH, different monitoring tools, and a parallel system for procurement. A five-year evaluation of the Global Fund also commented on the weak partnerships and country level harmonization:

“The Global Fund partnership model has opened spaces for the participation of a broad range of stakeholders. This progress notwithstanding, existing partnerships are largely based on good will and shared impact-level objectives rather than negotiated commitments or clearly articulated roles and responsibilities, and do not yet comprise a well functioning system for the delivery of global public goods” (Macro International Inc, 2009, p. 33).

Findings of these studies reinforce the findings from the systematic review that funding for vertical programmes such as those funded by the global health partnerships may not be successful without following the Paris principles and without addressing health systems issues.

Contracting NGOs:

Typically in their transition from relief to development fragile countries face a tension between the short-term need for ensuring essential services and the long-term need for state capacity building. Contracting NGOs in such a situation is often necessary. As Rocha-Menocal (2009) commented:

“In fragile and conflict-affected states, there may be tensions between the imperative to provide basic services to the population urgently, by any means, and the imperative to prioritise state-building ... at least in the short term, faced with weak state systems and low capacity, international agencies may find it necessary to further bypass the state by funding urgent services through international NGOs to generate quick and visible improvements in living conditions” (p.3).

Similarly Pantuliano (2009), in case of a fragile situation in South Sudan, noted:

“It is clear that the conventional aid architecture is ill equipped to cater for a situation that spans humanitarian and development needs. As in many ‘post-conflict’ countries, Southern Sudan still needs support for direct service delivery, alongside support for the building of government capacity” (p.1).
However, Witter (2012) warned that although providing short-term benefits, such contracting mechanisms, might have a negative impact on the building of sustainable service-delivery systems for the long term.

Reviewing the experience of vertical funding in Mozambique, Mussa et al. (2013) also identified three serious concerns related to channeling funding through international NGOs: difficulties in coordination and local control, inequalities created within the health system, and exodus of health workers from the public sector to international NGOs provoked by salary disparities.

Aid modalities in fragile situations each have inherent strengths and weakness but their success may depend on how they are implemented in a particular context and to what extent they follow the Paris principles of aid effectiveness.

### 2.5.4 Evaluation theories and techniques

Unlike other areas of social research, research on aid effectiveness has limited theoretical frameworks. Apart from the economic growth models, researchers have used other theories such as the “agency theory” that tries to explain the principal-agent relationship that revolves around resolving conflict and aligning two groups’ interest (Eisenhardt, 1989) and the “need-interest theory” that separates determinants of aid allocation into two models, namely recipient needs and donor interests (Alesina & Beatrice, 2002; Alesina & Dollar, 2000; Maizels & Nissanke, 1984; McKinlay & Little, 1977). While these models and theories help to see the aid motivations and empirical information for macro level effect of aid, they do not fully explain aid’s interactions with health systems and the results of those interactions to evaluate the impacts.

A realistic evaluation, which is also known as “realist evaluation” offers a framework that can accommodate the interaction between the Paris principles, health systems, and aid interventions. A realistic evaluation, according to Pawson and Tilley (1997), is based on a belief that the outcome of an intervention depends on the interaction between the mechanism and context of that intervention. That means the impact of an aid intervention will depend on the pathway of the intervention’s interactions with the health systems to produce both the intended and unintended effects. Analysis of such
interactions in the health systems involves “gathering data on health system inputs, processes, and outputs; and analyzing how these combine to produce the outcomes” (Berman & Bitran, 2011). Proponents of health systems thinking believe that comprehensive evaluation of a health intervention requires such analysis of health systems interactions and considers both intended and unintended outcomes in the health systems (Dodd et al., 2007; Olmen et al., 2010; Savigny & Adams, 2009).

An intervention’s interactions with the health systems and their outcomes can be explained by the intervention logic model. The intervention logic model (Cooksy, Gill, & Kelly, 2001) or the results-chain model (Margoluis & Salafsky, 1998) is also known as programme theory (Bickman, 1987, 1990; Bickman & Peterson, 1990), theory of change (Weiss, 1995, 1997), logical framework (Alter & Egan, 1997) and programme matrix (Chen, 1990). By eliciting the underlying logical chain, programme theory attempts to draw the causal relationships between resources (inputs), activities (process), outputs, outcomes and impact of an intervention (Caracelli & Greene, 1997; United Way of America, 1996; Wholey, 1994). Using the programme logic model, a “Corporate Scorecard Model” was suggested by the World Bank Operations Evaluation Department (Picciotto, 2002) to evaluate the aggregated performance of externally funded national development programmes. Using a theory-driven on-going performance management approach, Mayne (2001) suggested a “contribution analysis” for analyzing and attributing plausible contribution of the performance of an intervention on the results chain.

Victora et al. (2009) argue:

“Evaluation of large-scale programmes and initiatives aimed at improvement of health in countries of low and middle income needs a new approach. Traditional designs, which compare areas with and without a given programme, are no longer relevant at a time when many programmes are being scaled up in virtually every district in the world” (p.83).

Traditionally most quantitative techniques of impact evaluation involved a “counterfactual” and tried to mimic a randomized controlled trial (RCT) to ascertain if the changes in well-being are indeed due to the intervention and not because of any other factors (Khandker, Koolwal, & Samad, 2010). While RCTs are believed to be the “gold standard” in outcome efficacy trials for clinical interventions where “the causal chain between the agent and the outcome is relatively short and simple and where
results may be safely extrapolated to other settings” (Victora, Habicht, & Bryce, 2004), use of RCT in evaluating large scale national level interventions is often not practical due to the availability, ethical, and often political and economic challenges associated with having a control group.

Recent developments in the field of theory based programme evaluation and realistic evaluation suggest that impact evaluation of large-scale programmes can still be conducted without necessarily employing an RCT by theory based observational studies using adequacy and plausibility designs (Habicht, Victora, & Vaughan, 1999; Hughes & Hutchings, 2011; Imbens & Wooldridge, 2009; Victora et al., 2004).

Betts and Wood (2013) illustrated the process and methods used in the evaluation of the Paris Declaration undertaken by OECD. They highlighted, that confronted by the challenge of linking development results to the implementation of the Paris principles, the Paris Declaration Evaluation employed a theory-driven evaluation approach with programme theory, contribution analysis, and mixed methods across 21 country evaluations and 7 donor studies. Findings from these studies were then systematically analyzed and synthesized in the Evaluation report (Wood & Betts, 2013). While these works offer useful considerations and theoretical framework for evaluation, they do not provide an example of actual evaluation of the Paris principles at the intervention outcome level in a fragile state's health sector, as these studies attempted a high level sectoral analysis and did not include any fragile state in particular in their sample of evaluation.

2.5.5 Aid effectiveness evaluation

In an attempt to monitor the progress of the implementation of the Paris principles, a set of monitoring indicators and targets were set from the Paris Declaration by OECD. The list of these indicators and progress against their targets is presented in Table 2.4:
Based on these indicators of monitoring the progress of implementation of the Paris Declaration, OECD undertook a series of surveys in 2006 (OECD, 2007a), 2008 (OECD, 2008a), 2010 (OECD, 2011a), and 2011 (OECD, 2012) for the baselines and progress over time. As seen in Table 2.4, these indicators measure the progress in terms of the number of countries and donors meeting their commitments in implementing the Paris Declaration monitoring indicators and progress, 2010

<table>
<thead>
<tr>
<th>Paris Declaration Indicator</th>
<th>2010 Actual</th>
<th>2010 Target</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Operational Development Strategies</td>
<td>37% (of 76)</td>
<td>75%</td>
<td>Not met</td>
</tr>
<tr>
<td>2a Reliable public financial management (PFM) systems</td>
<td>38% (of 52)</td>
<td>50%</td>
<td>Not met</td>
</tr>
<tr>
<td>2b Reliable procurement systems</td>
<td>--</td>
<td>No Target</td>
<td>--</td>
</tr>
<tr>
<td>3 Aid flows are aligned on national priorities</td>
<td>41%</td>
<td>65%</td>
<td>Not met</td>
</tr>
<tr>
<td>4 Strengthen capacity by co-ordinated support</td>
<td>57%</td>
<td>50%</td>
<td>Met</td>
</tr>
<tr>
<td>5a Use of country PFM systems % of aid for the government sector using partner countries’ PFM systems</td>
<td>48%</td>
<td>55%</td>
<td>Not met</td>
</tr>
<tr>
<td>5b Use of country procurement systems % of aid for the government sector using partner countries’ procurement systems</td>
<td>44%</td>
<td>No Target</td>
<td>--</td>
</tr>
<tr>
<td>6 Strengthen capacity by avoiding parallel PIUs Total number of parallel project implementation units (PIUs)</td>
<td>1 158</td>
<td>565</td>
<td>Not met</td>
</tr>
<tr>
<td>7 Aid is more predictable % of aid for the government sector disbursed within the fiscal year for which it was scheduled and recorded in government accounting systems</td>
<td>43%</td>
<td>71%</td>
<td>Not met</td>
</tr>
<tr>
<td>8 Aid is untied % of aid that is fully untied</td>
<td>86%</td>
<td>More than 89%</td>
<td>Not met</td>
</tr>
<tr>
<td>9 Use of common arrangements or procedures % of aid provided in the context of programme-based approaches</td>
<td>45%</td>
<td>66%</td>
<td>Not met</td>
</tr>
<tr>
<td>10a Joint missions % of donor missions to the field undertaken jointly</td>
<td>19%</td>
<td>40%</td>
<td>Not met</td>
</tr>
<tr>
<td>10b Joint country analytic work % of country analytic work undertaken jointly</td>
<td>43%</td>
<td>66%</td>
<td>Not met</td>
</tr>
<tr>
<td>11 Results-oriented frameworks % of countries with transparent and monitorable performance assessment frameworks</td>
<td>20% (of 44)</td>
<td>36%</td>
<td>Not met</td>
</tr>
<tr>
<td>12 Mutual accountability % of countries with mutual accountability reviews in place</td>
<td>38%</td>
<td>100%</td>
<td>Not met</td>
</tr>
</tbody>
</table>

Notes:

a. Assessment against 2010 target uses data for all 78 countries participating in 2011 for which data were available. Where data are available for only a subset of these countries, the sample size is indicated in brackets.

b. Assessment against 2010 target uses data for the 32 countries participating in both the 2006 and 2011 Surveys, as the indicator target is formulated in relation to the 2005 baseline. Targets may differ from those published in previous years as baselines have been recalculated, omitting data from two countries (Nicaragua and Yemen) which formed part of the original panel of 34 countries participating in 2006, but which did not participate in 2011.

c. No targets are presented for indicators 2b (reliable procurement systems) and 5b (use of country procurement systems) as the sample of countries for whom data on the quality of systems are available is too small to allow for meaningful analysis.

d. The targets shown may differ from indicative targets published in previous years as a result of adjustments to historical data (e.g. indicator 8, where final data on tying led to adjustments to the underlying datasets after publication of reports on the previous surveys). The target for indicator 5a (use of country PFM systems) has been computed to consider the 2010 scores on the quality of PFM systems (indicator 2a), consistent with the approach agreed in the Paris Declaration and described in Chapter 3.

Source: (OECD, 2012, p. 19)
principles and do not actually investigate the effect of such implementation of the aid effectiveness principles at the aid intervention level. An independent evaluation of these surveys (Wood et al., 2011), however, claimed that there is growing evidence that implementing the Paris principles is providing better results in broad terms in the health sector.

As seen in the systematic review section, the number of published works on actual evaluation of the implementation of the Paris principles is very limited. Dodd and Olive (2011) attempted to study to what extent programmes and approaches in the health sector in Vietnam adhered to the Paris principles. Although their study identified barriers to such adherence, it did not evaluate the effect of the adherence to the Paris principles on health results.

The Working Party on Aid Effectiveness of OECD also undertook an evaluation of the implementation of the Paris Declaration, particularly in the health sector, using a series of country level evaluations. In their synthesis report of the evaluation of the Paris principles in the health sector, the Task Team of the Working Party on Aid Effectiveness, however, acknowledged the challenges of measuring and attributing the impact of aid effectiveness principles on health outcomes saying that “demonstrating and attributing the impact on aid effectiveness is challenging and there is no common understanding of what results can realistically be expected or how these will be measured” (OECD Working Party on Aid Effectiveness Task Team on Health as a Tracer Sector, 2011, p. 10).

A few other writers also attempted to find the pathway between the implementation of aid effectiveness principles and development results. An analysis by Dickinson (2011) suggested that aid effectiveness is contributing to development through creating conditions for sustainable impact:

“There is evidence that aid effectiveness is improving sector planning, budgeting, strengthening national systems and increasing resource allocations. More efficient funding of the health sector, through programme based approaches including SWApS is helping implement health sector reforms, which are contributing to better health results” (p.9).

Proposing an analytical framework based on the logical chain between implementation process, health system strengthening, and health outcomes in the context of monitoring
the results of adherence to the principles of aid effectiveness in the health sector in Mali, Paul et al. (2013) claimed that:

“We cannot prove a linear, causal link between implementation of the Paris principles (level 1), HSS [health system strengthening] (level 2), and health outcomes (level 3) but qualitative methods have helped us identify the most plausible factors that can explain results” (p.148).

This suggests a plausible impact pathway between the implementation of the Paris principles and health outcomes in the case of Mali.

Apart from these studies, the Quality of Official Development Report (Birdsall & Kharas, 2014; Birdsall, Kharas, & Perakis, 2009, 2011), developed by the Center for Global Development, measured donor performance across 31 indicators of aid quality to which donors have made commitments. These indicators were grouped into four dimensions including maximizing efficiency, fostering institutions, reducing the burden on partner countries, and transparency and learning associated with effective aid. However, like other studies this quality index did not show the health related outcomes of such quality of aid.

2.5.6 Methods and tools of comparative analysis

Theory driven programme evaluation such as the realist evaluation often employs both qualitative and quantitative research methods. It has been argued that such mixed methods approaches can be particularly useful in healthcare research (Bryman, 2007; Clarke & Yaros, 1988; Foss & Ellefsen, 2002; Steckler et al., 1992). As defined by Tashakkori and Creswell (2007), mixed methods research is “research in which the investigator collects and analyzes data, integrates the findings and draws inferences using both qualitative and quantitative approaches” (p.3). While evaluations often use descriptive and interpretative qualitative techniques, a number of quantitative techniques including balanced scorecard, cost-effectiveness analysis, and DEA are also used in operations research, health economics, and health care performance evaluation to help managers and policy makers in making informed decisions. As Ostlund et al. (2011) identify,

“Interest in, and expansion of, the use of mixed methods designs have most recently been fuelled by pragmatic issues: the increasing demand for cost effective research and the move away from theoretically driven research to research which meets policymakers’ and practitioners’ needs” (p.370).
**Balanced scorecards:**

Originated by Kaplan and Norton (1992, 1997, 2001), balanced scorecard is a performance measurement approach focusing on the corporate strategy from four perspectives: financial, customer, internal process, learning and growth. Balanced scorecards derive a composite performance score for each perspective by developing a set of key indicators for each and presenting the current status of those indicators compared with their agreed benchmarks.

The four perspectives in the balanced scorecard suggested by Kaplan and Norton (1992) were modified to adapt the approach as a tool for assessing the effectiveness of various application domains (Fairchild, 2002; Grebergen & Amelinckx, 2002; Martinsons, Davison, & Tse, 1999). For example, Inamdar, Kaplan, and Bower (2001) suggested the use of balanced scorecards for performance measurement in healthcare provider organisations by aligning the perspectives to the well-defined vision, mission, and strategies of the organization. A number of other researchers also found it useful and convenient to use the balanced scorecards method to evaluate and compare performance of health care systems (Chan et al., 2010; Chen et al., 2006; Commonwealth Fund Commission on a High Performance Health System, 2008; Gauld et al., 2011; Hansen et al., 2008; Peters et al., 2007; Pink et al., 2001; Schoen et al., 2006; Ten Asbroek et al., 2004; Voelker, Rakich, & French, 2001; Zelman, Pink, & Matthias, 2003).

In the development field, using the balanced scorecards concept, Niven (2011) proposed a step-by-step guide for government and nonprofit agencies as a useful communication tool, measurement system, and strategic management system. However, the balanced scorecards approach has not yet been adapted to measure and compare the performance of implementation of the aid effectiveness principles and development outcomes of aid interventions.

**Cost effectiveness analysis:**

Cost effectiveness analysis is increasingly being used as one of the useful components of impact evaluations of health programmes (Bouillon & Tejerina, 2006; Fiszbein & Schady, 2009; Holla & Kremer, 2009; Rawlings & Rubio, 2005; Zwane & Kremer, 2007). A number of researchers found it useful to use cost effectiveness analysis for comparative
effectiveness analysis of different health programmes or strategies for informed resource allocation decisions (Drummond et al., 2005; Gold et al., 1996; Jamison et al., 2006a, 2006d; Jamison et al., 1993; Levin & McEwan, 2001; Miguel & Kremer, 2004).

Cost-effectiveness analysis attempts to calculate the incremental cost per effectiveness unit for an intervention and compare this ratio with that of other options to see which option would be more cost-effective. For health interventions, the effectiveness is often measured by calculating DALYs saved by the intervention compared to a no-intervention scenario (McCabe, 2009).

As described in Musgrove and Fox-Rushby (2006), DALYs primarily measure disease and injury burden in terms of total number of years lost from the lives of a population due to disability and death from an illness compared to the total number of years the population would live in a perfect health condition. As explained by WHO (2013a) for the methods and data sources for global burden of disease estimates 2000-2011, DALYs for a specific cause are calculated as the sum of the Years of Life Lost (YLL) from that cause and the Years Lived in Disability (YLD) for people living in states of less than good health resulting from the specific cause:

\[
\text{DALY} = \text{YLL} + \text{YLD}
\]

The YLDs are calculated as the product of number of cases averted multiplied by the disability weight (DW) for particular disease sequelae (Vos et al., 2012).

The cost per DALY saved by an intervention can be compared with the cost per DALY saved by different interventions targeting the same cause to compare cost-effectiveness. The Disease Control Priority Project (Jamison et al., 2006a) provides examples of the most effective (“best buy”) interventions including examples from Africa and South Asia (Laxminarayan & Ashford, 2008).

Cost effectiveness of an intervention is also compared with the per capita GDP of a country. According to the WHO’s Commission on Macro-economics and Health (Walker, Hutubessy, & Beutels, 2010), an intervention that gains a year of healthy life (i.e., a DALY averted) at a cost that is less than GDP per capita is “very cost-effective”. An intervention is “cost-effective” if the cost per DALY averted by the intervention is within the range of
GDP per capita to three times of GDP per capita. Interventions whose costs per DALY averted are more than three times of the GDP per capita are not considered as cost-effective.

Data Envelopment Analysis (DEA):
Moreno-Sera and Smith (2012) believe that “inefficiency is intrinsically difficult to measure, as it represents the shortfall in performance from what could in principle be achieved, a concept that is manifestly open to challenge” (p.2). In the health sector, inefficiency is often inferred by constructing an estimate of the ‘health production frontier’ on the basis of the observed best performance amongst exemplar units of observation. The efficiency of a unit under evaluation is then measured against the best performing ‘health production frontier’ (Jacobs, Smith, & Street, 2006).

There have been two different approaches to constructing the best performance ‘health production frontier’: 1) the econometric methods that use various statistical techniques such as panel data models and stochastic frontier analysis (SFA), and, 2) the descriptive methods known as DEA (Moreno-Serra & Smith, 2012). Both these methods intend to use the observed behaviour of all production units to infer the maximum feasible level of attainment by a production unit (i.e. the production units that form the health production frontier) and try to estimate to what extent a production unit under evaluation falls short of that optimum level of production.

Based on the work of Farrel (1957), Charnes et al. (1978), Banker et al. (1984), DEA is a model that evaluates the relative efficiency of a production or “decision making unit” (DMU), within a context with multiple inputs and outputs. DEA identifies a subset of efficient "best-practices" to construct a virtual efficiency frontier. The efficiency of a DMU is then measured by comparing its performance to that virtual efficiency frontier.

For many years researchers have been using DEA techniques to evaluate efficiency and effectiveness in the healthcare sector. For example Banker, Conrad and Strauss (1986), Grosskopf and Valdmanis (1987) and Byrnes and Valdamanis (1996) used DEA for studying hospital efficiency. Such studies consider inputs such as the labor force (e.g. physician, nurse, paramedic, manager, and support staff), materials (e.g. medication,
nondrug materials, and other product related to health service delivery), and capital (e.g. land, building, medical equipment, vehicle, and bed) used by a hospital to produce outputs such as the quantity of emergency, inpatient, and outpatient services provided by the hospital. Considering the multiple inputs and multiple outputs, the DEA calculates technical efficiency of each hospital by comparing them with their best performing peers.

More recent studies include Zere et al. (2006), Aletras et al. (2007), Kirigia et al. (2007), Masiye et al. (2007), Nayar and Ozcan (2008), Hollingsworth (2008), Akazili et al. (2008), Sebastian and Lemma (2010), Sheikhzadeh et al. (2012), Guerra et al. (2012), Jehu-Appiah et al. (2014) and Kawaguchi et al. (2014) that have used DEA techniques to compare the performance of public health centers and healthcare systems in a wide range of countries including Namibia, Greece, Seychelles, Zambia, US, Ghana, Ethiopia, Scotland, Brazil and Japan.

2.6 Critical appraisal of studies dealing with impact of the Paris principles on health outcomes

As seen in section 2.4, the number of studies that actually attempted to see the impact of the compliance with the Paris principles on the health intervention outcomes is very limited. Table 2.5 presents a summary of critical appraisal of such studies from the sampled literature:
### Table 2.5: Critical summary of studies dealing with impact of the Paris principles

<table>
<thead>
<tr>
<th>Study</th>
<th>Summary</th>
<th>Methodology and limitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Camacho and Batista (2008)</td>
<td>This article claims that aiming at adherence to the Paris principles Spanish aid was promoting sector wide approach in its choice of aid instruments and saw better results compared to aid disbursed through other instruments. The article, however, noted that major changes were needed at donor and recipient levels for full implementation of the Paris Declaration.</td>
<td>This is a review article in Spanish language with abstract available in English. It is not clear how the claim of achieving better results through the SWAp than other aid modalities was substantiated with robust research methods and empirical evidence. Due to the language barrier, the methods followed in this article could not be assessed.</td>
</tr>
<tr>
<td>Dalil et al. (2014)</td>
<td>Using the Paris principles, this article assessed the history of aid effectiveness in Afghanistan’s health sector as the principles were followed by the donors under a strong stewardship by the Ministry of Public Health facilitated by a Grants and Contracts Management Unit established within the Ministry with donor funding. The article marked visible progress in the health sectors with improved health indicators, increased local capacity to run the health systems, sustained coordination between the government and donors, and stronger leadership and stewardship role by the Ministry of Public Health. The article claims that donor aid to Afghanistan appeared to be effective in supporting health gains because of increased adherence to the Paris principles, partnerships with the NGOs in health service delivery, using community based approach to extend access to health services, focusing on both public and private sector to build the entire health sector, and availability of skilled and dedicated technical expertise and human resources for health.</td>
<td>The article uses a qualitative historical analysis of the developments in the Afghanistan’s health sector. It attempts to link the process of implementation of aid effectiveness principles with the outcomes in the health sector by comparing the historical evidences (“before” and “after” comparison). However, the article focuses on the outcomes at the sectoral level and the pathways and contributions of the claimed factors are not analyzed or tested to substantiate the claims.</td>
</tr>
<tr>
<td>Dickinson (2011)</td>
<td>This paper proposes a programme theory based framework for analyzing the underlying results chain at the health sector level to see how the process of adhering to the Paris principles of aid effectiveness contributes to the health results. Using this framework, the study conducted a rapid synthesis of the evidence from selective studies related to the implementation of SWAps in the health sector. The paper claims that adherence to the Paris principles contributes to better health results by creating favourable conditions for sustainable impact such as improved sector planning and budgeting, strengthened national systems, and increased resource allocations.</td>
<td>This paper used a theory based analysis and synthesis approach to see the process and effect of the Paris principles in the health sector. However, the paper only used selected evaluations of SWAps in its synthesis and did not consider other aid modalities. As the selected studies mainly focused on the review of SWAp process, the claim of conditions for better health results remains as logical inference without any substantiated evidence of better results.</td>
</tr>
<tr>
<td>Dodd et al. (2009)</td>
<td>Based on a case study research in Lao PDR, this paper examines the potential of increased compliance with the Paris principles for improving the practice of human resources management in the health sector. The paper claims that the initiative of implementing the Paris principles improved the dialogue, partnerships, and coordination among the government and development partners in Laos for improved human resources management practices. However, the paper noted that working out the details of operationalization of the</td>
<td>This research used a case study approach using mixed research methods with both qualitative and quantitative analysis. However, the focus on this research is limited to the applicability of the Paris principles in the human resources management in the health sector. The research did not examine the effect of the Paris principles on the outcomes of particular health interventions.</td>
</tr>
</tbody>
</table>
principles of aid effectiveness in human resources management for health may threaten the broad consensus on the concept and rhetoric of the Paris principles.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Description</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dybul (2009)</td>
<td>This article provides an account of success of the PEPFAR programme. The article claims that PEPFAR followed the internationally agreed principles of aid effectiveness including the Paris principles. In promoting country ownership, PEPFAR undertook an approach of supporting the people of a country-- not just the government-- in tackling their problems. The article claims that with the philosophical foundation of country ownership and results based management, adequate resources, and a strong intellectual base, PEPFAR saved lives and provided lessons learned for effective development.</td>
<td></td>
</tr>
<tr>
<td>This article is mainly based on the author’s views and discussion. Although this articles referred to a few other studies to support its claims, it did not employ any methodological approach to substantiate the claims.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Martini et al. (2011)</td>
<td>This work indicates mixed results of the implementation of the Paris Declaration. It claims that good progress has been achieved in the health sector in terms of country ownership and coordination. However, recipients and donors lag behind in terms of alignment and use of country systems, managing for results, and mutual accountability. However, this work warns about the limitation of the evaluation methods that made it difficult to generalize and draw a causal link between the Paris principles and development effectiveness.</td>
<td></td>
</tr>
<tr>
<td>This work was presented as a Conference poster/abstract. It claims that it is based on a critical review of all the aid effectiveness surveys and evaluations published since 2005 by the OECD and IHP+. However, this work is based on the secondary sources and due to the wide variety of contexts and method, it could not come to any common definitive conclusion.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OECD Working Party on Aid Effectiveness Task Team on Health as a Tracer Sector (2011)</td>
<td>This is the final report of the OECD Working Party on Aid Effectiveness Task Team on Health as a Tracer Sector (TT HATS). It provides an overview of the progress of implementing the Paris principles in the health sector. While the report saw progress on the process indicators for implementation of the Paris principles, its marked limited evidence base on the link between the implementation of the principles of aid effectiveness and health results.</td>
<td></td>
</tr>
<tr>
<td>The report used a qualitative and synthesis approach based on an interim report, the outputs of the TT HATS work programme, the second phase evaluation of the Paris Declaration, the IHP+ results progress monitoring, review of related literature review, and interviews. However, the report provided a high level sectoral view and did not employ any agreed framework for evaluating the impact of aid effectiveness principles on health outcomes. As a result, its findings suffered from limited evidence base and possible confounding factors making it difficult to draw any causal link between the principles of aid effectiveness and health outcomes.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paul et al. (2009)</td>
<td>This work draws on frameworks for evaluation of Paris Declaration and IHP+ and author’s experience in documenting aid effectiveness results from implementation of a SWAp in Mali’s health sector. The work proposes a framework for evaluating aid effectiveness at three levels: first, at the process of implementation of Paris Declaration and change of stakeholder behaviour, second, how the donor support and compliance with the Paris principles contributed to health systems strengthening and service delivery, and third, at the health outcome level.</td>
<td></td>
</tr>
<tr>
<td>This is a conference presentation in line with theory driven evaluation. However, the focus is again at the sectoral level considering only the SWAp modality as a case of analysis.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paul et al. (2013)</td>
<td>This is the full paper based on aid effectiveness evaluation framework proposed in Paul et al. (2009) mentioned above and case studies conducted in Mali for implementation of a health</td>
<td>This paper used case studies methods with analysis of documents and statistics, interviews with a wide variety of stakeholders in the MoH (MoH) and</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Source: Author’s review</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As the above table shows, most of the studies mainly employed a theory-driven qualitative approach to see the results chain of the implementation of the Paris

<table>
<thead>
<tr>
<th>Table</th>
</tr>
</thead>
<tbody>
<tr>
<td>SWAp. Authors argue that the application of the proposed framework could not prove a linear causal link between implementation of the Paris principles (level 1), health systems strengthening (level 2), and health outcomes (level 3) but the qualitative methods helped to identify the most plausible factors that can explain results. These factors included scaling-up of high-impact strategies in health; improving access to health services; better availability of and access to medicines and health products; and increasing resources allocated to disadvantaged regions.</td>
</tr>
<tr>
<td>Taylor et al. (2013)</td>
</tr>
<tr>
<td>Wood and Betts (2013)</td>
</tr>
<tr>
<td>Wood et al. (2011)</td>
</tr>
</tbody>
</table>

The study used a meta-analysis and systematic review approach. However, the approach depends on the secondary sources and since the number of studies that actually track the outcomes of Paris principles is very limited, the approach was not successful in actually evaluating the outcomes of the Paris principles in maternal and reproductive health. Also the distinction between the Paris style aid and general aid has taken as dichotomous other than a continuum. The findings presented in this paper were derived through a participatory, theory-based evaluation approach using mixed methods across 21 country evaluations and 7 donor studies undertaken by OECD. Findings of all these studies were systematically analyzed and synthesized. However, the synthesis did not specifically focused on specific intervention outcomes and due to the methodological limitations and unavailability of data, any definitive generalizations were not possible. The evaluation mainly used qualitative technique with multiple sources of evidence and techniques including drawing upon some reliable quantitative data, where available. However, the focus of evaluation was mainly on the journey and progress of implementation of the Paris Declaration at the country level. Due to limited time of implementation of the Paris Declaration and methodological limitations, the evaluation could not trace results for development.
principles at the country and sectoral level and acknowledged the limitations of the methods and available evidence to draw any definitive conclusion. The studies in the sample mostly confined their evaluation within the SWAp modality grossly assuming that a SWAp readily complies with the Paris principles of aid effectiveness. Moreover, the studies attempted to differentiate the aid modalities that are believed to be following the Paris principles from the other modalities perceived as not following the Paris principles in a dichotomous way.

However, in reality a SWAp may not be implemented by fully complying with the Paris principles and complying with the Paris principles by different aid modalities can be more appropriately measured by their degree of compliance on various aspects of the Paris principles in a continuum than by a yes/no dichotomy. As the synthesis of the Paris Declaration Evaluation points out:

“With respect to aid modalities, the evaluation found that no single modality (e.g., budget or sector support, program[me]s, or projects) will automatically produce better development results, and a mix of aid modalities has continued to make sense for partner countries and donors” (Wood & Betts, 2013, p. 113).

The methodology of the Paris Declaration Evaluation (Betts & Wood, 2013) acknowledged the critical importance of the context and the fact that aid is one of the many contributing factors of development results. Since the above studies focused on the implementation of the Paris principles at the broader sectoral level, their synthesis actually missed the opportunity to consider the effect of the context and other confounding factors to analyze the development results at the intervention level.
2.7 Conclusion

The foregoing mapping of aid effectiveness literature, systematic review, and scoping show that although there is a sizable literature on aid effectiveness and various aspects of the implementation of the Paris principles in the health sector, none of these studies actually evaluates the effect of implementing the Paris principles at the health intervention level in a fragile state context. Grap-Pa Sante (2011) also pointed out this “notable dearth of rigorous analysis and country studies that look at the impact on health results, or that use the Paris principles as the starting point for analysis of results.” This review of related literature provides a theoretical background to establish the broad concepts, and framework and specific methods and tools used in this research. More specific discussion of the literature mentioned in this chapter can be found in other relevant chapters of this dissertation.
3 Country context

3.1 Introduction

For the purpose of having a basic understanding of the country context where interventions took place, this chapter presents the country profile of Timor-Leste in terms of its geography, people and culture, history, economy, politics, and the country’s performance in human development indicators including achievement of MDGs. Information presented in this chapter is based on a synthesis of the Government of Timor-Leste’s published documents; country profiles presented in the websites of WHO, USAID, AusAID, World Bank, CIA, and UNDP; and related published articles on Timor-Leste. In presenting these profiles this chapter also attempts to capture the contextual changes to see the trend of changes, especially during the period from 2007 to 2013, as the wider contextual analysis for Timor-Leste’s health sector.

3.2 Geography

The Democratic Republic of Timor-Leste, formerly known as East Timor is the youngest country in the southeastern Asia. The word ‘Timor’ came from Malay word meaning ‘East’. Timor-Leste falls at the eastern end of the Malay Archipelago and northwest of Australia. The Island of Timor is divided in two parts—the western part is known as West Timor and belongs to the Republic of Indonesia with its provincial capital in Kupang. The eastern part of the Timor Island belongs to Timor-Leste with its capital in Dili.
Timor-Leste has a total land area of 15,007 square kilometers that includes the eastern half of the island of Timor, Atauro Island located north of the capital Dili, Jaco Island located on the easternmost end of the island and the enclave of Oecussi situated within Indonesia on the northwestern side of the island. With a mountainous terrain, the country spans approximately 180 kilometers East-West and 75 kilometers North-South.

As a tropical country, the climate of Timor-Leste is mostly hot and humid with two distinct seasons—the rainy season and the dry season.

Timor-Leste is the eastern end point of the Malay Archipelago but geographically the island has been considered with high strategic importance due to its close proximity to Australia and due the location of Timor Sea, which is rich with oil and gas reserve. In fact the strategic geographic location of Timor-Leste was one of the reasons why the Dutch and Australian Allies fought with the Imperial Japanese force during the Second World War over the control of East Timor. For the ownership and control of the oil and gas reserve in the Timor Sea, Timor-Leste still has disputes with Australia over the demarcation of the sea border and sharing of oil and gas resources from the Timor Sea (Forbes, 2010).
### 3.3 People and culture

The people of Timor-Leste are known as Timorese. Ethnically Timorese population largely consists of Austronesian (Malayo-Polynesian), Papuan, and Chinese minority.

According to the National Statistics Directorate of Timor-Leste, the population of Timor-Leste in 2004 was 923,198 with an annual population growth rate of 5.3%. The 2010 Census showed that the population of Timor-Leste was 1,066,409 and the annual population growth rate was 2.4%. 41.4% of the population was below the age of 15 years (National Statistics Directorate of Timor-Leste, 2010). Table 4.1 provides a summary of population distribution in Timor-Leste in 2004 and 2010:

<table>
<thead>
<tr>
<th>Age Group</th>
<th>2004</th>
<th>2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-14 years</td>
<td>43.2%</td>
<td>41.4%</td>
</tr>
<tr>
<td>15-64 years</td>
<td>53.3%</td>
<td>53.9%</td>
</tr>
<tr>
<td>65+ years</td>
<td>3.5%</td>
<td>4.7%</td>
</tr>
<tr>
<td>Total Population</td>
<td>923,198</td>
<td>1,066,409</td>
</tr>
</tbody>
</table>

Data Source: (National Statistics Directorate of Timor-Leste, 2011a)

According to the World Fact Book of CIA, the 2013 population of Timor-Leste was estimated as 1,172,390 with the following distribution:

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Percentage</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-14 years</td>
<td>42.7%</td>
<td>257,340</td>
<td>243,174</td>
</tr>
<tr>
<td>15-24 years</td>
<td>19.7%</td>
<td>116,605</td>
<td>114,203</td>
</tr>
<tr>
<td>25-54 years</td>
<td>29.3%</td>
<td>166,048</td>
<td>177,024</td>
</tr>
<tr>
<td>55-64 years</td>
<td>4.8%</td>
<td>28,717</td>
<td>27,011</td>
</tr>
<tr>
<td>65 years and over</td>
<td>3.6%</td>
<td>20,428</td>
<td>21,840</td>
</tr>
<tr>
<td>Total</td>
<td>100%</td>
<td>589,138</td>
<td>583,252</td>
</tr>
</tbody>
</table>


It can be seen that an estimated 62.4% of the population is under the age of 25 years with an estimated median age of the population of 18.4 years. With an estimated 2.47% population growth rate, this very young population for the country means that Timor-Leste enjoys a substantial demographic dividend of having a constant supply of a productive work force. However, capitalizing on this advantage requires population planning, job creation, and investments in skills development to meet the needs of the labour market.
28.3% of Timor-Leste's population lives in the urban area with a slow urbanization rate of 4.25%. Most of the urban population of the country lives in the two largest urban centres of Dili and Baucau. Populations in rural areas often live in small, dispersed villages isolated by mountainous terrain and poor roads. Lack of roads and transportation facilities pose challenges for the rural population to access services at health care facilities (Ministry of Health of Timor-Leste, 2011n).

3.4 History

The history of Timor-Leste is marked with colonization, occupation, resistance and conflict. Traditionally the Timor Island was organized in small states ruled by two kingdoms known as Sorbians and Belos with people mostly holding indigenous spiritual beliefs (‘animism’) without following any organized major religion.

The Portuguese missionaries and traders started arriving the Timor Island in the early 16th century and colonized the island in mid 16th century. During mid 17th century the Dutch invaded the western end of Timor and took control over half of the island. Following a series of skirmishes between the Dutch and Portuguese, the two colonizing forces eventually reached a treaty in 1859 in which Portuguese withdrew from the western portion of the island and gave it to the Dutch and maintained the eastern portion of the island along with the Oeccusse enclave, Atauro island and Jaco island as their colony. The border between the Dutch Timor and Portuguese Timor still forms the basis for the border between the West Timor (now belonging to Indonesia) and East Timor (Timor-Leste).

During the Second World War, Imperial Japan occupied Timor from 1942 to 1945 and fought with the Australian and Dutch allied forces. Some tens of thousands of Timorese people lost their lives by fighting side by side with the Allies against the Japanese forces. In 1945, upon Japan's defeat in the Second World War and Indonesian independence, West Timor was integrated into Indonesian territory. However Portugal resumed their colonial authority over East Timor and maintained their administration in Timor-Leste until 1974 as one of the Portuguese provinces defying all international resolutions for the right to self-determination by the people of Timor-Leste.
However, with the dissolution of the Portuguese colonial empire in 1975, local liberation movements in East Timor further increased. On 28 November 1975 East Timor unilaterally declared itself independent from Portugal. This led to a civil war and nine days later Indonesian forces invaded East Timor. Through this invasion, Indonesia occupied East Timor and annexed this territory as their 27th province renaming it as Timor Timur in July 1976.

Following the Indonesian occupation, an unsuccessful campaign of pacification followed over the next two decades by the Indonesian authority but Timorese resistance movements became further consolidated during this time. It is estimated that nearly one quarter of the population died during the occupation as a consequence of conflict, forced migration, malnutrition and unattended public health needs (Ministry of Health of Timor-Leste, 2011n). According to a UNESCO estimation, “During the Indonesian occupation, some 200,000 people are believed to have died from war and famine. In the months following the invasion, some 60,000 were killed and those first years saw the establishment of nearly 150 concentration camps” (UNESCO, 2009, p. 5).

On 30 August 1999 a popular referendum was held in East Timor to determine if East Timorese would like independence or integration into Indonesia. The referendum was participated in by more than 90% East Timorese and 78.5% favored independence. However, this referendum result was followed by a large-scale violence organized by the anti-independence Timorese militias backed up by the Indonesian military, which were dissatisfied with the referendum result (Government of Timor-Leste, 2013).

The militias killed approximately 1,400 Timorese and forcibly pushed 300,000 people into western Timor as refugees. The violence saw about 70% of the country's total infrastructure including homes, water supply systems, irrigation systems, and schools destroyed. Nearly the whole of the country’s electrical grid was destroyed, 75% of the people were internally displaced, and 80% of health infrastructure and equipment destroyed in the three-week post-referendum violence.
On 18-20 September 1999, the United Nations (UN) led international peacekeeping troops were deployed in East Timor with the aim to disarm the militias and support the transition process and reconstruction of the country. From 20 September 1999 to 20 May 2002 Timor-Leste was supported by the UN Administration for East Timor (UNAET). With the progressive disarming of militias and beginning of the reconstruction of houses, schools and other infrastructure, the situation was slowly brought under control. Elections were also held for a Constituent Assembly that became responsible for drafting the Constitution of Timor-Leste. Finally on 20 May 2002, Timor-Leste's Constitution came into force giving the country its sovereignty as an independent state.


Following the recognition as an independent state, Timor-Leste experienced several bouts of internal conflict and violence in its brief post-independence history. The country fell back to a heightened level of tension and security risks when in 2006 a military strike led to a fresh violence and breakdown of law and order. As a result, around 150,000 people were internally displaced and were forced to live in Internally Displaced People (IDP) camps with sub-optimal facilities in Dili and surrounding areas. At the request of the Government of Timor-Leste, an International Stabilization Force (ISF) was deployed in Timor-Leste in 2006. The United Nation’s Security Council established the UN Integrated Mission in Timor-Leste (UNMIT) with authorized presence of UN Police personnel in Timor-Leste to support the local police to maintain law, order and stability.

The ISF and UNMIT along with the Timor Government restored the security situation that allowed successful presidential and parliamentary elections in 2007 and handing over power to the new constitutional government. However, the security situation deteriorated again in February 2008 when a rebel group, dismissed from the military in 2006, staged an attack against the President and the Prime Minister. In that attempt the President was shot and was seriously wounded. However, the rebel ringleader was killed in the attack and by April 2008 most of the rebels surrendered themselves to the Government.
Apart from occasional reports of crimes Timor-Leste has been enjoying stability over the last five years. The country successfully organized the Presidential and Parliamentary elections in 2012 and formed the new constitutional government. The UN Security Council voted to end peacekeeping mission in Timor-Leste and by the end of 2012, and both ISF and UNMIT departed the country leaving Timor-Leste to stand on their own feet (Central Intelligence Agency, 2013; World Bank, 2012d).

3.5 Economy

Timor-Leste is considered as one of the “Least Developed Countries” (UNCTAD, 2014). A per capita gross national income (GNI PPP, 2005 international dollar term) of $5,446 places Timor-Leste in the lower-middle income country category according to the World Bank. However, over 90% of state revenue is from the Petroleum Fund comprising of income from the oil and gas resources (Ministry of Finance of Timor-Leste, 2012). The non-oil per capita GDP was only US$810 in 2011. The actual living standard of most people in Timor-Leste is marked with high poverty and low achievement in human development indicators.

Timor-Leste is fortunate to have an offshore reserve of petroleum resources in the Timor Sea. The country depends on international petroleum companies to extract and sell non-renewable oil resources through cost recovery and a revenue sharing basis. In order to manage the income from exporting non-renewable petroleum wealth, Timor-Leste established a Petroleum Fund through a law approved in 2005 and revised in 2011 (McKechnie, 2013).

The Petroleum Fund is operationally managed by the Central Bank of Timor-Leste with specific guidelines on investments, “Estimated Sustainable Income”, fund transfer to the government, transparency and accountability. The balance of the Petroleum Fund at the end of December 2013 stood at US$ 14.9 billion (Petroleum Fund Administration Unit, 2014). The petroleum sector represents 76.9% of the country’s total GDP.

Thanks to increasing withdrawal from the Petroleum Fund, Timor-Leste government’s annual budget increased from $70 million in 2004 to $667.9 million in 2008 to $1,451.2 million in 2012. The increased expenditure was mainly for construction of roads and
buildings, producing electricity, providing social benefits to veterans and increased salaries and wages of government employees.

Among non-oil exports, Coffee represents the single largest export commodity giving the country approximately US$20 million a year. Table 4.3 and Figure 4.2 below provide Timor-Leste's actual and projected annual budget and sources of financing from 2008 to 2014:

**Table 3.3: Timor-Leste's annual budgets and sources of financing: 2008-2014 (in Million US Dollars)**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Domestic non-oil revenues</td>
<td>69.4</td>
<td>90.8</td>
<td>96.6</td>
<td>111.7</td>
<td>137.7</td>
<td>146.3</td>
<td>166.1</td>
</tr>
<tr>
<td>Financing from Petroleum Fund</td>
<td>414.5</td>
<td>512.8</td>
<td>663.7</td>
<td>985.4</td>
<td>1,016.8</td>
<td>1,457.6</td>
<td>1,282.9</td>
</tr>
<tr>
<td>Borrowing</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>43.1</td>
<td>43.6</td>
<td>51.0</td>
</tr>
<tr>
<td>Estimated ODA</td>
<td>184.0</td>
<td>221.0</td>
<td>263.9</td>
<td>283.9</td>
<td>253.6</td>
<td>203.4</td>
<td>117.9</td>
</tr>
<tr>
<td>Total Annual Budget</td>
<td>667.9</td>
<td>824.6</td>
<td>1,024.2</td>
<td>1,381.0</td>
<td>1,451.2</td>
<td>1,850.9</td>
<td>1,617.9</td>
</tr>
</tbody>
</table>


**Figure 3.2: Timor-Leste's annual budget and sources of financing 2008-2014 (in Million US Dollars)**

It can be seen that during the period from 2008 to 2014 Timor-Leste’s domestic revenues grew consistently at almost double-digit rates. ODA to Timor-Leste also grew from $184 million in 2008 to $283.9 million in 2011 and then started declining from 2012. Still the government had to finance most of the annual budget by substantially withdrawing from the Petroleum Fund beyond its estimated sustainable income level. Since the Petroleum Fund is dependent on the non-renewable natural resources, it is projected that the income of the Petroleum Fund will start declining from 2016 as the reserves from the oil and gas fields currently under production would be exhausted by 2025 (Scheiner, 2014). Therefore, the young country has the challenge ahead to create a non-oil dependent economy by creating sufficient employment and other economic activities.

3.6 Politics

The Democratic Republic of Timor-Leste follows a semi-presidential and a parliamentary system of government. The state executives are the President (head of state), the Prime Minister (head of government) and the Council of Ministers and the Superior Council of Defense and Security.

The head of the state is the President of the Republic, who is elected by popular vote for a five-year term. The president appoints the prime minister, protects the Constitution and State institutions and can mediate conflict resolution if necessary. The President also has the right to veto any legislation put forth by the government and approved by the parliament, dissolve the National Parliament and call national elections.

The head of the government, on the other hand, is the Prime Minister. After a legislative election, the President appoints the leader of the majority party or majority coalition as the Prime Minister to head the government. The Government is the Executive body of the State and is responsible for the development and implementation of the Government Programmes including the national budget. Like the President, the government is also appointed for a five-year term.

Timor-Leste’s Legislative is unicameral (single chamber) and called Parlamento Nacional (the National Parliament) with a number of seats varying from 52 to 65.
Members of the National Parliament are also elected by popular vote for a five-year term from the participating political parties on a modified representational basis.

Timor-Leste’s Judiciary is still developing. It consists of courts and supporting hierarchy with a civil law system based on Portuguese model. A Court of Appeal is functioning as the Supreme Court on an interim basis. Timor-Leste is administratively organized into 13 districts, 66 sub-districts, and 192 sucos (villages).

### 3.7 Millennium Development Goals (MDGs)

Despite rapid economic expansion and reconstruction of the health systems, human development outcomes in Timor-Leste still remain low. According to the UNDP Human Development Index (HDI), Timor-Leste’s HDI value for 2013 was 0.620— which was in the medium human development category— positioning the country at 128 out of 187 countries and territories. “Between 2000 and 2013, Timor-Leste’s HDI value increased from 0.465 to 0.620, an increase of 33.4 percent or an average annual increase of about 2.24 percent” (UNDP, 2014).

Table 3.4 summarizes Timor-Leste’s status and the trends of achievements in major human development indicators in 2000, 2008 and 2012.

**Table 3.4: Timor-Leste’s Human Development Indicators**

<table>
<thead>
<tr>
<th>Indicator</th>
<th>2000</th>
<th>2008</th>
<th>2012</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>830,100</td>
<td>1,079,700</td>
<td>1,187,200</td>
</tr>
<tr>
<td>Life expectancy at birth</td>
<td>56.2</td>
<td>61.1</td>
<td>62.9</td>
</tr>
<tr>
<td>Maternal mortality ratio</td>
<td>660 per 100,000 live births</td>
<td>557 (2009-10 DHS)</td>
<td></td>
</tr>
<tr>
<td>Infant mortality rate</td>
<td>100</td>
<td>75</td>
<td>45</td>
</tr>
<tr>
<td>Total fertility rate</td>
<td>7.1</td>
<td>6.5</td>
<td>6</td>
</tr>
<tr>
<td>Adult literacy rate</td>
<td></td>
<td></td>
<td>58.3%</td>
</tr>
<tr>
<td>Population living below poverty line</td>
<td>41% (below $0.55/day)</td>
<td>50% (below $0.88/day)</td>
<td>37.4% (below $1.25/day)</td>
</tr>
<tr>
<td>HDI</td>
<td>0.418</td>
<td>0.547</td>
<td>0.576</td>
</tr>
<tr>
<td>HDI ranking</td>
<td>162 of 182 countries</td>
<td>134 of 187 countries</td>
<td></td>
</tr>
</tbody>
</table>

Source: (Central Intelligence Agency, 2013; Ministry of Health of Timor-Leste, 2011n; UNDP in Timor-Leste, 2013; WHO, 2013h)

Although Timor-Leste has made some good progress in terms of increasing life expectancy at birth and decreasing maternal mortality, infant mortality and total fertility rate, the country has a long way to go to achieve the MDGs.
Since achievement of MDGs are closely interlinked to and mutually dependent on each other, this research attempts to see the context and performance of the health sector in relation to all MDGs.

3.7.1 MDG 1: Eradicate extreme poverty and hunger

In 2004, Timor-Leste set the national target of reducing poverty to 14% by 2015. However, poverty in Timor-Leste increased from the year 2000 when 41% of the population lived below the national poverty line of $0.55 a day to nearly 50% in 2008 when half of the population lived below the national poverty line of $0.88 a day.

Thanks to post 2007 growth of economy, some cash transfer, and social benefits programmes of the government, the poverty level reduced to 41% in 2010 according to a World Bank estimate (Ministry of Finance of Timor-Leste, 2013f). In 2012, still, 37.4% population lived below the poverty line of $1.25 a day.

The majority of those living below the poverty line reside in rural areas with little education and most of them depend on subsistence agriculture for their livelihood. The agricultural sector of the country is marked by low productivity and is vulnerable to climatic change.

With regards to hunger, Timor-Leste’s objective was to achieve food security and sustainability by 2014. The 2009-2010 National Demographic Health Survey of Timor-Leste indicated that 53% of all the children under 5 were stunted and 33% were severely stunted. It also found that 52% of the children under five were underweight (National Statistics Directorate of Timor-Leste, 2010). According to CIA World Fact Book (2013), Timor-Leste has the highest percentage (45.3%) in the world of having underweight children under 5. This indicates a prevalence of chronic malnutrition among children.

In spite of rapid economic growth, Timor-Leste’s labour market is not capable of creating enough jobs to absorb a large number of jobseekers entering the labour market every year. According to UNDP in Timor-Leste (2013), an estimated 15,000 young
people enter the labour market in Timor-Leste every year, while only 400 new jobs are created annually.

More than 40% of youth are currently unemployed. With half of the population below the age of 18 years and continued population growth, it is expected that unemployment will remain a pressing problem in Timor-Leste. Overall, Timor-Leste may find it highly challenging to achieve its MDG targets related to reducing extreme poverty and reducing the number of under-five children who are underweight.

**3.7.2 MDG 2: Universal enrollment in primary education**

Timor-Leste approved a national education law in 2008 to ensure that all children below 17 years of age access free basic education. However, people living in remote villages sometimes find it hard to reach a school.

Timor-Leste’s Standard of Living Survey in 2007 estimated a net enrollment rate of 65.6% in primary education (National Statistics Directorate of Timor-Leste, 2008). According to the Ministry of Education’s estimate, this rate rose to 82.7% in 2009. The 2011 Household Income and Expenditure Survey found a net enrollment rate of 84.7% in primary education with female enrollment rates higher than those of male both at the primary and secondary level (National Statistics Directorate of Timor-Leste, 2011c).

However, according to UNDP in Timor-Leste (2013), the dropout rate among primary school children in Timor-Leste is very high. Almost four in every ten primary school children in Timor-Leste drop out before completing their primary schooling.

**3.7.3 MDG 3: Promote gender equality and empower women**

Timor-Leste’s target for achieving a ratio of 100 per cent literate women to men requires further work. The ratio of literate women to men increased slightly from 96% in 2001 to 97% in 2004 but fell back to 93% in 2007. The trend of higher proportion of girls to boys being enrolled in the primary and secondary education also reverses at the tertiary level where the proportion of female enrollment is 83 to 100 male enrollment.
The male literacy rate in rural areas is higher than that in urban areas. In rural areas, only 62% of women are literate compared to 72% of men. The percentage of female wage earners in the non-agriculture sector was 36% in 2007, indicating a male dominance in this area.

In terms of having female parliamentarians, Timor-Leste ranks 16th in the world. In 2012, the country exceeded its target (35%) and had 38.5% of seats in the National Parliament held by women, which is the highest percentage of women parliamentarians in Asia and the Pacific (UNDP in Timor-Leste, 2013).

3.7.4 MDG 4: Reduce child mortality

Timor-Leste has made significant improvements in reducing infant mortality and under-five children mortality rates and achieved its targets against this goal. According to UNDP in Timor-Leste (2013), infant mortality rate in Timor-Leste decreased to 45 per 1000 live births and the under-five children mortality rate decreased to 64 per 1000 exceeding the MDG target of reducing the infant mortality (below 96 per 1000 live births).

The country has also made significant achievements in increasing coverage for measles and full childhood vaccinations. Measles vaccination coverage for one-year old children increased from 50% in 2001 to 68% in 2009. Full childhood vaccination also increased from 18% in 2003 to 53% in 2009 (UNDP in Timor-Leste, 2013). However, the country still has a long way to go to further reduce child mortality.

3.7.5 MDG 5: Improve maternal health

Timor-Leste saw one of the highest maternal mortality ratios in the world in 2000 when the maternal mortality ratio was estimated to be 660 per 100,000 live births. Although this rate is falling, high maternal mortality still remains a serious concern for Timor-Leste.

The Demographic and Household Survey of 2009-2010 (National Statistics Directorate of Timor-Leste, 2010) indicated 557 maternal deaths per 100,000 live births against the MDG target of reducing it to 252 deaths per 100,000 live births. The survey also showed
that while 86% of pregnant women received some antenatal care from a skilled provider, only 30% of births were attended by a skilled health professional.

Although much improved from the 2002 figures, use of modern contraceptives in Timor-Leste is still limited to only 20% women (National Statistics Directorate of Timor-Leste, 2010).

3.7.6 MDG 6: Combat HIV/AIDS, malaria and TB

HIV prevalence in Timor-Leste is estimated at 0.3%-0.68% of the adult population, 2.76% among female sex workers and 1.33% among men who have sex with men (Ministry of Health of Timor-Leste, 2010). While the prevalence of HIV/AIDS in general population in Timor-Leste is believed to be relatively low, the country has many of the factors that can cause a rise in the prevalence.

Malaria has been a major public health concern in Timor-Leste with malaria incidence rising until 2006 but then the situation improved and malaria incidence fell from 206 per 10,000 in 2007 to 112 per 10,000 in 2010.

In 2005, Timor-Leste had the highest per capita incidence of TB in Asia at 250 smear-positive cases per 100,000. Although Timor-Leste has been consistently achieving more than 85% treatment success rate, TB is still a major health issue in the country. Prevalence of all forms of TB in 2010 was estimated as 643 per 100,000. The 2013 estimation of all forms of TB prevalence was 447 per 100,000 people (WHO, 2013h).

3.7.7 MDG 7: Ensure environmental sustainability

The government of Timor-Leste has ratified the Kyoto Protocol of 1997 (United Nations, 1998), the Montreal Protocol of 1987 (United Nations Environment Programme, 1987), and the Vienna Convention of 1988 (United Nations Environment Programme, 1988) on environmental sustainability. However, it is not adequately documented what steps Timor-Leste has taken to manage its fragile natural resource base and preserve its rich land and marine bio-diversity. With the majority of Timorese people depending on subsistence farming and fishing, natural disasters such as droughts, flooding, erosion
and landslides might easily see severe impact on the livelihoods, growth and development of Timor-Leste.

Timor-Leste has a target to provide access to clean water to 78% of its population by 2015. According to a WHO–UNICEF Joint Monitoring Plan (JMP) study in 2011, 69% of the population in Timor-Leste had access to an improved drinking water source and only 39% of the population had access to an improved sanitation facility (UNICEF, 2013).

Timor-Leste has experienced slight loss of forest cover from 51% in 2001 to 50% in 2009. With 96% of its households depending on wood as source of energy (National Statistics Directorate of Timor-Leste, 2011c), it is unlikely that the country will reach its target of 55% land covered by forest by 2015.

**3.7.8 MDG 8: Develop a global partnership for development**

Timor-Leste has been a recipient of ODA over a decade. The country has made a “commitment to good governance, sustainable development and poverty reduction and focused particularly on establishing the foundations for an open, rule-based, non-discriminatory trading and financial system” (UNDP in Timor-Leste, 2013). More than 40 bilateral and multilateral agencies and hundreds of NGOs take part with the Timor-Leste government in addressing development needs.

In recent years Timor-Leste has strengthened its institutions and systems including Civil Service Commission, Anti-Corruption Commission, National Petroleum Authority, justice sector, taxation reform and public finance management system. The private sector of the country is small but growing, and public-private partnerships are still at a very early stage.

Timor-Leste has witnessed dramatic improvements in information and communication technology with approximately half of the population now enjoying access to a mobile phone and approximately 25% of the population has access to the Internet (UNDP in Timor-Leste, 2013).
3.8 Discussion and conclusion

Timor-Leste has developed a long-term national strategic development plan that aspires to create a diversified and socially inclusive economy, with improved quality of life, health and education for all Timorese by the year 2030 (Government of Timor-Leste, 2011). The contextual analysis of Timor-Leste’s geography, people and culture, history, economy, politics, and human development indicates some strengths, challenges and risk factors for the health sector of the country. In summary, the strengths include strong demographic dividend with a growing young population, relative stability over the last five years, rapid economic growth and assurance of funding from the oil-based economy, assistance from international development partners, and some good progress in human development indicators such as increased life expectancy and reduced infant mortality.

Emerging from the post-conflict fragility Timor-Leste had some commendable successes within a short period. The conflict in the country has been largely under control since 2008. Peaceful elections were held in 2012 and Timor-Leste has achieved the highest percentage (38.5%) of women parliamentarians in the Asia-Pacific region.

The establishment and successful management the Petroleum Fund is another success for Timor-Leste. Timor-Leste ranked 5th in the world in the governance of natural resource fund (the Petroleum Fund) in the 2013 Resource Governance Index (Revenue Watch Institute, 2013).

In the health sector Timor-Leste has also had some successes. Most of the health infrastructures are now being built up. The country has increased life expectancy, reduced child mortality and reduced incidence of infectious diseases from those of 2002 and 2008. Vaccination coverage has also been increased. However, despite rapid economic growth driven by the Petroleum Fund income and a large infrastructure budget, Timor-Leste is still facing a range of challenges and trying to overcome its weaknesses.
The country still needs to diversify its economy and find a non-oil dependent base for economic activities and growth. Jobs and equitable economic opportunities need to be created to cope with the demographic dividend of numbers of young people entering the labour market every year. The private sector and civil society need to be strengthened to complement government’s activities and to create stable and accountable governance.

Timor-Leste’s health and human development outcomes are still very low and many of the MDGs are unlikely to be met by the country by 2015. The maternal mortality is very high. The country ranks number one in the world in terms of the percentage of underweight children under five years of age. More than half of the country’s children under five are stunted indicating chronic mal-nutrition. 31% of the population does not have access to improved drinking water and 61% do not have access to improved sanitation facilities as per standards accepted in the MDGs.

The rapid economic growth, low human development outcomes and challenging achievements of MDGs, in fact, indicate Timor-Leste’s transition from the fragile state to a state of reconstruction and development. The country is overcoming its conflict affected past but is still in the transition to reach to a state of stability with both economic and human development.

With the presentation of the broader country context in this chapter, the next chapter will focus on analyzing the country fragility and health systems contexts to identify the factors that may influence the outcomes achieved by the selected aid funded interventions under evaluation.
4 Aid instruments and results framework: analysis of project backgrounds

4.1 Introduction

As identified by Leader and Colenso (2005), there are multiple aid instruments and modalities employed by donors and recipient countries in fragile states and situations. Such aid instruments and modalities include a wide range of programme aid, projects, Global Funds, technical co-operation, multi donor trust funds (MDTFs), social funds, humanitarian aid, and joint programmes with their own pros and cons (pp. 46-47). These modalities often fall in the continuum of “relief to development” approach (Haider, 2014) and “state-avoidance to state-partnership” approach in their operations (Canavan, Vergeer, & Bornemisza, 2008).

Like in other fragile states and situations, Timor-Leste’s health sector also saw multiple funding mechanisms and aid management modalities for external funding in the health sector. The Development Cooperation Report 2012 of Timor-Leste (Development Partnership Management Unit, 2013) identified 15 different bilateral and multilateral agencies that provided ODA to Timor-Leste’s health sector through different mechanisms.

Although Timor-Leste did not receive a “direct budget support” to cover the government budgetary shortfall to meet the health sector needs, there was a MDTF to support a SWAp in the health sector with funding initially from the AusAID and World Bank and later from the European Community. In line with the SWAp, funds from multiple external sources to MDTF were expected to contribute to a sector-specific umbrella such as National Health Strategic Plan of Timor-Leste (Ministry of Health of Timor-Leste, 2007) and partnership and coordination between the government and development agencies were expected to be strengthened with a clearer government leadership and state-building initiative (World Bank, 2007a).
While external funding under the SWAp were channelled through the government, there had been “state-avoidance” mechanisms too for channeling the external funding through some implementing agencies and stand-alone interventions. For example, USAID funding for children immunization was channelled and implemented through the IPL implemented by JSI.

In addition to these two different approaches of state-partnership and state-avoidance for external funding flow, Timor-Leste’s health sector also had a ‘hybrid’ funding mechanism for a Global Health Partnership such as the Global Fund. The Global Fund is an international health financing mechanism aimed at fighting AIDS, tuberculosis and malaria through partnerships between government, civil society, the private sector and communities living with the diseases. In Timor-Leste the principal recipient of Global Fund grants was the MoH but the funding supported the vertically managed national disease programmes for HIV/AIDS, TB and Malaria with some additional systems to meet the Global Fund requirements.

This chapter provides a description of the background, objectives, and performance targets of the three interventions selected for comparative evaluation in this research. It then attempts to analyze the relevance of the three interventions to the fragile state and health sector context and elicit their underlying programme theories.

4.2 Methods

Using the continuums of ‘relief to development’, ‘service delivery to systems building’, and ‘state avoidance to state partnerships’ Canavan et al. (2008) places different aid modalities of transitional funding in fragile situations in the following matrix:
development approaches have now superseded this transitional aid instrument, but it has provided many lessons on delivery of sustainable interventions and opportunities for development cooperation during protracted conflict.

Another innovative example of transitional aid instruments is the recent launch of the Sudan UNDP Recovery Fund, designed to expedite implementation of early recovery activities in Southern Sudan. This aid mechanism is perceived as a means to accelerate the delivery of essential services given the slow progression of the MDTF and includes a steering committee of government and international representatives with governance arrangements independent of the MDTF. Multi-year funds will be available, with a focus on support systems for delivery of basic services such as monitoring and evaluation and fiscal management of funds.

The World Bank's Post-conflict and Licus Trust Funds (which are being replaced by a new 'State and Peace-building' Fund) are dedicated to capacity building of government administrations based on a systems approach, while moving away from fragmented approaches that characterize responses to crises.

The funds are intended to complement other aid instruments (e.g. MDTF) whereby technical assistance is provided to core government departments for design and management of civil service structures.

Mix and Sequencing of Aid Instruments

Figure 9 illustrates the aid instruments described above, set along the relief to development continuum. All aid mechanisms used in the transition have their strengths and weaknesses (see Annex 6 for further details).

**Figure 4.1: Aid instruments in fragile situations**

Source: (Canavan et al., 2008, p. 31)

The aid instruments and the matrix can be further analyzed through monitoring indicators of the Paris principles of aid effectiveness (Paris Declaration, 2005). These indicators can be adapted in the form of the questions shown in the following table:

**Table 4.1: Monitoring indicators of Paris Declaration for intervention assessment**

<table>
<thead>
<tr>
<th>Indicator number</th>
<th>Assessment question</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Is the intervention based on an operational development strategy owned by the country?</td>
</tr>
<tr>
<td>2a</td>
<td>Does the intervention contribute to reliable PFM (PFM) systems?</td>
</tr>
<tr>
<td>2b</td>
<td>Does the intervention contribute to reliable procurement systems?</td>
</tr>
<tr>
<td>3</td>
<td>Are the aid flows for the intervention recorded in country budgets?</td>
</tr>
<tr>
<td>4</td>
<td>Is technical assistance through the intervention aligned and coordinated?</td>
</tr>
<tr>
<td>5a</td>
<td>Does the intervention use country PFM systems?</td>
</tr>
<tr>
<td>5b</td>
<td>Does the intervention use country procurement systems?</td>
</tr>
<tr>
<td>6</td>
<td>Does the intervention avoid parallel intervention implementation units (PIUs)?</td>
</tr>
<tr>
<td>7</td>
<td>Is the aid for the intervention more predictable?</td>
</tr>
<tr>
<td>8</td>
<td>Is the aid for the intervention untied?</td>
</tr>
<tr>
<td>9</td>
<td>Do the donors of the intervention use coordinated mechanisms for aid delivery?</td>
</tr>
<tr>
<td>10a</td>
<td>Do the donors of the intervention coordinate their missions?</td>
</tr>
<tr>
<td>10b</td>
<td>Do the donors of the intervention coordinate their country studies?</td>
</tr>
<tr>
<td>11</td>
<td>Does the intervention have sound frameworks to monitor results?</td>
</tr>
<tr>
<td>12</td>
<td>Does the intervention have mechanisms for mutual accountability?</td>
</tr>
</tbody>
</table>

Source: Adapted from (OECD, 2011a, p. 19)

First, this chapter presents the background, objectives, planned activities, results framework, and management modality of each project and uses the typology and the continuums of aid modalities (Canavan et al., 2008) to distinguish the different aid management modalities followed by three different projects. Second, each project was
assessed against the indicators in Table 4.1. Third, the underlying programme theory of each intervention is explored for the evaluations presented in subsequent chapters.

The sources of information for this chapter include a number of documents related to the project concepts, proposals, project agreements, budget, work plan, performance frameworks, as well as progress reports and any available reviews as listed in Annex 5 and referenced in this chapter. From review of these documents, underlying programme theory for each project was elicited by identifying their inputs, planned activities, outputs, outcomes and performance measurement indicators.

4.3 Health Sector Strategic Plan- Strengthening Project (HSSP-SP)

4.3.1 Background

The HSSP-SP was undertaken by the MoH of Timor-Leste in 2007 with initial funding from the AusAID and World Bank under a MDTF mechanism. According to a World Bank document, the intervention was designed in the context of persistent poor health status and outcomes; low utilization and quality of health services; inadequate demand for services and household health practices; pressing need for human resources development in the health sector; need for strengthening hospital referral network; need for strengthening planning, monitoring and evaluation; need for improving budget execution and expenditure management; and need for capacity building and innovative approaches (World Bank, 2007a, 2007g).

4.3.2 Goal and objectives

Through a sector-wide approach, HSSP-SP aimed to strengthen the MoH’s Health Sector Strategic Plan (Ministry of Health of Timor-Leste, 2011n) with the overall objective to improve the quality and coverage of preventive and curative health services, particularly for women and children, in order to accelerate Timor-Leste’s progress toward the health related MDGs.

The HSSP-SP had two parts—the first part was direct financial support to the MoH’s activities using the Health Sector Strategic Plan and the Mid Term Expenditure
Framework (MTEF). The second part was designed to address challenges and innovations needed for the health sector to be prepared “for the next generation of issues”.

Aligned with the Health Sector Strategic Plan, the specific objectives for the first part of the intervention were to 1) improve accessibility, demand and quality of health services; 2) strengthen support services, human resource development, and management; and 3) strengthen coordination, planning and monitoring.

The objective of the second part of the intervention was to promote innovation and programme development linked to priorities identified in the Health Sector Strategic Plan by providing resources to pilot and evaluate priority health sector innovations, such as demand-side and service-provider incentives, and public-private partnerships.

4.3.3 Activities

With a total budget of 20.3 million US dollars over 5 years, HSSP-SP had a range of planned activities under the following components:

Component 1: Health service delivery:
This component was intended to provide support to the MoH in the delivery of health services priorities identified in the Health Sector Strategic Plan. These priorities included: (a) implementation of a Basic Service Package for primary health care, including improved outreach services to remote areas; (b) strengthening community nutrition and health services; (c) improving district level planning and management capacity to support service delivery; (d) improving hospital care and the referral system; and (e) assuring quality of care throughout the health system, including improvement of infrastructure. Around 12 million US dollars were budgeted for these activities.

Component 2: Support services, human resource development, and management:
This component was intended to improve the capacity of the MoH in: (a) overall governance and management of the health sector; (b) human resource development and management practices in the health sector; (c) procurement, distribution and
management of essential drugs and supplies; (d) core health sector fiduciary (i.e., public financial management for health) and support functions, including planning and supervision of civil works; logistics and maintenance of infrastructure and equipment; financial management; procurement; and information and communication technologies.

Component 2 was also intended to strengthen technical and managerial capacity of the health workforce through: (a) strengthening the capacity of the Institute of Health Sciences; and (b) support for long-term and short-term training for health staff and managers. Around 4 million US dollars were budgeted for these activities.

**Component 3: Coordination, planning and monitoring:**
This component was planned to strengthen the capacity of the MoH in: (a) coordination and monitoring of donor programmes, including through the establishment of a Department for Partnership Management within the MoH; and (b) policy development and operational research.

Component 3 also intended to support the MoH in: (a) the implementation of improved health sector programme planning and budgeting procedures at national and district levels; and (b) implementation of a practical and cost-effective Health Sector Strategic Plan monitoring and evaluation system, including: (i) strengthening the Health Management Information System; and (ii) support for health sector surveys and evaluations. Around 2 million US dollars were budgeted for these activities.

**Component 4: Innovation and programme development:**
This component was planned to support initiatives, among others, to: (a) promote community demand for health services; (b) provide incentives to service providers; (c) establish effective public-private partnership options; and (d) pilot rapid results initiatives, to build local implementation capacity and strengthen the focus on results. Around 2.3 million US dollars were budgeted for supporting these activities.

**4.3.4 Management modality**
With an initial 19.3 million USD funding from the AusAID and 1 million USD grant from the World Bank/IDA, HSSP-SP used a MDTF mechanism managed by the World Bank. At
the donor level AusAID had administrative agreement with the World Bank, which was responsible for managing the MDTF. However, the biggest part of HSSP-SP was executed by the MoH of Timor-Leste following the World Bank rules.

The MoH of Timor-Leste had responsibility for HSSP-SP implementation, including procurement, disbursement, and financial management. Intervention policy and execution management was the responsibility of the Minister, who served as the Project Director of HSSP-SP. Responsibility for implementation of HSSP-SP financed activities rested with the various line departments of the MoH.

To facilitate integration of external financing with the national budget, the MoH created a new Directorate of Planning and Finance, and established within this Directorate a Department of Partnership Management. In addition, a Project Management Unit previously established within the MoH continued to provide operational support to the implementation, procurement and financial management of HSSP-SP.

4.3.5 Results framework

HSSP-SP planned to monitor the progress toward achievement of the HSSP-SP objectives by using a combination of sector-wide health service indicators such as vaccination coverage, contraceptive prevalence, and coverage of trained birth attendants. There was a set of intermediate indicators linked with activities and outputs supported by the Intervention such as number of children participating in community nutrition/health programmes, percentage of health clinics offering the full Basic Service Package (BSP), percentage of health center and district managers receiving management training, and availability of essential drugs at health facilities and so on. The following was the results framework of HSSP-SP:
Table 4.2: Results framework of HSSP-SP

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Baseline value and year</th>
<th>Target (31 December 2012)</th>
<th>Data Source/Remark</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage of children under 1 year of age vaccinated against (i) DPT3; and (ii) measles</td>
<td>63% 61% 31 Dec 2006</td>
<td>90% 90%</td>
<td>HMIS</td>
</tr>
<tr>
<td>Percentage of births attended by skilled health personnel</td>
<td>27% 31 Dec 2006</td>
<td>45%</td>
<td>HMIS</td>
</tr>
<tr>
<td>Percentage of pregnant women receiving four or more prenatal visits</td>
<td>36% 31 Dec 2006</td>
<td>55%</td>
<td>HMIS</td>
</tr>
<tr>
<td>Percentage of children (6-59 months) receiving vitamin A supplements</td>
<td>36% 31 Dec 2006</td>
<td>80%</td>
<td></td>
</tr>
<tr>
<td>Contraceptive Prevalence Rate</td>
<td>10% 31 Dec 2006</td>
<td>25%</td>
<td>HMIS</td>
</tr>
<tr>
<td>Number of formal MOH-private sector/NGO contracts signed</td>
<td>0 31 Dec 2006</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Availability of tracer essential drugs at (i) SAMES; (ii) health facilities</td>
<td>88 30 June 2008</td>
<td>98 95</td>
<td>HMIS and Central Warehouse Management System</td>
</tr>
</tbody>
</table>

Source: Adapted from (World Bank, 2011a) and (Lee, 2013)

The indicators specified in the Results Framework were planned to be collected through: (i) routine monitoring through the health management and information system (HMIS) and the MoH disease surveillance systems; (ii) household and beneficiary surveys; (iii) routine supervision and/or facility surveys; (iv) community-based monitoring; and (v) evaluation of pilot activities (World Bank, 2007g).

4.4 The Global Fund funded National HIV/AIDS and STI Programme (NAP)

4.4.1 Background

The Global Fund adopted a model that provides performance based funding to country governments and in-country stakeholders based solely on proposals and implementation plans designed by the countries themselves (The Global Fund, 2013).

In the context of low HIV prevalence but high risk for a concentrated HIV/AIDS epidemic, Timor-Leste successfully submitted a Round Five proposal to the Global Fund for funding the NAP of Timor-Leste for the period from June 2007 to December 2011 following the 2nd National HIV/AIDS and STI Strategic Plan: 2006-2010 (Ministry of Health of Timor-Leste, 2006).
In 2010, HIV prevalence in Timor-Leste was estimated at 0.30% to 0.68% among adult general population, 2.76% among female sex workers, and 1.33% among men who have sex with men (Ministry of Health of Timor-Leste, 2011j). However, as in previous studies (Pissany, 2004; UNSW, 2008), injection drug use behaviour and HIV prevalence among injection drug users in Timor-Leste were largely unknown. Timor-Leste developed the 3rd National HIV/AIDS and STI Strategic Plan for 2011-2016 (Ministry of Health of Timor-Leste, 2011a) and again successfully submitted proposal to the Global Fund for Round Ten funding for the NAP. The Round Ten funded NAP started implementation from January 2012 for a five-year period.

The total signed grant for the Round Five NAP was US$ 8,224,997 for the period from June 2007 to December 2011, whereas the total signed grant for the Phase 1 of Round Ten NAP was US$ 6,085,519 for the period from January 2012 to December 2013.

4.4.2 Goals and objectives

The primary goal of the NAP was to reduce HIV/AIDS and STI mortality and morbidity in Timor-Leste by enhancing related prevention and treatment services. The NAP sought to address critical gaps in supporting and scaling up the national response to HIV/AIDS and STI by activities around five objectives: 1) improve and increase coverage of strategic HIV/AIDS and STI prevention services; 2) strengthen and expand HIV/AIDS treatment and care services; 3) strengthen health systems related to HIV/AIDS and STI; 4) strengthen community systems related to HIV/AIDS and STI; and 5) build an enabling environment for implementation of the NAP.

4.4.3 Activities

Activities to achieve the objectives stated above were organized around each objective under some thematic categories termed as service delivery areas (SDA).

**Activities for HIV/AIDS and STI prevention (Objective 1):**
Activities under this objective included behaviour change communication (BCC) and community outreach with provisions for BCC to most at risk groups such as sex workers, men who have sex with men (MSM) and injection drug users (IDU); and provisions for BCC and community outreach to people with multiple sex partners including clients of
sex workers, vulnerable women and prisoners. This objective was further supported by
distribution and social marketing of condoms, and offering a number of HIV/AIDS
prevention services: HIV/AIDS counseling and testing, prevention of mother to child
transmission, management of sexually transmitted infections (STIs), safe blood
transfusion, and provisions for universal precaution and post-exposure prophylaxis.

**Activities for HIV/AIDS treatment and care (Objective 2):**
Activities to achieve this objective were organized around strengthening and expanding
service provisions antiretroviral treatment (ART), and prophylaxis and treatment for
opportunistic infections for people living with HIV/AIDS. This objective was supported
by enhancing care and support for people living with HIV/AIDS.

**Activities for strengthening health systems relating to STIs and HIV (Objective 3):**
Activities around this objective included strengthening systems related to the
management of NAP; monitoring, learning and knowledge management; HIV/AIDS and
STI related operations research and surveillance; and blood bank and laboratory
services and quality.

**Activities for strengthening community systems relating to STIs and HIV
(Objective 4):**
Activities to achieve this objective were organized around advocacy, communication and
social mobilization; building community linkages, collaboration and coordination;
human resources and skills development; and creating community based service
delivery options.

**Activities for building an enabling environment for implementation of STI and HIV
programme (Objective 5):**
Activities to achieve this objective were organized around social mobilization for
addressing stigma and discrimination; promoting gender equality; conducting policy
advocacy with legislation; institution building for HIV/AIDS governance and
coordination; and promoting a multi-sectoral response to HIV/AIDS.
### 4.4.4 Management modality

Following the Global Fund’s funding mechanism, Timor-Leste set up a Country Coordinating Mechanism (CCM) with representations from the Government, Development Partners, NGOs and people living with HIV/AIDS. The CCM coordinates development of funding proposals, submits proposal to the Global Fund, and oversees implementation of the Global Fund grants by the Principal Recipient.

The Principal Recipient of the Global Fund’s HIV/AIDS grant was the MoH of Timor-Leste. The NAP of the MoH managed and implemented the Global Fund grant funded activities through the MoH’s mechanisms as well as through the Sub Recipient NGOs. A specialized project facilitation unit named as the Division of Global Fund was set up by the MoH to support financial management and grant implementation of the Global Fund grants.

The Principal Recipient reports to the CCM as well as to the Global Fund through periodic progress update and disbursement requests. A Local Fund Agent (Pricewaterhouse Coopers, Jakarta) appointed by the Global Fund independently verifies the results, implementation status, and risks of the Principal Recipient and Sub Recipients and reports directly to the Global Fund.

Although implemented through MoH mechanisms, the NAP put in place a vertically direct reporting mechanism to collect the data from the health facilities and NGOs to be able to timely report to the Global Fund on the agreed performance indicators.

### 4.4.5 Results framework

The results framework of the NAP included the following impact, outcome and process indicators:
Table 4.3: Results framework for Timor-Leste’s NAP

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Baseline %</th>
<th>Baseline year</th>
<th>Target (Dec 2013)</th>
<th>Data source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage of young women and men aged 15-24 who are HIV infected</td>
<td>0.68%</td>
<td>2010</td>
<td>&lt;1.00%</td>
<td>National survey</td>
</tr>
<tr>
<td>Percentage of sex workers who are HIV infected</td>
<td>2.76%</td>
<td>2010</td>
<td>&lt;5%</td>
<td>IBBS</td>
</tr>
<tr>
<td>Percentage of MSM who are HIV infected</td>
<td>1.3%</td>
<td>2011</td>
<td>&lt;3%</td>
<td>IBBS</td>
</tr>
<tr>
<td>Percentage of adult and children known to be on treatment 12 months after initiation of antiretroviral therapy</td>
<td>62.5%</td>
<td>2012</td>
<td>80%</td>
<td>Patient record</td>
</tr>
<tr>
<td>Percentage of female sex workers reporting the use of a condom during penetrative sex with their most recent client</td>
<td>63.2%</td>
<td>2011</td>
<td>75%</td>
<td>IBBS</td>
</tr>
<tr>
<td>Percentage of men reporting the use of a condom the last time they had anal sex with a male partner</td>
<td>31.2%</td>
<td>2011</td>
<td>55%</td>
<td>IBBS</td>
</tr>
<tr>
<td>Percentage of uniformed personnel reporting the use of condom last time they had sex with a non-regular partner</td>
<td>21.8%</td>
<td>2010</td>
<td>28%</td>
<td>IBBS</td>
</tr>
<tr>
<td>Percentage of men who have sex with men (MSM) surveyed who both correctly identify ways of preventing the sexual transmission of HIV and who reject major misconceptions about HIV transmission</td>
<td>48.7%</td>
<td>2010</td>
<td>50%</td>
<td>IBBS</td>
</tr>
<tr>
<td>Percentage of female sex workers surveyed who both correctly identify ways of preventing the sexual transmission of HIV and who reject major misconceptions about HIV transmission</td>
<td>56.5%</td>
<td>2010</td>
<td>66%</td>
<td>IBBS</td>
</tr>
<tr>
<td>Number of men who have sex with men (MSM) reached through a basic package of HIV/AIDS prevention services</td>
<td>1083</td>
<td>Jun-2013</td>
<td>1050</td>
<td>NGO reports</td>
</tr>
<tr>
<td>Number of Female Sex Workers (FSW) reached through a basic package of HIV/AIDS prevention services</td>
<td>641</td>
<td>Jun-2013</td>
<td>800</td>
<td>NGO reports</td>
</tr>
<tr>
<td>Number of people who receive HIV testing and counseling including provision of test results</td>
<td>1480</td>
<td>2011</td>
<td>16500</td>
<td>Service centre reports</td>
</tr>
<tr>
<td>Number and percentage of HIV-positive pregnant women who received anti-retroviral to reduce the risk of mother-to-child transmission during the last 12 months</td>
<td>3/3</td>
<td>100%</td>
<td>2011</td>
<td>70/70 100%</td>
</tr>
<tr>
<td>Number and percentage of health facilities providing STI case management services as per national guidelines</td>
<td>43/71</td>
<td>61%</td>
<td>2012</td>
<td>57/71 80%</td>
</tr>
<tr>
<td>Number of adults and children with advanced HIV infection currently receiving antiretroviral treatment</td>
<td>110</td>
<td>2011</td>
<td>298</td>
<td>Service centre reports</td>
</tr>
<tr>
<td>Number and percentage of PLHIVs (who know their HIV status) reached with a basic package of care and support services</td>
<td>231/332</td>
<td>70%</td>
<td>2012</td>
<td>387/526 74%</td>
</tr>
<tr>
<td>Number and percentage of blood bank and laboratory facilities conducting HIV screening/ testing as per national guidelines</td>
<td>13/13</td>
<td>100%</td>
<td>2012</td>
<td>13/13 100%</td>
</tr>
<tr>
<td>Number of civil society/community based organizations received capacity development support</td>
<td>5</td>
<td>2012</td>
<td>10</td>
<td>Programme reports</td>
</tr>
</tbody>
</table>

Source: Adapted from (The Global Fund, 2014)

Baseline values of the impact indicators in the above table seem to have been already achieved before the implementation of the NAP. However, as HIV is, still, a non-curable life-long infection, and it is desirable to have no AIDS related deaths for the people living
with HIV/AIDS, for an early stage of the HIV/AIDS epidemic it is not realistic to set HIV/AIDS prevalence targets lower than the baseline. As the prevalence of HIV/AIDS would increase with new infections and reduced AIDS related deaths, NAP’s goal was to contain and reduce the rate of HIV transmission. It was assumed that without NAP intervention, the prevalence of HIV/AIDS among most at risk populations in Timor-Leste would be much higher than the targets set in the above table. While a few indicators depended on regular surveillance and the HMIS of the MoH, the NAP had a separate vertical system for collecting necessary data to be able to meet strict monitoring and reporting requirements of the Global Fund.

4.5 Immunizasaun Proteje Labarik (IPL)

4.5.1 Background

Known as Imunizasaun Proteje Labarik, IPL was implemented by JSI from April 2011 to October 2013 with US$ 2,639,250 funding through a USAID Threshold Grant from the MCC.

Timor-Leste’s Demographic and Health Survey in 2009/10 found 66.7% immunization coverage for DPT3 and 68.2% for measles (National Statistics Directorate of Timor-Leste, 2010). The IPL intervention was undertaken in this context of the lowest reported immunization coverage in the South-East Asia region of WHO. IPL worked at the national, district and local levels and specifically targeted 7 districts out of total 13 districts of Timor-Leste. These seven districts were Ainaro, Baucau, Dili, Ermera, Liquica, Manufahi and Viqueque. While the number of under 1 year old children in Timor-Leste was 38,915 in 2012, these seven focus districts had 26,357 under one year old children with lower immunization coverage rates than the other districts.

4.5.2 Goals and objectives

IPL worked in partnership with the MoH Timor-Leste with an aim to increase child immunization coverage for DPT3 and measles nationally. IPL’s overall goal was to assist the MoH of Timor-Leste to increase DPT3 and measles immunization coverage to a national average of 81.5% for children below one year of age by 2013. However, a complementary goal of the intervention was to strengthen the Government’s Expanded
Programme on Immunization (EPI) so that the country is able to sustain and expand the gains achieved from IPL intervention.

IPL was designed to collaborate with the MOH to achieve the following objectives of 1) strengthening service delivery to identify and reach unimmunized children at least five times a year; 2) strengthening district and Community Health Center (CHC) level programme management capacity and technical skills among government health personnel; 3) strengthening SISCa (a community based satellite service delivery mechanism) as an effectively functioning community-based outreach mechanism for providing immunization and other health services; and 4) strengthening programme monitoring and reporting through better collection of routine data and the routine analysis and use of data for decision-making and targeted action.

4.5.3 Activities

IPL focused its activities for strengthening various aspects of the immunization systems including community participation, local level engagement, communication, outreach, health staff capacity, vaccine transportation cold-chain management and logistics, micro planning, and district and national level partnerships.

IPL’s major activities included community leaders training and encouraging community participation. IPL provided orientation to junior high school students and supported satellite outreach services (SISCa), mobile clinics, and outreach work. The project introduced micro planning for immunization at the sub-district and village level, provided health worker training, and conducted supervision visits at different health facilities.

In order to fill a gap for skilled health personnel, IPL provided funding to hire 10 midwives from Indonesia to work in under-staffed health centres for a period of 17 months. Other activities of IPL included efforts for improving reporting and recording systems for immunization, improving coordination and collaboration between district and sub-district level health services, and conducting advocacy and facilitating immunization health partnerships at the national and strategic level (MCHIP, 2013g).
### 4.5.4 Management modality

IPL followed a ‘project funding’ model with services contracted out to an international NGO. The project was funded by MCC, with local oversight by USAID/Dili office. The management of IPL was under the umbrella of the USAID-funded Maternal and Child Health Integrated Programme (MCHIP), the USAID Bureau for Global Health's flagship maternal, neonatal and child health (MNCH) programme.

Under these arrangements, IPL was managed and administered by JSI with a team of international and local staff hired in Timor-Leste for implementing this project. The project team worked closely with the MoH of Timor-Leste and built on the MoH’s existing mechanisms and capacity for service delivery.

### 4.5.5 Results framework

Following were the coverage, outcome and output indicators adopted by IPL:

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Baseline</th>
<th>Baseline year</th>
<th>Target</th>
<th>Data source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average of DPT3 and measles immunization coverage rates</td>
<td>64.5%</td>
<td>2011</td>
<td>81.5%</td>
<td>HMIS</td>
</tr>
<tr>
<td>DPT1 vaccination rate</td>
<td>68%</td>
<td>Oct-Dec 2011</td>
<td>No target set</td>
<td>HMIS</td>
</tr>
<tr>
<td>DTP3 vaccination rate</td>
<td>64.8%</td>
<td>Oct-Dec 2011</td>
<td>No target set</td>
<td>HMIS</td>
</tr>
<tr>
<td>Drop-out rate between DPT1 and DPT3</td>
<td>5%</td>
<td>Oct-Dec 2011</td>
<td>No target set</td>
<td>HMIS</td>
</tr>
<tr>
<td>Measles vaccination rate</td>
<td>66.4%</td>
<td>Oct-Dec 2011</td>
<td>No target set</td>
<td>HMIS</td>
</tr>
<tr>
<td>Percentage of CHCs with current micro-plans that include catchment area map, fixed site and outreach strategies, based on coverage data, to reach all under ones in the catchment area.</td>
<td>20%</td>
<td>2011</td>
<td>No target set</td>
<td>IPL Baseline Assessment</td>
</tr>
<tr>
<td>Improved vaccinator ranking on quality measures in 2 CHCs (six monthly; done in the same quarter)</td>
<td>16%</td>
<td>Oct-Dec 2011</td>
<td>No target set</td>
<td>Surveillance</td>
</tr>
<tr>
<td>Number of staff in the focus districts who receive further/ refresher training in Immunization in Practice</td>
<td>28</td>
<td>Oct-Dec 2011</td>
<td>No target set</td>
<td>Intervention reports</td>
</tr>
<tr>
<td>Number of staff in the focus districts who receive further/ refresher training in Cold Chain and Vaccine Management</td>
<td>27</td>
<td>Oct-Dec 2011</td>
<td>No target set</td>
<td>Intervention reports</td>
</tr>
<tr>
<td>Number of staff in the focus districts who receive further/ refresher training in cold chain repair and maintenance</td>
<td>0</td>
<td>Oct-Dec 2011</td>
<td>No target set</td>
<td>Intervention reports</td>
</tr>
<tr>
<td>Number of nurses and midwives who receive in-service training</td>
<td>30</td>
<td>Oct-Dec 2011</td>
<td>No target set</td>
<td>Intervention reports</td>
</tr>
<tr>
<td>Percentage of health facilities (with functioning refrigerator) with proper vaccine management</td>
<td>23%</td>
<td>Oct-Dec 2011</td>
<td>No target set</td>
<td>Administrative data</td>
</tr>
<tr>
<td>Percentage of health facilities (with functioning)</td>
<td>30%</td>
<td>Oct-Dec</td>
<td>No target set</td>
<td>CHC reports/</td>
</tr>
</tbody>
</table>
refrigerator) reporting a vaccine stock-out in the past 3 months |  | 2011 | set | Intervention reports, Surveillance reports/HMIS
---|---|---|---|---
15 | Number of schools and churches received orientation on EPI | 15 | Oct-Dec 2011 | No target set | Intervention reports
16 | Percentage of CHCs with an updated list of unreach children by Suco | 0% | Oct-Dec 2011 | No target set | Intervention reports
17 | Proportion of poor performing sucos in priority sub-districts whose health committee of the Suco council or Suco council including its PSF has been oriented on EPI | 18% | Oct-Dec 2011 | No target set | Administrative data
18 | Proportion of CHC holding review meeting quarterly with participation of sub-district administrator and Chefe Suco, PSF where Microplan, SISCa and EPI performances were reviewed | 11% | Oct-Dec 2011 | No target set | Meeting minutes/ Intervention report/ participants list
19 | Proportion of CHCs with an active system for identifying and following up unreach and dropped out children | 0% | Oct-Dec 2011 | No target set | Intervention report
20 | Proportion of Suco council meetings of the identified poor performing sucos held to address getting more Suco children immunized | 2% | Oct-Dec 2011 | No target set | Intervention report

Source: Adapted from Project Monitoring Plan (Immunizasaun Proteje Labarik, 2012)

It can be seen that IPL only set target for the high-level coverage indicator and did not have any specific targets set for the process indicators. Although IPL kept regular track of their progress against the baseline of all process indicators, their performance framework allowed flexibility of implementation by not setting targets against the lower level activities. IPL used the existing MoH systems and HMIS for some indicators, it also has its vertical systems for collecting monitoring data on its project level activities.

### 4.6 Analysis and discussion

A number of works acknowledge that the Paris principles of aid effectiveness are very much applicable to the fragile states and situations (Cox & Hemon, 2009; DFID, 2010; Leader & Colenso, 2005; NZAID, 2008; OECD, 2007d, 2011c; OPM/IDL, 2008; Vergeer et al., 2009). Using the monitoring indicators of the Paris Declaration, as adapted in Table 4.1, the design and operational arrangements of the three selected projects are analyzed in Table 4.5:
Table 4.5: Timor-Leste's selected health interventions in light of monitoring indicators of the Paris Declaration

<table>
<thead>
<tr>
<th>Assessment question</th>
<th>HSSP-SP</th>
<th>NAP</th>
<th>IPL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Is the intervention based on an operational development strategy owned by the country?</td>
<td>Yes (based on a sector strategic plan)</td>
<td>Yes (based on a national programme plan)</td>
<td>No (based on donor assessment)</td>
</tr>
<tr>
<td>2a Does the intervention contribute to reliable PFM (PFM) systems?</td>
<td>Yes (supports PFM)</td>
<td>No (bypasses PFM)</td>
<td>No (bypasses PFM)</td>
</tr>
<tr>
<td>2b Does the intervention contribute to reliable procurement systems?</td>
<td>Yes (supports govt. procurement system)</td>
<td>Yes (uses govt. procurement system)</td>
<td>No (bypasses govt. procurement system)</td>
</tr>
<tr>
<td>3 Are the aid flows for the intervention recorded in country budgets?</td>
<td>Yes (but not integrated)</td>
<td>No (uses separate accounts)</td>
<td>No (uses separate mechanism)</td>
</tr>
<tr>
<td>4 Is technical assistance through the intervention aligned and coordinated?</td>
<td>Yes (through sector coordination)</td>
<td>No (stand alone for the project)</td>
<td>No (stand alone for the project)</td>
</tr>
<tr>
<td>5a Does the intervention use country PFM systems?</td>
<td>Yes (uses govt systems)</td>
<td>No (set up a parallel system with Ministry)</td>
<td>No (uses different system)</td>
</tr>
<tr>
<td>5b Does the intervention use country procurement systems?</td>
<td>Yes (subject to WB policy compliance)</td>
<td>Yes (subject to GF policy compliance)</td>
<td>No (uses project’s mechanism following USAID rules)</td>
</tr>
<tr>
<td>6 Does the intervention avoid parallel intervention implementation units (PIUs)?</td>
<td>No (has a project management unit)</td>
<td>No (has a Division of Global Fund)</td>
<td>No (has a separate project team)</td>
</tr>
<tr>
<td>7 Is the aid for the intervention more predictable?</td>
<td>Yes (subject to currency exchange rate fluctuations)</td>
<td>Yes (subject to performance based funding)</td>
<td>Yes (as per agreement)</td>
</tr>
<tr>
<td>8 Is the aid for the intervention untied?</td>
<td>Yes (untied grant)</td>
<td>Yes (untied grant)</td>
<td>No (USAID procurement source rules apply)</td>
</tr>
<tr>
<td>9 Do the donors of the intervention use coordinated mechanisms for aid delivery?</td>
<td>Yes (through sector coordination)</td>
<td>Yes (through country coordinating mechanism)</td>
<td>No (direct funding to project)</td>
</tr>
<tr>
<td>10a Do the donors of the intervention coordinate their missions?</td>
<td>Yes (organized joint missions)</td>
<td>No (did not take part in any joint mission)</td>
<td>No (did not take part in any joint mission)</td>
</tr>
<tr>
<td>10b Do the donors of the intervention coordinate their country studies?</td>
<td>Yes (coordination between WB, AusAID, EC &amp; MoH)</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>11 Does the intervention have sound frameworks to monitor results?</td>
<td>No (results framework not clearly linked with project activities)</td>
<td>Yes (has performance framework clearly linked with funding)</td>
<td>Yes (has PMP linked with funded activities)</td>
</tr>
<tr>
<td>12 Does the intervention have mechanisms for mutual accountability?</td>
<td>Yes (MoH held accountable for implementation)</td>
<td>Yes (Principle Recipient held accountable for results)</td>
<td>No (accountability of JSI towards MoH not clear)</td>
</tr>
</tbody>
</table>

Source: Author’s analysis

An analysis of project background documents show that all three projects considered in designing their interventions Timor-Leste’s post-conflict country context of the fragile situation, internal violence, political instability, poverty, poor human development indicators, and weaknesses in institutional capacity. All three interventions tried to address the critical gaps in the country needs and responses in achieving one or more
MDGs related to reducing child mortality (MDG 4), improving maternal health (MDG 5) and combating HIV, TB and Malaria (MDG 6). However, while HSSP-SP was based on Timor-Leste’s national Health Sector Strategic Plan, and the NAP was based on a national HIV/AIDS strategic plan, IPL was “designed to improve Timor-Leste’s performance in the policy areas measured by MCC’s Control of Corruption and Immunization Rates indicators” (MCC, 2010, p. 1).

In terms of the design process, the three interventions differ significantly from each other. For example, HSSP-SP was designed by a “Joint Design Mission” with participation from the MoH of Timor-Leste, World Bank, AusAID and World Bank hired external consultants. The Global Fund proposal, on the other hand, was led by a country process and developed by in-country stakeholders with no involvement of the Global Fund in the proposal development process other than reviewing the proposal by an independent Technical Review Panel. IPL was predominantly designed by USAID based on a review of Timor-Leste’s eligibility for accessing MCC funding with the implementing NGO selected through an open international tender.

In terms of aligning to existing country systems, HSSP-SP tried to contribute to strengthening public financial management system (PFM) and procurement system by using the existing systems of the MoH subject to compliance with minimum requirements of the World Bank policies and procedures. The NAP, on the other hand, used the MoH’s procurement system with additional support provided by the Division of Global Fund and subject to compliance with the Global Fund’s procurement policies and standards. The NAP, however, employed a separate financial management system within the Division of Global Fund. Being contracted out to an international NGO, IPL did not use the PFM or procurement systems.

Although reporting to the Aid Effectiveness Directorate of the Ministry of Finance of Timor-Leste, the budgets and funding of none of these projects were integrated with the government budget. While HSSP-SP used the existing MoH’s bookkeeping systems for project expenditure accounting, the NAP used separate books and accounts for project expenditure. IPL’s expenditure was also managed and accounted for separately outside the MoH’s systems.
For the management modalities, HSSP-SP went through the regular structure of the MoH with crosscutting functions and regular activities of different units of the MoH being supported by the project funding through a ‘horizontal programming’ approach. The NAP also went through the MoH, but implemented a vertically designed stand-alone programme with a set of agreed time bound activities and specific process monitoring indicators. IPL, on the other hand, was implemented by an international NGO outside the MoH’s management mechanism and in parallel to the immunization programme run by the MoH.

In line with the horizontal and vertical programming approaches, the technical assistance provided by HSSP-SP was aligned to the MoH’s needs and priorities, whereas technical assistance provided by the NAP and IPL mostly followed a stand-alone approach.

It can be seen that all three interventions used specialized “Project Implementation Units” (PIU). Although ‘horizontally designed’, HSSP-SP had a Project Management Unit (PMU) under the Department of Partnership Management of the MoH staffed with a Project Manager, Technical Advisers, Finance and Procurement staff. The NAP was also supported by a Division of Global Fund staffed with a Programme Manager, Programme Management Adviser and specialized finance, procurement and administrative staff. The Division of Global Fund also structurally fell under the Department of Partnership Management of the MoH. As mentioned earlier the IPL was also implemented by a Project Team hired by JSI with a Country Director and specialized staff.

In terms of predictability of aid, all three projects had agreed budget and total amount of funding expressed through the project agreements. However, total funding available for the HSSP-SP fluctuated due to the currency exchange rate fluctuation for Australian dollar (AUD) and US dollar (USD). Initially the total funding available for HSSP-SP from the 23 million AUD AusAID funding was estimated as 21 million USD. However, in order to manage the exchange rate risks, the initial allocation was set at 15.7 million USD in the Grant Agreement. World Bank later increased this allocation from 15.7 million USD to 20 million USD for HSSP-SP (World Bank, 2013k).
Similarly, a maximum total amount of funding for the NAP was indicated in the Global Fund Grant Agreement in USD. However, actual release of funding was subject to the satisfactory performance of the NAP. IPL did not have any such currency exchange rate fluctuation or performance based funding restriction issue.

Although there were some conditions that needed to be met by the projects before funding could be released, funding for both HSSP-SP and the NAP was untied to any procurement source and origin related conditions. IPL had some procurement source and origin related conditions as generally imposed by USAID on all USAID funding.

On the questions of harmonization, HSSP-SP followed a coordinated mechanism for aid delivery, organizing joint assessment missions, and joint country studies. For the NAP, funding proposal submission to the Global Fund was coordinated through the CCM. However, the Global Fund never took part in any joint missions or joint country study in Timor-Leste with other donors and the government. Funding for IPL was also not delivered through a coordinated mechanism and IPL also did not take part in any joint donor mission or joint country study.

On the question of managing for results, three interventions employed three different mechanisms. The HSSP-SP employed a flexible approach and allowed the MoH to develop a project work plan but MoH had to obtain a “No objection letter” from the World Bank in cases of difference from the approved work plan. For HSSP-SP the World Bank quarterly produced project status reports and internally evaluated the project performance in terms of their process and outcomes based on the agreed results framework.

The Global Fund, on the other hand, used a results based funding mechanism with rigorous monitoring of financial and programmatic performance. The Global Fund engaged an external agency called Local Fund Agent (LFA) to conduct periodic verification of monitoring reports of the Global Fund grant funded activities. Any failure by the Principal Recipient to account with financial agreement and/or to meet performance targets was likely to affect continuation of funding by the Global Fund.
Although IPL worked closely with the MoH in implementing the project activities and collecting monitoring data, IPL was monitored by JSI and local USAID’s monitoring procedures against the agreed results indicators.

On the question of mutual accountability, donors held the implementing agencies accountable for implementation of agreed donor funded activities. While accountability of the project implementation teams of HSSP-SP and the NAP to the MoH were clearly established through the reporting lines, IPL’s project team remained outside the reporting line of the MoH.

Use of different aid management modalities in fragile situations is not new. The health systems of fragile states often have to go through the phases of relief, rehabilitation, reconstruction and development (Newbrander et al., 2003). Due to the lack of legitimacy, accountability, capacity and transparency of the state mechanisms and the need for essential services, peace building, and state building, donors often employ different aid instruments and aid management modalities in fragile situations.

The three selected aid interventions in Timor-Leste’s health sector, in fact, represent three major external fund management modalities in a fragile situation. Management Sciences for Health (2007) identified some advantages and disadvantages of different donor financing options for health services in fragile states, which are also applicable for Timor-Leste’s situation. These advantages and disadvantages of different aid modalities are summarized in the following table:
Table 4.6: Advantages and disadvantages of donor financing options for health services in fragile states

<table>
<thead>
<tr>
<th>Donor financing option</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>General budget support</td>
<td>Donor alignment with govt’ priorities</td>
<td>Potential for money being pulled away from health services to other govt’ services</td>
</tr>
<tr>
<td></td>
<td>Supports donor-govt’ accountability</td>
<td>Not targeted to those in greatest need</td>
</tr>
<tr>
<td></td>
<td>Aligns support with govt’ priorities</td>
<td>Dilutes donor attribution</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Impact on improving health is weakened</td>
</tr>
<tr>
<td>SWAps: Pooled donor funding</td>
<td>Donor alignment with govt’ priorities</td>
<td>Lack of govt’ capacity to coordinate</td>
</tr>
<tr>
<td></td>
<td>Harmonization among donors</td>
<td>Lost opportunities for broader impact on govt’ financial, logistics, and service delivery systems</td>
</tr>
<tr>
<td></td>
<td>Enhances donor-govt’ accountability</td>
<td>Dilutes donor attribution</td>
</tr>
<tr>
<td></td>
<td>Efficiencies—reduces transaction costs</td>
<td>Difficulty in getting all major donors to participate</td>
</tr>
<tr>
<td></td>
<td>Aligns budgeting with priorities</td>
<td></td>
</tr>
<tr>
<td>Contracting with NGOs for service delivery</td>
<td>Services extended quickly</td>
<td>Potential for donors to bypass govt’in contracting with NGOs since contracting</td>
</tr>
<tr>
<td></td>
<td>Promotes govt’ role of steward, overseer, and regulator of health sector</td>
<td>requires govt’ capacities and systems to adequately manage contractors</td>
</tr>
<tr>
<td></td>
<td>Promotes a basic package of health services for delivery throughout the country</td>
<td>Dependent upon NGOs being able to scale up their service provision capacity quickly</td>
</tr>
<tr>
<td></td>
<td>May be more cost efficient than govt’ provision of health services</td>
<td>Cost and sustainability questions arise when main contractors are international NGOs or</td>
</tr>
<tr>
<td></td>
<td>Leverage for monitoring NGOs’ performance in extending access and providing quality care</td>
<td>local private providers</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NGOs’ contractual relationship with donor or govt’ may compromise their perceived “honest broker” role</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Issue of sustainability for the long term</td>
</tr>
<tr>
<td>Global health partnerships</td>
<td>Widens coverage, especially of the poor, for provision of certain disease-specific services or prevention</td>
<td>System, management, and sustainability issues similar to contracting with NGOs [limited capacity for absorbing resources]</td>
</tr>
<tr>
<td></td>
<td>Can fill gaps of service provision</td>
<td>Requires strong leadership and management capacity</td>
</tr>
<tr>
<td></td>
<td>Addresses imbalances in equity and access</td>
<td>Potential parallel or duplicative mechanisms</td>
</tr>
<tr>
<td></td>
<td>Standardized approaches help promote faster implementation</td>
<td>Disease-specific interventions may create nonintegrated services</td>
</tr>
<tr>
<td></td>
<td></td>
<td>May not support capacity building throughout the health system</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Requires a coordination mechanism within MoH</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Global mechanisms may not be flexible enough for fragile states</td>
</tr>
</tbody>
</table>

Source: [Management Sciences for Health, 2007, p. 29]

Timor-Leste’s health sector did not see any ‘general budget support’ mechanism. However, the HSSP-SP was a SWAp with pooled donor funding through a MDTF mechanism. The IPL was the mechanism for contracting with NGOs for service delivery and the NAP was a global health partnerships mechanism.

As the related literature showed, a SWAp provides opportunities for increased harmonization and alignment but comes up with the risk of slow implementation and difficult institutional compatibility of different large systems (Barakat, 2009; Chansa et al., 2008; Hill, 2002; Hutton, 2004; Hutton & Tanner, 2004; Martinez, 2008; McNee,
In fragile situations, on the other hand, donors and the recipient country often have to rely on vertical funding mechanisms such as the Global Health Initiative and contracting out to NGOs (Baudienille et al., 2010; Leader & Colenso, 2005; Newbrander et al., 2003; White & Cliffe, 2000). Lagarde and Palmer (2009) found that contracting out services to non-state providers may be appropriate in post-conflict or fragile situations as it may increase access to and utilization of health services. However, analyzing experiences in Mozambique, Mussa et al. (2013) found that there is value and importance in vertical funding to the health sector to achieve programmatic successes although there are three serious concerns around the vertical funding approach:

“1) difficulties coordinating external resources and challenges to local control over the use of resources channelled to international private organizations; 2) inequalities created within the health system produced by vertical funds channelled to specific services while other sectors remain under-resourced; and 3) the exodus of health workers from the public sector health system provoked by large disparities in salaries and work” (p.1).

Waddington et al. (2009) also criticized the well funded global health partnerships such as the Global Fund and GAVI saying that:

“They [global health partnerships] produce results, but are limited by the coverage and quality of health systems. Their achievements will be expensive to sustain. The Global Fund in particular has created many parallel systems, which have created inefficiencies in implementation” (p.4).

With advantages and disadvantages of different aid management modalities in fragile situations a common view among both donors and recipient governments is that “there is virtue in using a diversity of modalities as a means of spreading risk” in fragile situations (Booth & Fritz, 2008). The first principle for Good International Engagement in Fragile States and Situation (OECD, 2007d) also recommends for a mix of different aid instruments according the need of the particular fragile state context saying that, in fragile situations, “International actors should mix and sequence their aid instruments according to context, and avoid blue-print approaches” (p.1).

4.7 Conclusion

In line with these views and by applying the typology used by Canavan et al. (2008) we can see that HSSP-SP, the NAP, and IPL employed three different aid management modalities. HSSP-SP employed a development aid, systems building, and state partnership approach facilitating the MoH’s leadership in a SWAp. IPL, on the other hand, employed a relief aid, service delivery, and seemingly state avoidance approach by
contracting out the project outside the government. The NAP, by using the MoH as the Principal Recipient but by relying on a vertical management system of the Global Fund grants and operation, falls in-between these different continuums of relief to development, service delivery to systems building, and state avoidance to state partnerships. The management modality of the NAP allowed the Global Fund to channel the fund through the government but without fully following the government’s existing mechanisms and systems.

Putting the labels on the intervention by using these typologies, however, can lead to oversimplification of the facts. Service delivery and systems building may not always fall at the two opposing ends of a continuum—rather they can be complementary to each other. In fact, in Timor-Leste’s case, all three selected projects had service delivery as well as systems building components. Eldon et al. (2008) found that strengthening service delivery and health systems in fragile situations contribute to state building in the health sector.

Although the three selected interventions focused on achieving improvement of population health outcomes, all of them had health systems strengthening components. For example, the original and revised plans of HSSP-SP focused on strengthening health support services, human resources, and PFM and pharmaceutical management systems. The NAP focused on strengthening health systems and community systems related to HIV/AIDS and STI. IPL also focused on strengthening service delivery systems and management capacity related to immunization at the district, community health centre and village level. It is, therefore, useful to see how these different aid modalities interact with the state fragility and health systems to produce intended and unintended outcomes.
5 Theoretical framework and research methods

5.1 Introduction

As discussed in Chapter 2, despite a sizable body of literature on aid effectiveness and remarkable recent progress in the field of impact evaluation methodologies, aid effectiveness evaluation is still challenged with the lack of a common theoretical framework and published work on aid effectiveness evaluation in a fragile state’s health sector context. The reasons for such a disturbing situation include great variability in thinking and understanding of the concept of aid effectiveness, and non-consideration of systems interactions and health systems outcomes in relation to longer lasting effects of aid in the health sector.

Bringing together the diversified streams of thinking, intervention logic model, and health systems framework, this research developed a methodology for a “realist evaluation”* of aid effectiveness in the health sector. This chapter presents the broad theoretical framework and an overview of methods used in this research. Since a broad range of methods were used, for easier reading each subsequent chapter of this thesis provides a more detailed description and explanation of the methods used for that particular chapter.

This chapter is organized into three sections. In the first section, building on the changing concept of aid effectiveness from econometric studies to the Paris Declaration and further to the Busan Outcome and impact evaluation, an attempt was made to settle on a comprehensive view of aid effectiveness incorporating the Paris principles as well as programmatic outcomes of an aid intervention. The second section draws on programme theory and health systems thinking and specifies the theoretical framework

* Pawson and Tilley first used the term “realistic evaluation” (Pawson & Tilley, 1997) to propose their evaluation methods. However, they later switched to using the term “realist evaluation” (Pawson & Tilley, 2004) due to preferred nomenclature of other authors.
used in this research. The third section then outlines the specific steps and methods used for this research including measures for quality control, validity, and robustness.

5.2 Aid effectiveness and development effectiveness

As discussed in the previous chapter, followed by the Paris Declaration (2005), the spotlight on aid effectiveness shifted from the economic growth effect of aid to the process of improving aid disbursement and management. Prior to and followed by the Fourth High Level Forum on Aid Effectiveness in Busan (2011), the focus of aid effectiveness shifted further to the developmental outcomes or development effectiveness of an aid intervention. In fact, the evaluation of the Paris Declaration (Wood et al., 2011) tried to measure the effect of adherence to the Paris principles at the country levels on the development outcomes in the health sector. This leaves the study of aid effectiveness with the question if there is a causal relationship between adherence to the Paris principles and programmatic outcomes of an aid intervention in the health sector.

These paradigm shifts denote that there are several diversified concepts inherent to the quality of aid effectiveness: the quality of aid policy, instruments and programme design; the quality of aid implementation process and operations; and the quality of aid intervention outcomes. Since a combination of all these aspects is important in maximizing impact of aid on development, they can be conceptually put together in an analytical-assimilative way to derive a more comprehensive concept of aid effectiveness in the following manner:

I. Quality of aid delivery policy, instruments and programme design + Quality of implementation process and operations = Paris principles of aid effectiveness
   (Aid effectiveness)

II. Quality of implementation process and operations + Quality of intervention outcomes = Programmatic effectiveness

III. Paris principles of aid effectiveness + Programmatic effectiveness = Development effectiveness of an aid intervention

With the principles of country ownership, alignment, harmonization, managing for results, and mutual accountability, the Paris principles of aid effectiveness (Paris
Declaration, 2005), in fact, focused on the aid delivery policy, instruments, intervention design as well as the implementation process and operations of the aid intervention. It was expected that adhering to these principles would produce better programmatic outcomes or programmatic effectiveness.

The Busan Outcome (4th High Level Forum on Aid Effectiveness, 2011), on the other hand, tried to link more explicitly the Paris principles of aid effectiveness to programmatic effectiveness under the concept of development effectiveness. As Lisk et al. (2013) pointed out:

“...by focusing on the concept of development effectiveness or, more precisely, how the aid effectiveness principles can improve development outcomes, the Busan agreement represents a mechanism to transform aid relationship into spurs of development” (p.125).

These concepts of aid effectiveness and programmatic effectiveness underlying the concept of development effectiveness can be presented in Figure 5.1:

![Figure 5.1: A comprehensive view of aid effectiveness and development effectiveness](source: Author’s diagram)

Following this growing emphasis on the effect of aid effectiveness principles on development effectiveness, this research was also particularly interested in linking the adherence of an aid intervention to the Paris principles to the outcomes of that intervention in a fragile state’s health sector context and seeing if such adherence is correlated with the effectiveness of the intervention in terms of achieving its objectives.
5.3 Theoretical framework for aid effectiveness evaluation

In their definition of evaluation, the Organization for Economic Co-operation and Development (OECD) emphasized the design, implementation and results of an aid intervention:

“Evaluation is the systematic and objective assessment of an on-going or completed project, programme or policy, its design, implementation and results. The aim is to determine the relevance and fulfilment of objectives, development efficiency, effectiveness, impact and sustainability” (OECD-DAC, 2002) (unpaginated web based publication).

While an evaluation aims to systematically investigate the relevance and achievement of objectives along with the efficiency, effectiveness, impact, and sustainability of an intervention, evaluation also depends on a framework of predefined standards or values against which the worth of an intervention outcome is judged. According to Suchman (1967), evaluation is a process of determining the value or worth of something by judging it against explicit, predetermined standards.

Evaluators, however, often vary widely on the issues of standards and approaches to evaluation. Like the paradigm shifts in the concept of aid effectiveness, there have been at least two opposing philosophical underpinnings of evaluation theory and approaches, which is referred as positivists-constructivists debate or ‘paradigm war’ (Armytage, 2011).

The positivist view to evaluation advocates for a traditional scientific approach by trying to establish the validity and reliability of data, adopting experimental methods with counterfactual measurements. The constructivist approach, on the other hand, tries to interpret the reality from the ‘voice of stakeholders’ by adopting participatory methods, case-studies and observations and by often refuting the scientific approach as being costly, impractical and irrelevant (Cracknell, 2000).

A third approach that tries to assess not only the implementation of an intervention and its effectiveness, but also the causal mechanisms and the contextual factors that underlie outcomes of an intervention, is broadly referred to as theory-driven evaluation (Chen, 1990). A theory-driven evaluation maps out the causal chain or ‘programme theory’
from inputs to outcomes and impact of an intervention and tests the underlying assumptions behind such theory to answer the question why and how it works.

As explained by Greenhalgh et al. (2009), a programme theory or results-chain logically shows the assumptions and causal linkages between the resources (input), activities (process), immediate result (outputs), medium term results (outcomes), and direct and indirect longer-term results (impact) of an intervention. Thus the programme theory helps the evaluator to deconstruct the design of an intervention into its different components and then reconstruct it with the causal chain and expose the underlying mechanisms and influence of the context in which the intervention takes place.

Using a results-chain model the Operations Evaluation Department of the World Bank suggested a “Corporate Scorecard Model” (Picciotto, 2002) to evaluate the aggregated performance of an externally funded national development programme:

![Figure 5.2: Corporate scorecard model using results-chain](image)


The Corporate Scorecard Model provides an example of an impact evaluation framework for national level large-scale interventions by identifying the causal chain from inputs such as budget resources and skills; to outputs such as financial and knowledge products; to reach such as partnership processes; to outcomes such as country programme results; and finally to impacts such as contribution to MDGs. The scorecard model also suggests for analyzing attribution of results to exogenous factors,
global policies, partner performance, and agency performance for an effectiveness evaluation of an agency or intervention.

In this tradition of theory-driven evaluation practice, Pawson and Tilley (1997) developed the realistic evaluation, which was later more widely known as ‘realist evaluation’ (Pawson & Tilley, 2004). Realist evaluation argues that rather than answering the question of whether an intervention works, evaluations need to provide more detailed information for decision makers and indicate ‘what works, how, in which conditions and for whom’ (Pawson & Tilley, 2004). To realist evaluators, outcome of an intervention not only depends on its ‘mechanism’ or the underlying programme theory, but it also depends on the ‘context’ of the intervention (Marchal et al., 2012).

These positivist, constructivist, and realist approaches to evaluations can be understood by adapting Scriven’s terminology of ‘black box’, ‘grey box’ and ‘white box’ evaluations (Scriven, 1994). Kazi (2000) compares the positivist approach to evaluation to the ‘black box’ evaluation, in which the researcher concentrates on scientifically evaluating the effect of an intervention without addressing the contents and context of that intervention. The constructivist or ‘interpretivist research’, on the other hand, is comparable to the ‘Grey box’ evaluation where the components of an intervention are discerned, “but their inner workings or principles of operation are not fully revealed” (Kazi, 2000).

It can be seen that a realist evaluation tries to go beyond the ‘black box’ and ‘grey box’ to evaluate the mechanism, context and outcome of an intervention. As Kazi (2000) puts it:

“Scientific realism [such as realist evaluation] attempts Scriven’s ‘white box’ evaluation, which not only addresses the effects, but also the inner workings and operations of the components of a programme and how they are connected. Such a perspective has a great deal of promise for utilizing both quantitative and qualitative approaches” (p.765).

In a fragile situation, interventions are often complex and tend to be influenced by their contexts. A growing number of researchers argue that a realist evaluation provides a sound framework to examine how context and mechanisms influence the outcomes of an intervention making it suitable for evaluating such complex interventions (Blaise & Kegels, 2004; Byng et al., 2008; Clark et al., 2005; Greenhalgh et al., 2009; Maluka et al.,
Applying the realist evaluation approach to the context of a health sector in a fragile situation, we can argue that the direct and indirect results of health interventions are influenced by their interactions with the health systems as they are placed within the state fragility. According to the WHO (2007):

"A health system consists of all organizations, people and actions whose primary intent is to promote, restore or maintain health. This includes efforts to influence determinants of health as well as more direct health-improving activities. A health system is therefore more than the pyramid of publicly owned facilities that deliver personal health services" (p.2).

Since any health intervention is placed in the broader context of a health system, there has been growing emphasis on considering the effect of health systems on a project as well as effect of the project on the health systems to bring into the analysis of contributions and impacts of a health intervention (Adams & Savigny, 2012; Berman & Bitran, 2011; Brown et al., 2013; Coker et al., 2008; Dodd et al., 2007; Eldon et al., 2008; Galichet et al., 2010; Huntington, Banzon, & Dy Recidoro, 2012; Kruk et al., 2010; Leischow et al., 2008; Olmen et al., 2010; Savigny & Adams, 2009; Trochim et al., 2006; van Olmen et al., 2012). WHO (2007) analyzes the process, organization and goals of a health system in terms of six “system building blocks” as shown in Figure 5.3:

**Figure 5.3: WHO health systems framework**

Source: (WHO, 2007, p. 3)
Analysing health system building blocks, systems thinking calls for a paradigm shift from conventional ‘static’ thinking of “systems as effect” to a ‘dynamic’ thinking of “systems as cause” for the intervention results (Savigny & Adams, 2009). Applying health systems thinking, it can be argued that any intervention in the health sector has interactions with the different building blocks of the health system to produce a series of results chains and net impact. Proponents of systems thinking, therefore, argue that it is essential to analyze the impact pathway of a health intervention’s interactions with the health system to consider both the intended and unintended effects of the intervention in the health sector (Dodd et al., 2007; Olmen et al., 2010; Savigny & Adams, 2009).

While health systems analysis provides necessary information for context and mechanism analysis, a realist evaluation tries to elicit a programme theory to see how the intervention interacts with the health systems. According to the realist evaluation approach,

“Programmes are theories incarnate. They begin in the heads of policy architects, pass into the hands of practitioners and, sometimes, into the hearts and minds of programme subjects” (Pawson & Tilley, 2004, p. 3).

In addition to having an underlying programme theory, the realist evaluation also believes that programmes are ‘embedded’ in their social systems and context having a set of activities that interact with the context; and intervention outcomes may be subject to unanticipated events (Pawson & Tilley, 2004). That means an aid-funded intervention has its underlying programme theory or theory of change and has interactions with the health system’s context to achieve programmatic objectives. However, achievement of programmatic outcomes may be subject to unanticipated events and depend on how the programme beneficiaries (programme subjects) use the outputs.

As it can be seen from the foregoing discussion, by drawing attention to the aid disbursement mechanism and aid management modalities, the Paris principles of aid effectiveness influence the ‘mechanism’ of an intervention while health systems provide the particular context of a health intervention. Therefore, both the Paris principles of aid effectiveness and health systems thinking fit into a theory-driven realist evaluation framework for evaluating the outcomes of adherence to the Paris principles as well as outcomes of a health intervention in a particular fragile state context.
Combining the concept of the Paris principles of aid effectiveness with health systems thinking into a theory-driven realist evaluation approach, we can draw the following evaluation framework:

<table>
<thead>
<tr>
<th>Strategic level</th>
<th>Framework</th>
<th>Indicators</th>
<th>Evaluation Focus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Design and Inputs</td>
<td>Paris principles of aid effectiveness</td>
<td>Ownership</td>
<td>Project contexts</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Harmonization</td>
<td>Adherence to Paris</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Alignment</td>
<td>principles</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Managing for results</td>
<td>Relevance</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mutual</td>
<td>Economy of inputs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>accountability</td>
<td></td>
</tr>
<tr>
<td>Process and Outputs</td>
<td>Results framework; Health systems interactions</td>
<td>Project outputs against performance indicators</td>
<td>Efficiency against targets</td>
</tr>
<tr>
<td>Outcomes and Impacts</td>
<td>Logical framework Impact pathway</td>
<td>Intervention outcome and impacts Health system outcomes</td>
<td>Effectiveness of intervention; Likely Sustainability of results</td>
</tr>
</tbody>
</table>

**Figure 5.4: Theoretical framework for aid effectiveness evaluation**

Source: Author's diagram based on programme logic model

As it can be seen in the above framework, it is assumed that greater adherence to the Paris principles of aid effectiveness (i.e., country ownership, alignment, harmonization, managing for results, and mutual accountability) will work at the intervention design and implementation process level contributing to greater sense of ownership, greater alignment and harmonization of similar efforts, greater policy coordination and increased emphasis on results and accountability. This will, in turn, contribute to increased economy of scale and reduced transaction costs, increased programmatic efficiency, and increased efforts in health systems strengthening. The outputs from the greater adherence to the Paris principles and intervention results will then contribute to increased effectiveness of achieving intervention objectives and increased health system strengthening outcomes and ultimately to the sustainable impact of the intervention. Thus the Paris principles of aid effectiveness, together with the programmatic effectiveness, will contribute to the development effectiveness in terms of achieving programmatic objectives and health systems strengthening.

In summary, the justifications for drawing this evaluation framework in the tradition of theory-driven realist evaluation are:
• Such a theory-driven realist evaluation framework is suitable for evaluating a set of complex programmes in a fragile state's health sector by evaluating their design, process, and outcomes within their own contexts;
• The framework provides a generic conceptual model, which can be used for evaluating diverse interventions without the need for developing a separate evaluation framework for each;
• The framework incorporates the Paris principles of aid effectiveness and their interplay with the design, mechanism and context of each intervention and, thus, links the concept of aid effectiveness with the concept of development effectiveness;
• The framework allows flexibility to choose from a range of qualitative and quantitative research methods suitable for evaluating different parts of the evaluation.

5.4 Research design and methods

The following sections describe the particular design and methods followed by this evaluative research by using the theory-driven realist evaluation framework specified above.

5.4.1 Aim

The aim of this research was to pilot test an evaluation approach to critically understand if and how the Paris principles of aid effectiveness followed by aid interventions contribute to the development effectiveness of that intervention specifically in a fragile state’s health sector context.

5.4.2 Objectives

The specific objectives of the research were to pilot test the feasibility of using the theoretical framework and analytical tools selected for this evaluation to:
• Analyze intervention contexts of the three projects to identify possible implications for the intervention process and outcomes;
• Analyze project designs and evaluate to what extent the three different aid management modalities adhered to the Paris principles of aid effectiveness;
• Elicit underlying programme theories of the three aid interventions and assess their execution and outcomes to evaluate programmatic efficiency and effectiveness;
• Investigate if, why and how adherence to the Paris principles are correlated to the development effectiveness of an intervention.

5.4.3 Setting

The setting of this research was Timor-Leste’s health sector with a period of interest from June 2007 to December 2013 and the selection of three specific aid project initiatives.

5.4.4 Interventions to evaluate

As the research objectives indicate, this research required a number of components to be able to do a comparative evaluation and analysis of relevant aspects of each of three projects selected and determine if there is a correlation between adherence to the Paris principles and development effectiveness. Balancing the need for sizable sample projects for evaluation, and availability and access to data the following three projects were selected from Timor-Leste’s health sector for comparative evaluation:
• The AusAID and World Bank funded HSSP-SP
• The Global Fund funded NAP, and
• USAID funded IPL

As discussed in Chapter 4, these three projects provided three different aid management modalities operated in the same context of Timor-Leste’s health sector. Funded by the AusAID and World Bank through a MDTF mechanism, HSSP-SP aimed to support the MoH of Timor-Leste in implementing Timor-Leste’s health sector strategic plan to achieve health related MDGs. This project employed a sector-wide approach and undertook sector coordination and health system strengthening activities to contribute to the state-building goals.

Unlike the HSSP-SP, the NAP was a disease specific, vertical national programme aiming at reducing disease burden related to HIV/AIDS and STI in Timor-Leste. The principle recipient of the Global Fund grants for HIV/AIDS was the MoH of Timor-Leste. However,
this project depended on a separate project management facilitation unit and did not follow the regular MoH’s systems for financial management and monitoring.

IPL, on the other hand, was implemented outside the government’s mechanism through a contract by the USAID with an international NGO (JSI). However, this project aimed to increase the capacity of the MoH of Timor-Leste to achieve increased national coverage of immunization for DPT3 and measles for children under their first birthday.

A realist evaluation assumes that:

$$\text{Context (C) + Mechanism (M) = Outcome (O)}$$ (Pawson & Tilley, 2004)

Following the principles of realist evaluation, the underlying assumption of this evaluative research sample was: in a common given context of Timor-Leste’s health sector, the three selected interventions following varying degrees of adherence to the Paris principles of aid effectiveness would trigger three different mechanisms (i.e. aid management modalities and underlying programme theories) and they would produce different outcomes. Therefore, the mechanisms and outcomes would be fairly comparable to each other when exploring what worked, why they worked, under which conditions they worked, and if adherence to the Paris principles had any correlation with the outcomes of an intervention.

### 5.4.5 Research questions

Using the concepts of aid effectiveness and development effectiveness specified above the evaluation approach piloted in this research attempted to evaluate and compare the three selected projects from Timor-Leste’s health sector to answer the following key questions:

- Whether and how the contexts of the three projects influenced their intervention processes and outcomes;
- To what extent the design of the three projects and their aid management modalities adhered to the Paris principles of aid effectiveness;
- What are the underlying programme theories of the three projects and how these theories were executed to produce the efficiency of outputs and effectiveness of outcomes;
• If, why and how adherence to the Paris principles are correlated to the efficiency and effectiveness of the three projects.

5.4.6 Research methods

As Pawson and Tilley (2004) suggest, a ‘realist research cycle’ involves the stages of eliciting and formulating the programme theory underlying an intervention; collecting data on appropriate mechanism, context, and outcome; analyzing outcome patterns to test the programme theory systematically; and revising the understanding of context-mechanism-outcome configurations as a prelude to further rounds of theory refinement and testing. Following these suggestions this research employed the steps, approaches, methods and tools summarized in Table 5.1:

Table 5.1: Summary of research design and methods

<table>
<thead>
<tr>
<th>Step</th>
<th>Method/approach</th>
<th>Source of information</th>
<th>Analysis and expected outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1: Context analysis: Analyze the country context, state fragility, and health systems contexts of the interventions under evaluation</td>
<td>Qualitative approach through realist synthesis methods</td>
<td>Related literature, documents from the Timor-Leste government, WHO, UNDP, AusAID, and USAID</td>
<td>Identify factors, risks, and contextual interactions that can influence intervention outcomes (Chapter 3 and 6)</td>
</tr>
<tr>
<td>Step 2: Project design analysis: Analyze project designs and elicit underlying programme theories and intervention parameters</td>
<td>Qualitative-interpretivist approach through impact pathway analysis</td>
<td>Project documents: Project proposals, Project agreements, Work plan, performance framework, progress reports</td>
<td>Elicit underlying programme theories of each project and evaluate relevance of such theories to the context (Chapter 4)</td>
</tr>
<tr>
<td>Step 3: Mechanism analysis: Analyze perspectives of the stakeholders about design, process, results, efficiency, effectiveness, and likely sustainability of each project under evaluation</td>
<td>Constructivist-mixed methods approach through semi-structured interviews of stakeholders; focus group discussions; qualitative and quantitative analysis of responses using balanced scorecards</td>
<td>Semi-structured interviews with 21 key informants (stakeholders) using a ranking questionnaire with 1-5 likert scale; 3 focus group discussions with the project implementation teams</td>
<td>Comparative analysis using balanced scorecards methods on the aspects of adherence to the Paris principles, economy, efficiency, effectiveness, and likely sustainability of each project (Chapter 7)</td>
</tr>
<tr>
<td>Step 4: Outcome analysis: Analyze project outcomes and their values in terms of resources used</td>
<td>Quantitative analysis of cost-effectiveness of outcomes using mathematical modeling exercise with actual, counterfactual and optimum scenario modeling and economic measurement methods</td>
<td>Project reports, account statements, and information available from programme reviews and survey results</td>
<td>Analysis of cost effectiveness and technical efficiency of each project (Chapter 8)</td>
</tr>
</tbody>
</table>
| Step 5: Comparative outcome patterns analysis: Compare outcome patterns of each project for economy, efficiency, effectiveness, and any underlying correlations | Quantitative analysis using DEA technique and statistical analysis of correlations | Project results against performance indicators, results from stakeholder interviews in Step 3, and results from cost effectiveness analysis | Triangulate findings by comparing the programme theory, empirical evidences from qualitative analysis, and empirical evidence from quantitative analysis; compare context-
in Step 4
mechanism-outcome configurations of the projects; investigate correlations between the Paris principles and programmatic effectiveness (Chapter 9)

Step 6: Synthesis, assessment and interpretation of analysis
Realist synthesis approach with critical analysis of research findings
Research findings from Step 1 to Step 5.
Comments and discussion on appropriateness of an aid management modality and correlation between aid effectiveness and development effectiveness in a fragile state’s health sector context (Chapter 10 & 11)

Source: Author

The following sections describe specific methods and tools used in each step:

**Step 1: Context analysis:**
This research used a range of qualitative techniques to understand the country context, state fragility, and health systems context. The commonly used analytical tools included mapping actors, force-field analysis, SWOT (strengths, weaknesses, opportunities and threats) analysis, and PEST (political, economic, social and technological environment) analysis. Synthesis of these analyses was achieved through a ‘realist synthesis’ approach, as summarized by Pawson and Tilley (2004):

“Realist synthesis is an approach to reviewing research evidence on complex social interventions, which provides an explanatory analysis of how and why they work (or don’t work) in particular contexts or settings. It complements more established approaches to systematic review, which have been developed and used mainly for simpler interventions like clinical treatments or therapies” (p. iv).

Due to the complex fragile situation, national level interventions with multiple objectives, and subtle interplay of aid effectiveness principles and health systems with the intervention mechanisms, the realist synthesis approach was thought to be appropriate for filtering and synthesizing available information and findings from multiple sources.

**Country context:**
Based on the related literature review and qualitative analysis of information available on the Government of Timor-Leste’s website and the websites of the World Bank, AusAID, USAID, WHO, and UNDP, the country context analysis provided the key
considerations of the country's geography, history, culture, economy, politics and human development status in terms of achieving the MDGs.

**State fragility:**

Analysis of state fragility of Timor-Leste also followed the realist synthesis method based on Timor-Leste’s self assessment of state fragility (Fragility Assessment Team at Ministry of Finance, 2013), the State Fragility Index by Fund for Peace (2014), and by employing a set of internal and external factors of state fragility as identified by Mcloughlin (2010) based on an extensive literature review focused on fragile states.

Timor-Leste’s self assessment of state fragility (Fragility Assessment Team at Ministry of Finance, 2013) provided the country’s progress and current status against five peace and state-building goals (PSG) such as legitimate politics, security, justice, economic foundations, and revenues and services. The Fragile States Index of the Fund for Peace (2014), on the other hand, is an independent assessment that considers a number of social, economic, political and military indicators to measure a state's fragility or degree of failure every year.


While Timor-Leste’s self-assessment of state fragility and Fragile State Indexes provided good indications of overall status of the country, the underlying factors of fragility were analyzed by assessing the factors identified in prevalent literature on fragility (Mcloughlin, 2010). Analysing Timor-Leste’s situation and key significant events, the chapter on fragility assessment attempted to identify the relevance, presence, and possible interplay of some of these factors in the outcomes of the selected interventions.
**Health systems:**

The overview of Timor-Leste's health systems also followed the qualitative analytical technique based on the health systems framework proposed by WHO (2007) and the systems thinking espoused by Adams and Savigny (2012).

The structure and organization of Timor-Leste's health sector were analyzed in terms of six building blocks of health governance and leadership by using available published work and documents from the MoH and the Ministry of Finance of Timor-Leste, WHO, World Bank, AusAID and USAID. The six building blocks were: 1) health governance and leadership, 2) health service delivery, 3) health information, 4) health work force, 5) health commodities and technology, and 6) health financing. The structure and organization of the health systems were then compared with the state fragility context and health system's performance in achieving MDGs to identify the strengths, weaknesses and risks for the health sector to note any possible effect of the context on the performance of the projects under evaluation.

**Critical analysis of Step 1 techniques:**

Step 1 attempted a synthesis and analysis of available information for the country context, state fragility, and health systems. While this approach is useful in building background scenario necessary for understanding the aid interventions, it largely depends on the availability and correctness of information. Luckily, in Timor-Leste's case, information was available from the government and donor websites and in a number of reports. However, this may not be the case for some fragile situations where information is not readily available.

The realist synthesis method followed in this step also depended on the researcher and maybe subjective in some cases. Nevertheless, use of some analytical techniques such as SWOT analysis, state fragility analysis, and health systems analysis provided useful considerations for evaluating aid interventions' outcomes.

**Step 2: Project design analysis:**

As seen in Chapter 4, this research conducted an analysis of the intervention designs in question to identify their underlying programme theories, assess the programmatic
mechanisms they planned to apply, review the identified needs they planned to address, and relevance of their objectives, results framework and impact pathways. This analysis considered a number of documents related to the project proposals, project agreements, budget, work plan, performance frameworks, as well as progress reports and any available reviews. Annex 5 provides a list of these project documents. From review of these documents, underlying programme theory for each project was elicited by identifying their inputs, planned activities, outputs, outcomes and performance measurement indicators.

**Critical analysis of Step 2 techniques:**
Step 2 employed a qualitative-interpretivist approach through impact pathway analysis. This technique was useful in understanding the underlying programme logic or theory of change behind each intervention. Again, the quality of this analysis depends on the availability and articulation of the intervention documents. For example, some projects may have well articulated theory of change and project design, while others may not have well-expressed programme logic. Therefore, inferring and analyzing the programme logic may depend on the researcher and may differ from what the project designers actually had in mind.

**Step 3: Mechanism analysis with balanced scorecards:**

**Approach:**
In order to understand the ‘mechanisms’ triggered by the interventions under evaluation, this research used a constructivist approach to involve the stakeholders and construct an understanding based on stakeholder perspectives Guba and Lincoln (1989) and described by Stufflebeam (2008): this research collected, analyzed and evaluated the perspectives of the stakeholders about design, process, results, efficiency, effectiveness, and likely sustainability of each intervention under evaluation.

**Sampling of stakeholders:**
The sampling and selection of stakeholders to take part in the interviews were purposive based on the relevance and participants’ knowledge about the interventions. A total of 22 stakeholders including donor representatives, policy makers, technical experts, project implementation team members, sector specialists, people with oversight
function, and representatives from different sub-systems of the health sector were selected for the interviews based on their informed consent and availability.

Table 5.2 shows a list of different categories of participants that took part in the interviews.

**Table 5.2: Sampling of stakeholder interviews**

<table>
<thead>
<tr>
<th>Category</th>
<th>Number of participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Sector coordination representative from Country Coordinating Mechanism</td>
<td>1</td>
</tr>
<tr>
<td>2 Health sector specialist from Academia</td>
<td>1</td>
</tr>
<tr>
<td>3 MoH senior management and policy makers</td>
<td>4</td>
</tr>
<tr>
<td>4 Members of project implementation teams</td>
<td>7</td>
</tr>
<tr>
<td>5 Counterpart representatives from relevant sub-systems of the health system</td>
<td>4</td>
</tr>
<tr>
<td>6 Technical advisers and donor representatives</td>
<td>4</td>
</tr>
<tr>
<td>7 NGO</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>22</strong></td>
</tr>
</tbody>
</table>

During the 40 to 60 minute interview process with each participant, a ranking questionnaire was filled in by the researcher based on the participant response and opinion and the interview was electronically recorded by consent. Participants were also given the option to remain anonymous in the research documents. Based on this option, all participants opted to remain anonymous in the research documents.

**Questionnaire:**

In order to collect the stakeholder views a ranking questionnaire was developed to compare and rank different aspects of each intervention against the assumed best practices as outlined in the detailed criteria of ranking in Annex 1.

The questionnaire contained 22 different questions with each question having a rank of 1 to 5 to indicate the status of the intervention on that particular aspect; 5 represented the best practice and most desirable status whereas 1 represented the worst case. Each rank for each question was explained to the interview participants by circulating the questionnaire in advance. A semi-structured interview was then conducted with each participant where the questions and ranks were explained again by the researcher and they were asked to rank the status of each intervention against those explained ranks indicating the number of steps away from the best practice by an intervention. In
addition to the ranking, additional open-ended probing questions were asked to confirm the participants' views and clarify any points made.

**Importance weighting:**
Since the importance of different aspects of the interventions would vary to the policy makers in terms of their relevance and value judgment from the managerial point of view, this research also tried to capture the importance weight of each item of the questionnaire. Three senior level policy makers and managers from the MoH of Timor-Leste were interviewed. They were asked to assess independently the importance of each item of the questionnaire on a scale of 1 to 10 with 10 representing the highest and 1 the lowest level of importance from the MoH’s point of view.

**Focus group discussions:**
In addition to the stakeholder interviews, three focus group discussions were conducted with the project implementation teams of the selected interventions. Each focus group discussion was participated by 3 members from the project implementation team representing an intervention. The participants of the focus group discussion were asked open-ended questions about the design, programme strategy, implementation, outcomes, drivers, barriers, lessons learned, strengths and weaknesses of their project.

Information obtained from the focus group discussions was used to mark any factual error provided by any participant during the stakeholder interview process.

**Data analysis:**
Scores obtained from the participants’ responses on the ranking questionnaire and importance weight interviews were compared with the recorded interviews and cross-checked with the project documents and focus group discussions to clarify any confusion or factual error. The scores were then transferred to a spreadsheet for descriptive statistical analysis including standard deviation, weighted scores, total scores, and average scores on each item and group of items representing an aspect of the interventions.
As can be seen in Annex 1, the ranking questionnaire used for collecting stakeholder views employed a Likert scale with scores ranging from 1 to 5 for each item of the questionnaire. These scales require equal interval properties (i.e., equal distance between the ranks) as a basic requirement for descriptive statistical analysis such as summing up, calculating mathematical averages and determining standard deviations. However, for the purpose of analyzing stakeholders’ opinions, this thesis did not determine the scaling properties of the questionnaire. As discussed in Chapter 7, even if some of the indicators might not have equal interval scales, for the sake of an indicative analysis, it was assumed that these scales did have equal interval properties.

**Presentation and interpretation of scores:**

As will be shown in Chapter 7, this research used the ‘Balanced scorecard’ method to present and interpret the findings from the stakeholder interviews, focus group discussion, and importance weight interviews to compare the degree of adherence to the Paris principles, economy, efficiency, effectiveness, and likely sustainability of each intervention. Results obtained from the stakeholder interviews were then compared with and validated by findings from other available sources such as project documents listed in Annex 5 and other published reviews related to the findings.*

**Critical analysis of Step 3 techniques:**

Step 3 employed a constructivist and mixed methods approach through semi-structured interviews of stakeholders, focus group discussions, and qualitative and quantitative analysis of responses using balanced scorecards. While this approach is useful in analyzing and presenting multiple aspects of a project, the challenges for such approach come from the source of information. As the indicators in balanced scorecards are often scored subjectively by management consultants [https://balancedscorecard.org/Resources/About-the-Balanced-Scorecard accessed 31 January 2016](https://balancedscorecard.org/Resources/About-the-Balanced-Scorecard), the validity of the empirical evidence presented through this approach

* In cases of discrepancies and suspected error in scoring in one individual case, the author went back to the interviewee to double-check and correct the score. For the summary and overall scores, the results and discussion sections of this thesis presents findings from multiple sources for comparison and triangulation.
can be challenged. As discussed above, statistical analysis of scores obtained through the likert scale presents additional challenge related to the assumption of equal distance between the ranks or status of a project in terms of the selected indicators for their analysis. Nevertheless, using stakeholders in scoring the indicators provides a good approach of perception analysis that can be further compared and analyzed with other available information and further analysis through cost effectiveness analysis and data envelopment analysis.

**Step 4: Outcome evaluation with cost-effectiveness analysis:**

*Approach:*
In addition to collecting and analyzing stakeholder views, this research employed cost-effectiveness analysis to analyze project outcomes and their values in terms of resources used and number of DALYs saved by each intervention.

*Methods:*
The cost effectiveness analysis in this research attempted to calculate the cost per DALY saved by each selected intervention. This was done by measuring the outcomes of each intervention and comparing the outcomes with two other scenarios: 1) the null or control scenario without the intervention; and 2) the optimum scenario if all the intervention targets were achieved and all the allocated resources were fully utilized by the intervention.

This research used the overall DALY for Timor-Leste as calculated by Global Health Estimates (WHO, 2014) as a basis for estimating DALY saved by HSSP-SP. Therefore, being consistent with the DALY measures used by WHO in Global Health Estimates (WHO, 2014), this research used the simplified methods of calculating DALYs without the time discounting and with uniform age weights compared to the highest level of life expectancies (92 years) projected for the year 2050 by the World Population Prospects 2012 (United Nations Department of Economic and Social Affairs Population Division, 2013).
**Data sources and calculations:**
The cost effectiveness analysis was done based on a synthesis of all available information including project documents such as project appraisals, budgets, performance targets, expenditure accounts, results reports, other available evaluation, national statistics, as well as any survey or review findings.

As explained in Chapter 8 in more detail, three different methods were used for calculating total DALYs saved or that would be saved by each project in different scenarios. DALYs saved by HSSP-SP were estimated from the estimates of total disease burden for Timor-Leste based on Global Burden of Disease (Institute for Health Metrics and Evaluation, 2013) and Global Health Estimates (WHO, 2014). DALYs saved by the NAP were estimated by using Modes of Transmission (UNAIDS, 2012) tool, Spectrum epidemiological modeling software (Futures Institute, 2014a) and mathematical calculations based on natural history of disease and possible disease outcomes. DALYs saved by the IPL were estimated by using programmatic coverage data, vaccine effectiveness, case fatality ratio, and mathematical calculations based on susceptibility of disease in absence of vaccination and possible disease outcomes based on natural history of diseases.

Total DALYs saved by each intervention were then divided by the total expenditure incurred by each intervention to calculate the cost per DALY saved as a cost-effectiveness measure.

Based on the cost-effectiveness analysis, technical efficiency of each intervention was calculated for further comparative analysis of technical efficiency of each intervention. Total number of DALYs saved by each intervention was compared with the resources used or likely to be used by each intervention for the actual and optimum (full targets achieved) scenarios to calculate the cost per DALY saved in each scenario and measure their technical efficiency.

It was assumed that each project’s budget and performance targets were designed to achieve the maximum results by utilizing minimum resources. Therefore, in the
optimum scenario the project’s performance would be 100% efficient if the project had fully achieved all its targets by utilizing 100% of the allocated budget.

Based on this assumption, the technical efficiency of each selected project was measured by comparing the cost per DALY saved by the project with the likely cost per DALY saved for the optimum scenario.

**Critical analysis of Step 4 techniques:**

Step 4 employed a quantitative analysis of cost-effectiveness of outcomes using mathematical modeling exercise with actual, counterfactual and optimum scenario modeling and economic measurement of outcomes through DALY. While DALY provides a common measurement of outcomes for different projects with different objectives, a comparative analysis solely based on DALY can be dubious. This step, therefore, further calculated the technical efficiency of each project by comparing the cost per DALY saved against the likely cost per DALY saved in optimum performance scenario. While the cost and performance related data for each project are used for analysis, as explained in Chapter 8, this technique also relies on some assumptions for the different performance scenario and disease outcomes. As can be seen in the uncertainty analysis in Chapter 8, the analysis provides only some indications of DALY saved and technical efficiency of each project and they cannot taken as absolute measurement to draw a conclusion.

**Step 5: Comparative outcome patterns evaluation with DEA and correlations:**

**Approach:**

Further analysis of operations and outcome patterns of each intervention was conducted by comparing the selected projects’ relative economy (inputs/cost of inputs), efficiency (outputs/cost of inputs), and effectiveness (outcome/cost of inputs). As explained in detail in Chapter 9, this research used a non-parametric linear programming technique, such as DEA, for comparing the outcome pattern and relative efficiency of each intervention in terms of output-input balance.

**Method:**

As explained in Chapter 9, the DEA model was developed by using the ‘Solver’ add-in in the Microsoft Excel programme. In order to calculate the relative efficiency of an
intervention for a particular aspect, the DEA model asked for solving the problem of maximizing the output of that intervention in terms of single units of inputs by changing the relative weight of all inputs and outputs used by that intervention and other comparing interventions through linear programming. The DEA model considered the constraints that the relative weights of the inputs and outputs cannot be negative and weighted outputs cannot be greater than the weighted inputs.

The DEA model was run for calculating relative efficiency of each intervention in terms of economy of procurement of goods and equipment, economy of hiring staff and technical assistance, efficiency of obtaining project results, and efficiency of project outcomes.

**Data sources and calculations:**

**Economy:** In this research, economy was calculated as a ratio of inputs and cost of inputs. However, this calculation was not possible in cases where the inputs were very different from each other. This research, therefore, only indicatively tested the costs of some common items in calculating the economy of procurement of those items. These items included laptop and desktop computers, printers, motorbikes, and projectors, which were procured by all three projects. This was done by selecting these items from the asset lists of each project and collecting the information on their total number and cost of procurement.

Economy was also compared for acquiring human resources by each project. This was done by collecting the list of local staff and technical assistance used by each project during January to December 2012 with their total amount of inputs (in terms of person-months) and total salaries and fees paid for procuring the inputs.

The sources of inputs and outputs data for the analysis of economy was project asset lists, project staff list, and project financial reports.

**Efficiency:** The efficiency was calculated as a ratio of outputs and inputs. The outputs of each project was measured for achievement of project results against their performance
targets and achievement of health system strengthening related outputs as perceived by the stakeholders during the stakeholder interview process in Step 3.

The source of information for measuring project results was the performance framework and progress reports of each project with baseline data, targets, and achievements against agreed indicators. Using these data the rate of change from the baseline or achievement by the project against an indicator was compared with the likely rate of change if the project achieved the target fully to derive the percentage of project’s output achievement value against an indicator. In cases of negative achievements and over achievements, the percentage values were adjusted by converting the negative achievements to 0% and over achievements to 100% to keep them within the range of 0% to 100%. Achievements against all indicators were then considered to derive total project outputs.

For the health system strengthening outputs, the average scores obtained by each project in the relevant questions in the stakeholder interview questionnaire were considered. The average scores were then converted to percentages by comparing them with the highest possible scores (i.e., highest scores in the ranking questions) to input the data in the DEA model.

In order to calculate the inputs used by each project, the percentage of expenditure incurred by each project was calculated from the total expenditure and budget of each project.

**Effectiveness:** Like the efficiency the effectiveness of an intervention was calculated as a ratio of the value of outcomes (both project level and health system level outcomes) and inputs.

The value of project outcomes was measured by the DALYs saved by each project as calculated in Step 4 for cost-effectiveness analysis. For considering health systems related outcomes, the DEA model used the scores from the related questions of stakeholder interview questionnaire as obtained in Step 3 and converted the average scores to a percentage by comparing them against their best possible scores (i.e. highest...
ranks in the ranking questionnaire) to obtain the health systems outcome performance ratio. This health system outcome performance ratio was then applied to the total health systems strengthening related investments by each project.

To calculate the total investment related to health system strengthening by each project, information was collected from the project financial reports and audit reports to identify the total investment by a project technical assistance, training, durable goods, equipment and infrastructure. Regardless of the objective of the project, it was assumed that these investments on technical assistance, training, durable goods, equipment and infrastructure would have longer lasting effects on the health system’s performance improvement. The health system outcome performance ratio of each project was then applied to the total investment on health system strengthening by each project to derive a health system strengthening outcome value.

Both the value of project outcomes (i.e., DALYs saved) and health system outcomes values were then compared with the total expenditure by each project. Data related to the total expenditure were obtained from project reports.

**Interpretation of findings:**
Results from the DEA model were interpreted as ‘relative’ in comparison to the performance of each other. These results were compared to the stakeholders’ views as obtained in Step 3 and cost-effectiveness analysis results obtained in Step 4 to test the validity of constructions.

This was complemented by a statistical analysis of correlation and hypothesis testing for the results for compliance with the Paris principles as obtained in Step 3 (balanced scorecards) and other variable measures including cost-effectiveness, economy, efficiency, effectiveness, and sustainability. Statistically significant correlations were then compared with their underlying programme theories and the findings from other studies to infer any causal relationships between the variables.
**Critical analysis of Step 5 techniques:**

Step 5 employed further quantitative analysis using DEA technique and statistical analysis of correlations. A DEA technique can compare the performance and technical efficiency of a project against the performance of its peers. However, DEA works well with a large number of projects with same outcome so that they can be compared against a best performance frontier and against their use of inputs. However, use of DEA in this research was seriously limited by the sample size of only 3 projects. As 2 of these projects would form a best performance frontier anyway, the technique did not provide much information about these two projects. However, it still indicated at least one sub-performing project comparing to the performance of the other two projects.

Similarly, the correlation analysis was highly constrained by the sample size and could only provide some indicative and inconclusive findings. Nevertheless, the research indicated that both DEA and correlation analysis would be more useful with a larger sample size.

**Step 6: Synthesis, assessment and interpretation of analysis:**

This final step of the research attempted a synthesis of all the findings from the context, mechanism, outcomes, and outcome pattern analysis in order to be able to evaluate them holistically. This employed a realist synthesis approach with mixed methods using statistical analysis of correlation and hypothesis testing and critical analysis of findings.

Through the comparison of findings in Step 1 to Step 5, the comparative evaluation attempted to seek answers to the main research questions of the thesis. As can be seen in Chapter 10 and 11, this final step of the evaluation process involved analysis and assimilation of all findings and judging them against their underlying programme theories, contexts, and possible impacts for health outcomes and the health systems in comparison to each other.

**Critical analysis of Step 6 techniques:**

Step 6 employed a realist synthesis approach with critical analysis of research findings from all other steps. This step provided comparison, interpretation, and evaluation of findings from the series of pilot investigations and captured the lessons learned.
However, like Step 1 and Step 2, this step also depended on the researcher and can be subjective. Nevertheless, with some indication of the evaluation of three aid funded projects, this step also evaluated the feasibility of using these evaluation approach and techniques in evaluating large scale national level health projects in a fragile state situation.

5.5 Quality control of research

Apart from following the six steps described above, this research employed several quality control measures to ensure the robustness, validity and reliability of the research findings. These measures included obtaining ethical approval, triangulation, and checking against triple-A requirements as explained below.

5.5.1 Ethical approvals

The research proposal including the study protocol, stakeholder interview questionnaire, and informed consent form for participants was reviewed and approved by the Research Advisory Committee of the Department of Preventive and Social Medicine of the University of Otago, the Technical and Ethical Committee for Health Research in Timor-Leste, and the University of Otago Research Ethics Committee. In addition, approval was obtained from the Minister of Health of Timor-Leste to access data and interview MoH officials. Each participant of the stakeholder interview took part based on informed consent with option for anonymity. Copies of informed consent form and approval letters can be found in Annex 2, 3 and 4.

5.5.2 Triangulation

This research used triangulation to increase validity of the research findings. As defined by Jick (1979), triangulation is the use of multiple methods, including qualitative and quantitative techniques, to study the same phenomenon for the purpose of increasing the credibility of a study.

At the data level, the triangulation techniques used in this research involved cross-examination and double-checking of data obtained through different methods and sources to ensure the validity of findings and to reduce any possible factual error.
However, at the synthesis and interpretation level, the triangulation techniques were used to compare the empirical evidence obtained from the qualitative and quantitative methods with the programme theories. This can be illustrated by using the ‘triangulation triangle’ adapted from Erzberger and Kelle (2003):

![Triangulation Triangle](image)

**Figure 5.5: Illustrating the triangulation technique**

*Source: (Ostlund et al., 2011, p. 371)*

This research elicited the underlying programme theories of three interventions under evaluation and employed both qualitative and quantitative methods to collect evidence from multiple sources. Findings from these multiple sources are presented side by side to compare to what extent they agree to or differ from each other with possible explanations. The high-level triangulation allowed testing of programme theories and logical inferences against empirical evidence for increased validity.

### 5.5.3 Checking against ‘triple-A’ requirements

The validity and credibility aspects of an evaluation also need to meet the criteria for alignment, aggregation, and attribution, which are sometimes known as ‘triple-A’ requirements (Leeuw, 2003; White, 2005).

**Alignment:**

Alignment is whether the data collected at the micro level are aligned to greater objectives at the macro level to contribute to the impact. In this case of evaluative research, all data collected were aligned to the underlying programme theory and
evaluation framework. The data analysis was further checked through the process of triangulation for greater alignment with the programme theory and underlying objectives. Chapter 9 provides a comparison of results from different methods for validity check whereas Chapter 10 attempts to synthesize and interpret the findings in the context of underlying programme theories.

**Aggregation:**
Aggregation refers to the issue whether the reporting systems produce data of various activities that can be meaningfully added up for comparison. While the balanced scorecards method employed in Step 3 measured stakeholder views on various aspects of the interventions in a likert scale, the cost-effectiveness analysis in Step 4 used a single measure such as cost per DALY saved for outcome values. The DEA technique used in Step 5, on the other hand, was capable of analyzing multiple inputs and multiple outputs in different forms and categories. Therefore, this research also met the criteria for ‘aggregation’.

**Attribution:**
The third requirement is related to the issue of attribution of an intervention’s activities to the outcomes at the country level. Since this research tried to evaluate national level interventions with the presence of government and other funded interventions with similar objectives, linking a percentage of outcomes to the intervention activities was a challenge.

Developed by Mayne (2001), “contribution analysis” provides an alternative approach of dealing with such attribution issues especially in the situations where a positivist experimental design of research is not possible or appropriate.

By inferring an underlying programme theory and then focusing on the question to what extent the observed results (positive or negative) are the consequence of an intervention activity, contribution analysis assesses causal questions and inferences to derive a more plausible contribution story each intervention is making. Mayne (2001) offered the following step-by-step approach to contribution analysis designed to help
managers, researchers, and policymakers reach the conclusion about the contribution an intervention has made to particular outcomes:

Table 5.3: Steps in contribution analysis

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1: Develop the results chain</td>
<td>Describe the programme theory model/programme logic/results chain describing how the programme is supposed to work. Identify as well the main external factors at play that might account for the outcomes observed. This programme theory should lead to a plausible association between the activities of the programme and the outcomes sought. Some links in the results chain will be fairly well understood or accepted. Others will be less well understood or subject to explanations other than that the programme was the “cause.” In this way you acknowledge that attribution is indeed a problem.</td>
</tr>
<tr>
<td>Step 2: Assess the existing evidence on results</td>
<td>The results chain should provide a good idea of which intended results (outputs, intermediate and end outcomes) could be measured. What evidence (information from performance measures and evaluations) is currently available about the occurrence of these various results? The links in the results chain also need to be assessed. Which are strong (good evidence available, strong logic, or wide acceptance) and which are weak (little evidence available, weak logic, or little agreement among stakeholders)?</td>
</tr>
<tr>
<td>Step 3: Assess the alternative explanations</td>
<td>Outcomes by definition are influenced not only by the action of the programme but also by external factors — other programmes, as well as social and economic factors. In addition to assessing the existing evidence on results, there is a need to explicitly consider the extent of influence these external factors might have. Evidence or logical argument might suggest that some have only a small influence and that others may have a more significant influence on the intended results.</td>
</tr>
<tr>
<td>Step 4: Assemble the performance story</td>
<td>With this information, you will be able to set out your performance story of why it is reasonable to assume that the actions of the programme have contributed (in some fashion, which you may want to try and characterize) to the observed outcomes. How credible is the story? Do reasonable people agree with the story? Does the pattern of results observed validate the results chain? Where are the main weaknesses in the story? There always will be weaknesses. These point to where additional data or information would be useful.</td>
</tr>
<tr>
<td>Step 5: Seek out additional evidence</td>
<td>To improve your performance story you will need additional evidence. This could involve information on both the extent of occurrence of specific results in the results chain and the strength of certain links in the chain. A number of strengthening techniques that you might be able to adopt are outlined in this work.</td>
</tr>
<tr>
<td>Step 6: Revise and strengthen the performance story</td>
<td>With the new evidence, you should be able to build a more credible story. One that a reasonable person will be more likely to agree with. It will probably not be foolproof, but will be stronger and more credible.</td>
</tr>
</tbody>
</table>

Source: Adapted from (Mayne, 2001, p. 9)

Acknowledging the complexity of situation and non-experimental design of evaluation, this research employed contribution analysis as suggested above to assess the outcomes attributable to each intervention.

5.6 Conclusion

Against a backdrop of an evolving concept of aid effectiveness and development effectiveness, changing paradigms in the ontology, epistemology, and methodology of impact evaluations, this chapter presents the overview and justifications of theoretical framework and methods used in this research. Through a series of pilot studies, this research evaluated and compared different components of the selected interventions and investigated if adherence to the Paris principles of aid effectiveness were actually
correlated with the development effectiveness of an intervention in a fragile state context. For this purpose the research employed a theory-driven evaluation approach following the realist evaluation tradition with mixed methods.

While this chapter presents an overview of the research design and methods used in different steps of this research, the subsequent chapters of this thesis explain in further detail the interventions being evaluated, particular techniques used in different steps, their data requirements, feasibility, and strengths and weaknesses for using them in a comparative evaluative research.
6 State fragility and health systems: a case study of Timor-Leste

6.1 Introduction

The relationship between achievement of health indicators and health systems in fragile state contexts is well documented (Berman & Bitran, 2011; Canavan et al., 2008; Carment, Prest, & Samy, 2008; Chauvet & Collier, 2004; Cometto, Fritsche, & Sondorp, 2010; Eldon et al., 2008; Galichet et al., 2010; Glassman & Savedoff, 2011; Goodhand & Sedra, 2010; Gruber, 2011; Haar & Rubenstein, 2012; Kaplan, 2008; Kruk et al., 2010; Leischow et al., 2008; Lockhart, 2011; Management Sciences for Health, 2007; Nagai et al., 2007; Newbrander, 2007). The first principle of the Principles for Good International Engagement in Fragile States and Situations underscores the need for taking “context as the starting point” (OECD, 2007d). In order to evaluate the effectiveness of any health sector intervention in a fragile state context, it is, therefore, very important to have a sound understanding of the country’s fragility and health systems context including different enablers and constraints that have a likely effect on the performance of the intervention in the health sector.

This chapter attempts to identify and assess Timor-Leste’s health sector performance risks associated with the factors of state fragility and health. This chapter is organized in three parts-- Part A provides a state fragility analysis for Timor-Leste; Part B provides a health system analysis; while Part C further analyzes the key considerations from the fragility and health system analysis for their possible influence on the outcomes of health interventions under evaluation.

6.2 State fragility analysis

Timor-Leste has been classified as a post-conflict fragile state. It is a core member of the group g7+ comprising of 19 fragile and conflict affected countries. Since its independence referendum in 1999, the country experienced some conflict or instability on average every two years (World Bank, 2012d). However, there has been relative stability over the last five years and the country has established a plan to transition from
a phase of internal security to the next development phase of state building (Government of Timor-Leste & United Nations Integrated Mission in Timor-Leste, 2011). Despite this relative stability, the economic and public health situation of Timor-Leste can be affected by the remaining fragility and post-conflict reality of the country, if the country cannot mitigate the risks and remaining stressors of fragility (World Bank, 2013a).

6.2.1 Methods of state fragility assessment

As Zoellick (2008) points out, “Academic analysis of how to measure, categorise or rank fragility vary, but three main characteristics stand out: ineffective government, poverty and conflict” (p.69). A few institutions including the World Bank, Center for Systemic Peace, and Fund for Peace regularly publish indices to measure and rank state fragility.

The World Bank assesses countries’ fragility through a “harmonized list of fragile situations”. This list considers World Bank’s and Asian Development Bank’s scores on ‘Country Policy and Institutional Assessment’ (CPIA) and the presence of a United Nations and/or regional peace-keeping or peace-building mission during the past three years. If a country either has a harmonized average equal to or lower than a cut-off point (which is 3.2 currently), or presence of a United Nations or a regional peace building or peacekeeping mission in the past three years, then the country is considered fragile (World Bank, 2015).

The Center for Systemic Peace, on the other hand, uses an index based on a country’s scores on the effectiveness and legitimacy in security, political, economic and social areas. Based on the overall score, this fragility index classifies countries into five categories: 0 to 3 (no fragility), 4 to 7 (low fragility), 8 to 15 (medium fragility), 16 to 19 (high fragility) and 20 to 25 (very high fragility).

The Fragile States Index of the Fund for Peace (2014) also considers a number of social, economic, political and military indicators to measure a country’s fragility or degree of failure every year. Social indicators of the Fragile State Index include demographic pressure and pressure on the population due to disease and natural disaster, pressure related to refugees and internally displaced people, presence of group grievances
including tensions and violence between groups, and human flight and brain drain due to stress and lack of opportunities in the country. Economic indicators include uneven economic development with ethnic, religious or regional disparities, and poverty and economic decline. The political and military indicators include state legitimacy including corruption and lack of representativeness in the government; status of public services such as health, education and sanitation services; status of human rights and the rule of law; strength of security apparatus; presence of factionalized elites; and degree of external intervention in internal affairs.

Meanwhile, g7+, the self-identified group of 19 fragile and conflict affected countries, developed a fragility spectrum tool ‘to assist fragile and conflict affected states to identify the nature of their own fragility and plan a pathway of transition towards stability and development’ (International Dialogue on Peacebuilding and Statebuilding, 2011). Aligned with five agreed Peace-building and State-building Goals (PSG) such as legitimate politics, security, justice, economic foundation, and revenues and services, this diagnostic tool calls for a “country-led assessment on the causes and features of fragility and sources of resilience as a basis for one vision, one plan” for a fragile country.


While the indices and self-assessments related to state fragility provide good indications of overall status and symptoms of state fragility, they often miss the underlying factors of fragility. Mcloughlin (2010) identified a list of causes and characteristics of state fragility that are more prevalent in the related literature. Based on that list, this chapter employs the following set of internal and external factors of state fragility to the history and current situation of Timor-Leste to see how the presence and interplay of these factors contribute to Timor-Leste’s state fragility:
Table 6.1: Common factors of state fragility

<table>
<thead>
<tr>
<th>Factors of fragility</th>
<th>Indicators to assess</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structural and economic factors</td>
<td>Poverty, low income and economic decline, violent conflict, presence of armed insurgents, natural resource wealth or lack of natural resource wealth, geography ('bad neighbours'), and demographic stress (including urbanization);</td>
</tr>
<tr>
<td>Political and institutional factors</td>
<td>Crises of state legitimacy and authority, bad governance, repression of political competition, weak (formal) institutions, hybrid political orders, institutional multiplicity, political transitions, succession and reform crises in authoritarian states, state predation, and neo-patrimonial politics;</td>
</tr>
<tr>
<td>Social factors</td>
<td>Horizontal inequalities, severe identity fragmentation, social exclusion, gender inequality, lack of social cohesion (including lack of social capital), and weak civil society;</td>
</tr>
<tr>
<td>International factors</td>
<td>Legacy of colonialism, international political economy, climate change, global economic shocks (including food prices).</td>
</tr>
</tbody>
</table>

Source: Adapted from (Mcloughlin, 2010, p. 16)


6.2.2 State fragility indices for Timor-Leste

The state fragility indices from both the World Bank and the Center for Systemic Peace show gradual improvement of fragile situations in Timor-Leste. The average harmonized score for Timor-Leste in the World Bank’s “Harmonized List of Fragile Situations” improved from 2.77 in 2010 to 2.96 in 2012 and further to 3.16 in 2013. Timor-Leste’s average harmonized score in this list in both 2014 and 2015 is 3.2 (World Bank, 2010, 2012a, 2013g, 2014, 2015).

The Harmonized List of Fragile Situations in 2013 (World Bank, 2013g) consisted of a list of 35 countries with a harmonized average score ranging from 1.13 (Somalia) to 3.69 (Nepal). According to this harmonized list, a country is considered fragile if a country’s average harmonized score is equal to or less than 3.2 or if there is a presence of United
Nations or regional peacekeeping or peace building mission in that country in the last three years. The United Nations Peacekeeping mission withdrew itself from Timor-Leste in December 2012. With a current harmonized score of 3.2, it looks like Timor-Leste is moving towards a graduation from the fragile situation.


The State Fragility Index and Matrix evaluates all independent countries in the world in which the total country population is greater than 500,000 against their scores on both Effectiveness and Legitimacy in four performance dimensions: Security, Political, Economic, and Social, giving a total of eight matrix indicators. Each of the matrix indicators is rated on a four-point fragility scale: 0 “no fragility,” 1 “low fragility,” 2 “medium fragility,” and 3 “high fragility” with the exception of the Economic Effectiveness indicator, which is rated on a five-point fragility scale (including 4 “extreme fragility”). The State Fragility Index, then, combines scores on the eight indicators and evaluates the overall situation of a country ranging from 0 “no fragility” to 25 “extreme fragility” (Marshall & Cole, 2012).

Timor-Leste's scores on the Center for Systemic Peace's State Fragility Index and Matrix indicate that Timor-Leste is in the category of medium fragility compared to other countries in the world. However, Timor-Leste's fragility situation is gradually improving.

Apart from the World Bank's Harmonized List of Fragile Situations and the Center for Systemic Peace's State Fragility Index and Matrix, the Fragile State Index of the Fund for Peace provides indications for a country's fragility situation against a number of similar indicators. Table 6.2 and Figure 6.1 below show Timor-Leste's progress in Fragile States Index from 2007, 2008, 2009 and 2013 compared to world average in 2013:
Table 6.2: Fragile State Index for Timor-Leste 2007-2013

<table>
<thead>
<tr>
<th>Year</th>
<th>Demographic Pressures</th>
<th>Refugees and IDPs</th>
<th>Group Grievance</th>
<th>Human Flight</th>
<th>Uneven Development</th>
<th>Poverty and Economic Decline</th>
<th>Legitimacy of the State</th>
<th>Public Services</th>
<th>Human Rights</th>
<th>Security Apparatus</th>
<th>Factionalized Elites</th>
<th>External Intervention</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>2007</td>
<td>8.1</td>
<td>8.5</td>
<td>7.1</td>
<td>5.3</td>
<td>6.5</td>
<td>8.5</td>
<td>9.5</td>
<td>7.9</td>
<td>6.9</td>
<td>9.0</td>
<td>8.8</td>
<td>8.8</td>
<td>94.9</td>
</tr>
<tr>
<td>2008</td>
<td>8.1</td>
<td>8.6</td>
<td>7.1</td>
<td>5.3</td>
<td>6.5</td>
<td>8.2</td>
<td>9.0</td>
<td>8.0</td>
<td>6.9</td>
<td>8.8</td>
<td>8.5</td>
<td>8.8</td>
<td>93.8</td>
</tr>
<tr>
<td>2009</td>
<td>8.4</td>
<td>9.0</td>
<td>7.3</td>
<td>5.7</td>
<td>6.8</td>
<td>8.4</td>
<td>9.4</td>
<td>8.4</td>
<td>7.0</td>
<td>9.0</td>
<td>8.8</td>
<td>9.0</td>
<td>97.2</td>
</tr>
<tr>
<td>2013</td>
<td>8.7</td>
<td>7.4</td>
<td>6.8</td>
<td>6.4</td>
<td>6.7</td>
<td>7.9</td>
<td>8.0</td>
<td>8.5</td>
<td>6.0</td>
<td>8.3</td>
<td>8.5</td>
<td>8.5</td>
<td>91.5</td>
</tr>
<tr>
<td>2013 World Average</td>
<td>6.1</td>
<td>5.1</td>
<td>6.0</td>
<td>5.4</td>
<td>6.3</td>
<td>5.8</td>
<td>6.3</td>
<td>5.7</td>
<td>5.9</td>
<td>5.8</td>
<td>6.3</td>
<td>5.8</td>
<td>70.5</td>
</tr>
</tbody>
</table>


According to the Fund for Peace's "Fragile States Index", Timor-Leste ranked 32 out of 178 countries in 2013 with a total index of 91.5. Timor-Leste ranked 20 out of 176 countries with an index of 94.7 in 2007 and made some improvement in the 2008 index and ranked 25 out of 177 countries with a slightly improved index of 93.8. However, the conflict and security situation in 2008 affected the country's progress and in 2009 Timor-Leste again ranked 20 out of 177 countries with the index deteriorating to 97.2 (Fund for Peace, 2014).

Figure 6.1: Timor-Leste's scores on Fragile State Index 2007-2013

The Fragile States Index rates a country’s fragility against the twelve indicators presented in the table and figure above in a scale of 0 (no fragility) and 10 (extreme fragility) and then compiles a total score out of 120. It can be seen that between 2009 and 2013 indexes, Timor-Leste made some good progress in addressing refugee and internal displacement issues. In fact, conflict in 2006 saw a high number of people in Timor-Leste internally displaced and were forced to live in the Internally Displaced People (IDP) camps. Thanks to using the Petroleum Fund the country provided cash benefits to the internally displaced people to rehabilitate them in their homes during 2008 and 2009.

Timor-Leste also made good progress in increasing the legitimacy of state with stable, democratic political process reflected in the conflict free election in 2012. Good progress can also been seen in the human rights area and with an index of 6.0 the country is very close to the world average of 5.9 in this area.

Thanks to cash injection and social benefit programmes of the government, Timor-Leste made some progress in a number of other areas such as reducing group grievances, reducing poverty and economic decline and reducing gaps among factionalized elites. The state security apparatus such as police and military were strengthened and with withdrawal of the UNMIT and ISF in 2012, the level of external intervention was further reduced.

However, with high population growth and a large young population the country’s demographic pressure is continuously increasing. The country experienced increased human flight due to lack of opportunities in country. With increased pending cases in the justice sector and lack of public services in remote rural areas, the country’s index on public services slightly deteriorated. The country also made very little progress in addressing uneven economic development and high level of disparities exist between rural and urban areas.
Except for the human rights indicator, Timor-Leste’s indexes on all other indicators are still much higher than the world averages indicating that the country still needs to improve a lot to overcome the factors and remaining stressors of state fragility.

### 6.2.3 Timor-Leste’s self assessment of fragility

Based on the New Deal for engaging with fragile states (International Dialogue on Peacebuilding and Statebuilding, 2011), Timor-Leste conducted a self-assessment of state fragility by using an in-country participatory, consultative process involving 41 institutions including “State Institutions, Development Partners, Civil Societies Organization, Local Authorities from two selected districts and Universities” (Fragility Assessment Team at Ministry of Finance, 2013).

The Fragility Assessment of Timor-Leste assessed Timor-Leste’s current status against five peace and state-building goals (PSG) such as legitimate politics, security, justice, economic foundations, and revenues and services and tried to measure the country’s progress against each PSG against five stages where Stage 1 represented crisis and Stage 5 resilience. Stage 2, 3 and 4 respectively represented rebuilding and reformation, transition, and transformation. Progress against each PSG was measured on various dimensions. The following table summarizes overall findings of Timor-Leste’s fragility assessment:

**Table 6.3: Summary of Fragility Assessment of Timor-Leste, 2012**

<table>
<thead>
<tr>
<th>PSG</th>
<th>Fragility stage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Stage 1 Crisis</td>
</tr>
<tr>
<td>1. Legitimate Politics</td>
<td>3</td>
</tr>
<tr>
<td>2. Security</td>
<td>3.5</td>
</tr>
<tr>
<td>3. Justice</td>
<td>2</td>
</tr>
<tr>
<td>4. Economic foundation</td>
<td>2.5</td>
</tr>
<tr>
<td>5. Revenues and services</td>
<td>3</td>
</tr>
</tbody>
</table>

Source: (Fragility Assessment Team at Ministry of Finance, 2013)

Overall, the fragility assessment found Timor-Leste’s highest progress in stabilizing the country’s security situation (PSG 2) and some good progress in the areas of legitimate politics (PSG 1) and revenues and services (PSG 5). However, the assessment identified
the areas such as justice (PSG 3) and economic foundation (PSG 4) requiring more attention to improve in future (Fragility Assessment Team at Ministry of Finance, 2013).

As led by a country-driven process, Timor-Leste’s self-assessment of state fragility provides a good insider’s view of the current status of the country’s fragility situation. However, for a more in-depth analysis, we need to assess the presence and interplay of some common factors of state fragility as found in scholarly literature, applying them in Timor-Leste’s situation.

6.2.4 Factors of fragility in Timor-Leste

**Structural and economic factors:**
Timor-Leste has a combination of structural and economic factors of fragility including poverty and inequality, conflict, natural resource vulnerability, and demographic pressure.

**Poverty and inequality:**
As we have seen in Chapter 3, thanks to increasing withdrawal from the Petroleum Fund, Timor-Leste had massive economic growth in recent years. However, the benefits of this economic expansion did not reach the districts and rural areas. Still more than one third of the population is living below the poverty line. More than two thirds of the population lives in rural areas with a much lower literacy rate and poor access to good drinking water, education and basic healthcare services. High inflation, rising commodity price, and lack of economic opportunities are making their livelihoods increasingly more difficult.

**Conflict:**
As we have seen in the state fragility indices, Timor-Leste is emerging out of conflict. The country had long resistance movements and guerrilla warfare from 1975 to 1999. The 1999 independence referendum was followed by armed violence by the pro-Indonesian militia backed up by the Indonesian military.

During 1999 and following the restoration of independence in 2002, much effort was made to disarm the militias and war veterans. However, fresh conflict started in 2006
when about 40% of the military force deserted the military, carrying arms with them posing huge threat to the country’s stability and security. The situation continued till 2008 when there was an attack on the President and the Prime Minister of the Republic by a group of renegade forces led by their ringleader. The ringleader, however, was killed in the clash with the President’s security troupe and after the incident the government took enhanced efforts to resolve the issue by offering the renegade soldiers legal protection, cash incentives and jobs in private security companies to encourage them to surrender arms and get rehabilitated into normal life.

**Natural resource wealth and vulnerability:**

Although Timor-Leste’s Petroleum Fund is independently managed by the central bank and it is one of the best practices of the world in terms of transparency, government’s over dependency on the Petroleum Fund to meet the state expenses can lead to longer-term fragility.

Timor-Leste is the second most oil depended economy in the world. However, unlike other oil rich middle-income countries such as those in the Middle-East, Timor-Leste’s oil reserves are limited and may last for only another 15 years (Scheiner, 2014). This leaves the country with the enormous challenge of using oil income to create a more sustainable diversified economy by 2025 to avoid a so-called “resource curse”* and “Dutch disease†” (Lundahl & Sjoholm, 2006).

* In economics, the term “resource curse” denotes a situation in which a country with an abundance of non-renewable resources experiences negative economic consequences from stagnant growth or adverse effect on the other economic sectors. The resource curse occurs if the country begins to focus all of its energies on a single industry, such as petroleum, and neglects other major sectors resulting in an apparent relationship between increased exploitation of natural resources and a decline of production in the manufacturing or agriculture sector.

† A commonly cited example of the resource curse is the “Dutch disease”. The “Dutch disease” was a situation that occurred in the Netherlands following a large natural gas discovery in 1959 and the country began to target its economic focus on this high-income industry. This resulted skilled workers from other sectors transferring to the resource sector with higher wages that made the national currency as well as production from other manufacturing sectors less competitive. So other industries especially the manufacturing sector began to suffer. (Source: [http://www.investopedia.com/terms/r/resource-curse.asp](http://www.investopedia.com/terms/r/resource-curse.asp) accessed in November 2013)
Demographic pressure:
As discussed in Chapter 3, 62% of Timor-Leste’s population is under the age of 25. With one of the highest birth rates in the world and a large young population, Timor-Leste’s demographic pressure is mounting. The number of unemployed youth is increasing every year with a new batch of young people entering the labour market but very few new jobs created.

Political and institutional factors:
Political and institutional factors of Timor-Leste’s fragility can be traced through an analysis of state legitimacy, institutional weaknesses, and the presence of a neo-patrimonial politics.

State legitimacy:
Timor-Leste set up a number of institutions such as the Public Service Commission, Central Bank, Anti-Corruption Commission, National Petroleum Authority and strengthened the national police and military forces. This helped the country gain increasing legitimacy and authority as a state. However, state institutions in Timor-Leste are just taking shape and still need trained people and further capacity building.

Institutional weaknesses:
The 2006 crisis in Timor-Leste shows institutional weaknesses of the police and military forces when both forces engaged in internal conflicts against each other. Timor-Leste’s 2012 fragility assessment also found institutional weaknesses of the law enforcing authorities in terms of insufficient human resources, inadequate intellectual capacity of the members of the law enforcing authorities, and lack of capacity to enforce laws and to deal with community cases.

As discussed earlier, the justice sector is weak with inadequate human resources, high numbers of pending cases, and lack of land law to deal with the land disputes over three regimes—the Portuguese, Indonesian and Timor-Leste. There is also a prevailing perception of impunity among the elite and members of the law enforcing authorities (Fragility Assessment Team at Ministry of Finance, 2013).
The 2012 Fragility Assessment also identified centralization, weak parliamentary supervision, limited security forces, and interference of different sovereign bodies to each other as ongoing challenges for the country's stability.

There is still some confusion over the institutional roles and authorities of the national police and military forces. Most of the government institutions are still territorial with very little coordination between them over public service delivery.

**Neo-patrimonial politics:**
Fragile states discussion often uses the term “neo-patrimonial politics” related to the African states (Bratton & van de Walle, 1997). ‘Patrimonialism’, as explained by Max Weber ((1922) 1978), is “a system of patron-client rule in which elites exploit public resources and distribute them to political followers in return for loyalty. Neo-patrimonialism describes a situation in which patrimonial and formal bureaucratic rules co-exist” (Mcloughlin, 2010, p. 13). Timor-Leste's current situation indicates some practices similar to neo-patrimonial politics of using public resources for the benefits of a loyal group. For example, 1) after the 2012 elections the ruling party coalition formed a 57 members government with 17 ministers, 14 vice-ministers, and 26 secretaries of state. For a population of just over one million, the size of the Timor-Leste’s government was bigger than many other countries. 2) In Timor-Leste a members of the parliament after serving part of a term are entitled to lifetime salaries and benefits. 3) Total cash benefits given to veterans were much higher than the country’s total investment in the education sector. These facts indicate an uneven distribution of public resources to some privileged groups.

**Social factors:**
Social factors of Timor-Leste’s fragility can be discussed in terms of horizontal inequality, presence of non-reconciled past, remaining refugees in West Timor, language diversity and weak civil society.
**Horizontal inequality between east and west:**

According to Douma (2006) and Kaplan (2008), social exclusion and lack of social cohesion particularly based on gender, ethnicity, class or religion based inequalities and severe identity fragmentation may lead to conflict and fragility. When societal or political groups are systematically deprived or discriminated, they may seek to challenge the system or the state and may seek recourse to violent opposition.

By using a conflict theory framework, Harrington (2007) identified that one of the root causes of the 2006 conflict in Timor-Leste was the cultural difference and horizontal inequalities between the people from the eastern regions and western regions of Timor-Leste and associated land ownership disputes.

According to Harrington (2007), the 2006 conflict was triggered by about 400 western region members of the military submitting a petition to their head seeking remedy against discriminatory treatment. This was followed by mass desertion of their post by about 600 members. However, the chief of the military, with approval of the then prime minister, sacked those 600 members out of the 1500 member military force for their alleged failure to return to their posts. This resulted in demonstrations, threats, conflicts and clash between military, police and civil militia and attacks on the properties and people from the eastern regions by people from the western regions in Dili.

Although the issue seemed submerged after 2008, still the land disputes have not been resolved and there remains confusion between the mandates of the military and police forces as identified in the self-assessment of state fragility.

**Non-reconciled past:**

From 1974 to 1999, Timor-Leste suffered huge human rights violations by the Indonesian military and their supporters in East Timor in the form of killing, torture, rape and disappearance. The country set up a Commission for Reception, Truth and Reconciliation (CAVR - the Portuguese acronym) in 2001 to undertake truth seeking and facilitate community reconciliations of these crimes. In 2005 the Commission presented its definitive report to the President and the Parliament with recommendations for bringing the criminals to the national and international courts for prosecution and
justice. However, the perpetrators, who are mostly based in Indonesia, remain out of the reach of justice.

CAVR was dissolved in December 2005 and replaced by a Post-CAVR Technical Secretariat. In 2008 the Parliament approved a resolution for implementing the CAVR report and gave the authority to the Parliament for implementation. The 2012 Fragility Assessment of Timor-Leste (2013) identified lack of prosecution of human rights violation and serious crimes as one of the challenges for the country’s stability and rule of the law and recommended to implement the recommendations of CAVR.

Refugees in West Timor:
Closely linked to the 1974-1999 human rights violations is the issue of remaining refugees in West Timor. According to International Crisis Group (2011), as many as 250,000 East Timorese were forced to take shelter in West Timor in 1999. Although many of them returned to Timor-Leste, many refugees still remain in West Timor in poor condition and without being well integrated in the host communities. Due to the lack of legal assurance and uncertainty over how they would be accepted in Timor-Leste, they are not sure of their return. The International Crisis Group think that “the unresolved status of thousands of former refugees who fled across the border following a 1999 vote for independence remains a challenge to Timor-Leste's long-term stability” (http://www.crisisgroup.org/en/regions/asia/south-east-asia/timor-leste/B122-timor-leste-reconciliation-and-return-from-indonesia.aspx).

Language diversity:
Timor-Leste is a land of language diversity. Whereas Portuguese and Tetum are national languages and English and Bahasa Indonesia are working languages, the country has another 15 local languages (Hattori et al., 2005). While the presence of so many languages indicate a rich diversified local culture, in conflict situations they may sometimes contribute to fragmentation with possible lack of cooperation, trust, reciprocity and belief in collective wellbeing that are vital in forming a social contract.
Weak civil society:
According to Chabal and Daloz (1999), civil society has a role to act as a check on the accountability of country leaders. Weak civil society, and its lack of capacity to perform this role contribute to an environment conducive to the continuation of fragility. Apart from a few active NGOs and a handful of journalists, civil society in Timor-Leste is still very weak. Except a few who mostly wrote in Portuguese or English, there are no known Timorese writers, poets or intelligentsia. The history of Timorese literature is also not well documented. The country still does not have a publishing house.

The Fragility Assessment of Timor-Leste identified presence of NGOs and civil society as strengths in monitoring government activities and ensuring checks and balances. However, the assessment raised concerns over the sustainability of NGO activities at the end of external funding. The Fragility Assessment report also identified lack of political knowledge of voters, poor access to draft laws for general people, and weak monitoring of public expenditure as other weaknesses of Timor-Leste related to the state fragility (Fragility Assessment Team at Ministry of Finance, 2013).

International factors:
Since fragile states lack the capacity to mitigate the negative effects of external shocks, they are often vulnerable to international factors such as international geo-politics, global economic shock, and climate change. Such factors are also relevant to Timor-Leste’s situation.

International geo-politics:
Timor-Leste is small country with big neighbours such as Indonesia and Australia. Although Timor-Leste’s border dispute with Indonesia is almost resolved to 97% (Ministry of Finance of Timor-Leste, 2013a), it has ongoing issues with Australia over the ownership of the Greater Sunrise oil field in the Timor Sea. Indonesia is the number one source of imports for Timor-Leste. Malaysia and Singapore are interested in trade and investments in Timor-Leste. Portugal also has vested interest in Timor-Leste as the past colonizer and having shares in a bank and telecommunication company in Timor-Leste. China is also strengthening its ties with Timor-Leste (Timor-Leste European Community, 2008).
Timor-Leste is a member of ASEAN Regional Forum and also a member of the CPLP—a multilateral forum for the former Portuguese colonies and Portuguese speaking countries. The country is increasing its ties with the Pacific Island countries.

**Global economic shock:**
Since Timor-Leste’s economy is mostly dependent on petroleum and coffee and the country imports most items, it is highly vulnerable to price changes of petroleum, coffee and import items (including food prices) in the international market. Timor-Leste uses the US Dollar as its currency. So the fluctuating exchange rate of US Dollars against the currency of Timor-Leste’s trade partners also affects its economy. Apart from this, global recession and low interest rates have reduced the rate of return from the investment of its Petroleum Fund.

**Climate change:**
Timor-Leste has a rich bio-diversity but it can be easily affected by climate change-linked events. Global warming, rising sea level and change in the rain pattern can highly affect Timor-Leste’s ground water reserve and agriculture. According to a recent study by Asian Development Bank (2013), economic losses in Timor-Leste from the impact of climate change could reach as much as 10% of Timor-Leste’s annual GDP by 2100 with potential impacts on agriculture, fisheries, tourism, coral reefs, and human health. If not managed proactively, Timor-Leste’s vulnerability to climate change can potentially overstretch its adaptive capacities and lead to mass displacement, destabilization and violence that will repeat the cycle of fragility.

### 6.3 Health system analysis
Understandably Timor-Leste’s health sector has been transitioning through the stages of post-conflict reconstruction, to early recovery, to growth and stabilization. The priorities were for reconstruction and rehabilitation of the country’s health infrastructure, providing essential health services, and developing necessary human resources for health. Against this backdrop, the following is an analysis of the organisation and challenges of Timor-Leste’s health systems in the context of state fragility.
6.3.1 Methods of health systems analysis

As presented in Figure 6.3, this chapter uses health systems analysis framework of WHO, World Bank, GAVI, and the Global Fund (2009) to analyze the history, organization, processes, outputs, and challenges of the six building blocks of Timor-Leste’s health systems:

![Evaluating health systems reform and strengthening: A stepwise approach](image)

**Figure 6.2: Health systems assessment framework**

Source: (WHO et al., 2009, p. 22)

As shown in the above diagram, this chapter analyzes the inputs, outputs, outcomes, and impact related to the health governance and leadership, health service delivery, health information, health work-force, health commodities and technology, and health financing in their broader context of country’s fragility. This analysis is based on the relevant government documents of Timor-Leste (Government of Timor-Leste, 2011; Government of Timor-Leste & United Nations Integrated Mission in Timor-Leste, 2011; Ministry of Finance of Timor-Leste, 2014g; Ministry of Health of Timor-Leste, 2011e, 2011n); Timor-Leste’s country profile, country plans, and country studies as available on the official websites (Central Intelligence Agency, 2013; Nathan Associates, 2008; OECD, 2010; Timor-Leste European Community, 2008; UNDP, 2014; UNDP in Timor-
Leste, 2013; UNESCO, 2009; USAID, 2013; World Bank, 2012d, 2013a); and other published documents as referenced in the following sections.

6.3.2 Health governance

Article 57 of Timor-Leste's Constitution states, “all Timorese citizens are entitled to health care and the State has a duty to promote and protect this right free of charge, in accordance with its capabilities and in conformity with the law” (The Government of Timor-Leste, 2002). The Constitution also asks for running health services under a decentralized participatory management setting as far as possible.

Timor-Leste is a signatory to various international human rights conventions and treaties. The Constitution of Timor-Leste provides an opportunity to translate these human rights commitments to practice by protecting the right to health, medical care and healthy environment for its people and by protecting other human rights that have an impact on health such gender equality; non-discrimination; protection of persons with specific needs (people with disabilities, the elderly, women, children and youth); and health protection at workplace including maternity protection, and occupational health.

The country’s health system is led by the central service unit of the MoH, and includes 6 referral hospitals including a tertiary level national hospital; 13 district health services units; 1 central procurement, storage, and distribution warehouse for drugs, medical products and equipment; 92 community health centers; and 192 health posts. The following figure shows the organization of the management of the health system:
Historically, Timor-Leste’s health sector faced a huge challenge of reorganizing and reconstructing. The conflict in 1999 and departure of the Indonesian health workforce from Timor-Leste left the country with 80% of health infrastructure and equipment destroyed and only a few general medical practitioners left in the country.

After arrival of the international relief agencies, UN agencies, NGOs and donors, the country established the Interim Health Authority in February 2000. The Division of Health Services was set up in August 2000 and the MoH started functioning from September 2001. From there the country is slowly emerging into an organized health systems with constitutional obligations to ensure “healthy people in healthy Timor-Leste” (Government of Timor-Leste, 2002).

The country’s health governance system is still challenged with a complex organization and multiple organizational hierarchies. The Central Service Unit of the MoH is headed by a Director General reporting to the Vice-Minister and Minister of Health. However,
there are a number of other units such as the National Hospital, five Referral Hospitals, National Laboratory, National Blood Centre, National Medical Warehouse (SAMES) and National Institute of Health Sciences that are considered autonomous and headed by a Director General post reporting directly to the Minister. This makes coordination among different units sometimes difficult.

Asante et al. (2011) pointed out the weaknesses of the middle management of Timor-Leste’s health system due to lack of institutional training in management, poor role clarity and multiple priorities making it difficult for the system to function efficiently.

6.3.3 Health service delivery

The public sector in Timor-Leste plays the major role in providing healthcare at all levels with an expanding private sector comprising of national and international NGOs, faith-based organizations, cooperatives and a few private clinics. Table 6.4 summarizes the organization of health services delivery in Timor-Leste.

<table>
<thead>
<tr>
<th>Facility</th>
<th>Location</th>
<th>Public</th>
<th>Private</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health Posts</td>
<td>Suco/Parish</td>
<td>192</td>
<td>0</td>
<td>192</td>
</tr>
<tr>
<td>Community Health Centers</td>
<td>Sub-district</td>
<td>66</td>
<td>26</td>
<td>92</td>
</tr>
<tr>
<td>(CHCs)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maternity Clinics</td>
<td>Selected hospitals and CHCs</td>
<td>42</td>
<td>1</td>
<td>43</td>
</tr>
<tr>
<td>Secondary Referral Hospital</td>
<td>In selected five districts</td>
<td>5</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Tertiary Referral Hospital</td>
<td>In capital Dili</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td></td>
<td>264</td>
<td>26</td>
<td>290</td>
</tr>
</tbody>
</table>

Source: National Health Sector Strategic Plan 2011-2030 (Ministry of Health of Timor-Leste, 2011)

Timor-Leste has a three-tiered healthcare and referral system started from the village level health posts to sub-district level community health centers providing primary level of care, and then 5 district level secondary referral hospitals for secondary level of care, and a national hospital for tertiary level of care. In addition, the MoH in 2008 introduced a village level satellite and outreach health services through Servisu Integrado du Saude Comunidade - Integrated Community Health Services - (SISCa) to improve access to primary health services at the village level (Martins & Trevena, 2014).
Primary health care in Timor-Leste includes a basic health service package (BSP) and national programme services such as immunization, maternal and child health, malaria, nutrition, tuberculosis and HIV/AIDS. A health post is normally staffed with a nurse and a midwife and works as the first level of contact for the community offering a minimum level of curative, preventive and promotive care.

The next level of contact is the community health centre (CHC) that has a physician on the staff. CHCs are designed to provide inpatient and outpatient services and a few CHCs also have dental and laboratory services.

The MoH introduced the Servisu Integrado du Saude Comunidade (SISCa) in 2007/2008 to provide access to basic health services package and preventive measures at the remote suco (village) level. SISCa offers satellite health services at around 474 designated SISCa posts where a team of health care professionals and community volunteers from the nearby CHCs visit on a monthly basis to organize activities and offer services related to: family registration; nutrition assistance and child health promotion; maternal health and family spacing; hygiene, sanitation and malaria prevention; ambulatory primary care; and health promotion (Martins & Trevena, 2014).

Timor-Leste currently has a network of one national tertiary hospital and five district referral hospitals designed to offer emergency, outpatient and inpatient care. Referral hospitals are staffed with general practitioners and specialists in surgery, paediatrics, gyneco-obstetrics and internal medicine.

The only tertiary level facilities for specialized services are available at the national hospital in Dili. For further complicated cases where services are not available in country, the national hospital has referral linkages with facilities in Australia, Indonesia and Singapore. Transportation of a patient from Dili by plane takes around one hour ten minutes to Darwin (Australia), two hours to Bali (Indonesia), and three and a half hours to Singapore.

Closely linked with the referral hospitals, there is also a network of a national blood centre located in Dili and five regional blood centres located at the five referral
hospitals. Blood centres are equipped with blood collection, screening and storing facilities.

In addition to the referral hospital laboratories, there is a national reference laboratory located in the capital Dili for supervision, testing quality control and performing specialist tests.

Health service facilities at the non-government sector in Timor-Leste, mainly run by the faith-based institutions and not-for-profit organizations, are not robust. The MoH recognized 26 non-government clinics at the status of Community Health Centres (Ministry of Health of Timor-Leste, 2011n). In addition, there are a few for profit health clinics in two large cities—Dili and Baucau offering a range of primary health care and dental services. A number of pharmacy shops also sell medicine and medical supplies. However, these clinics and pharmacies are yet to be regulated for quality control by the government.

Apart from the setting up the service facilities, other major challenges related to the health service delivery in Timor-Leste include difficulty to access health care facilities by people living in the remote areas; lack of standard operating procedures in the clinics and hospitals; and lack of role clarification and coordination at the district level between district health services office, hospitals, CHCs and national disease programmes such as HIV/AIDS, malaria, and TB.

6.3.4 Medicine and technologies

Like all other sectors, the health sector of Timor-Leste in recent years focused on building necessary infrastructure for health care services. Efforts for infrastructure building included renovation and construction of health facilities and offices, procuring medical equipment, ambulances and vehicles and ensuring electricity and water supply.

Timor-Leste does not have any in-country facility to manufacture medical equipment or pharmaceutical products. The country has to import 100% of its medical equipment, pharmaceutical products and medical supplies. There is an autonomous Central Medical
Warehouse (SAMES) responsible for procuring and distributing all medicines, pharmaceutical products and equipment.

Pharmaceutical system strengthening is one of the major challenges in Timor-Leste’s health sector as the country is still challenged with a range of issues including registration and selection of drugs, forecasting, budgeting, procurement, storage, logistics management, distribution, quality control and ensuring rational use of drugs and pharmaceutical products. The country has yet to set up pharmaco-vigilence and drug resistance monitoring systems.

6.3.5 Health information management system

The Central Office of Health Management Information System and Surveillance (HMIS-SE) at the MoH collects comprehensive data from all public and private health facilities on a regular basis on key health indicators through a number of prescribed reporting forms. The health posts report to the community health centers (CHCs). The CHCs compile reports from the health posts under them and add their own report and then send the compiled report to the District Health Services office. The District Health Services office then compiles all reports from the CHCs under them and sends the compiled reports to the central HMIS-SE office. The referral hospital and national hospital, on the other hand, send their reports directly to the central HMIS-SE office.

The HMIS-SE enters data from all reports in their database and produces comprehensive health indicators and surveillance reports on an annual basis for programme managers and policy decision makers.

In order to facilitate data collection, validation and reporting the HMIS-SE introduced their own staff at the CHC and district level and provided computers up to the CHC level. However, due to difficulty with the Internet connection and supply of electricity at the district level, data are still reported on paper-based formats requiring multiple levels of data entry at the CHC, district and central level.

Although regular data collection and validation system exists, the health information management system of Timor-Leste is still marked with a number of challenges
including lack of a common understanding and agreement on the minimum level of data collection requirements at various levels; lack of a standard data collection format; multiple reporting requirements by different donor funded and national programmes; lack of internet connection, software, capacity and formal training; and lack of an efficient HMIS database to streamline health information management with regular service delivery for increased efficiency.

6.3.6 Health workforce

Following the 1999 conflict Timor-Leste’s health sector faced a serious health workforce crisis. In 2001 the total health workforce in Timor-Leste shrunk to only 642 with only 3 physicians and 1 dentist left in the whole country.

The table below provides an indication of growing health workforce in Timor-Leste from 2001 to 2008 to 2011:

<table>
<thead>
<tr>
<th>Category/Sub-category</th>
<th>2001</th>
<th>2008</th>
<th>2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>General Medical Practitioners</td>
<td>3</td>
<td>37</td>
<td>75</td>
</tr>
<tr>
<td>Medical Specialists</td>
<td>0</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>Dentists</td>
<td>1</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>Dental Nurses and Technicians</td>
<td>9</td>
<td>29</td>
<td>39</td>
</tr>
<tr>
<td>Pharmacists</td>
<td>0</td>
<td>4</td>
<td>15</td>
</tr>
<tr>
<td>Pharmacy Technicians and Assistants</td>
<td>3</td>
<td>44</td>
<td>116</td>
</tr>
<tr>
<td>Nursing and Midwifery Practitioners</td>
<td>474</td>
<td>1104</td>
<td>1283</td>
</tr>
<tr>
<td>Non-medical Public Health Practitioners</td>
<td>25</td>
<td>85</td>
<td>122</td>
</tr>
<tr>
<td>Medical Technologists</td>
<td>34</td>
<td>107</td>
<td>180</td>
</tr>
<tr>
<td>Optometrists</td>
<td>1</td>
<td>3</td>
<td>13</td>
</tr>
<tr>
<td>Physiotherapists</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Health Management and Support Staff</td>
<td>92</td>
<td>541</td>
<td>621</td>
</tr>
<tr>
<td>TOTAL</td>
<td>642</td>
<td>1958</td>
<td>2481</td>
</tr>
</tbody>
</table>

Source: Human Resources for Health: Country Profile (Ministry of Health of Timor-Leste, 2011e)

Until end of 2011, Timor-Leste had a health workforce of around 2,500. Over two thirds of these staff numbers were posted at District and sub-district level. The peripheral health staff at the community level were supported by a network of health volunteers referred to as PSF. In addition to the PSF network, local community leaders at District,
sub District and community level (aldeia) assist to register populations and mobilize populations to attend the SISCa (integrated Community Health Services Posts).

In order to fill in the health workforce gap Timor-Leste had a medical cooperation agreement with the government of Cuba. Under this agreement that started in early 2004, a few hundred medical practitioners from the Cuban Medical Brigade were deployed in Timor-Leste to provide essential medical services and 1,000 Timorese students were given scholarship to study for a medical degree in Cuba. Additionally a programme was also set up at the National University of Timor-Leste in cooperation with the Cuban Medical Brigade where around 180 Timorese students are currently studying medicine (Asante et al., 2012).

Despite significant expansion of the health workforce from the period of crisis, Timor-Leste’s health sector still faces shortages with low numbers of nurses, midwives, radiologists, physiotherapists, pharmacists, laboratory technicians and managers. Medical students studying in Cuba have started graduating and returning to Timor-Leste from 2010. Absorbing and inducting these new graduates in Timor-Leste’s health sector scenario is another challenge for the country (Ministry of Health of Timor-Leste, 2011e).

Apart from the medical training through the cooperation agreement with Cuba, the Faculty of Health Sciences of the National University of Timor-Leste, the Institute of Health Sciences of the MoH and a private university provide pre and in-service training for healthcare providers.

6.3.7 Healthcare financing

The following table summarizes Timor-Leste’s status and changes in health financing from 2008 to 2013:
Table 6.6: Healthcare financing in Timor-Leste 2008-2013

<table>
<thead>
<tr>
<th></th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total MoH Budget (Government contribution in thousand US dollars)</td>
<td>30,799</td>
<td>32,893</td>
<td>35,692</td>
<td>38,198</td>
<td>44,480</td>
<td>62,058</td>
</tr>
<tr>
<td>Donor funding in Health Sector (in thousand US dollars)</td>
<td>27,700</td>
<td>20,001</td>
<td>30,314</td>
<td>35,782</td>
<td>31,109</td>
<td>36,028</td>
</tr>
<tr>
<td>Total Health Expenditure (Govt and donors in thousand US dollars)</td>
<td>58,499</td>
<td>52,894</td>
<td>66,006</td>
<td>73,980</td>
<td>75,589</td>
<td>98,086</td>
</tr>
<tr>
<td>Percentage of ODA in total health expenditure (excluding out of pocket expenses)</td>
<td>47.35%</td>
<td>37.81%</td>
<td>45.93%</td>
<td>48.37%</td>
<td>41.16%</td>
<td>36.73%</td>
</tr>
<tr>
<td>Estimated per capita health expenditure (Govt &amp; donors)</td>
<td>$54</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Govt per capita expenditure on health</td>
<td>$29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Appropriation of gov't health budget of total gov't budget</td>
<td>6.36%</td>
<td>5.45%</td>
<td>4.69%</td>
<td>3.48%</td>
<td>3.71%</td>
<td>3.77%</td>
</tr>
<tr>
<td>Percentage of ODA in total combined budget of the country</td>
<td>27.55%</td>
<td>26.80%</td>
<td>25.77%</td>
<td>20.56%</td>
<td>17.48%</td>
<td>10.99%</td>
</tr>
</tbody>
</table>


It can be seen that government expenditure for health increased steadily from around $30.8 million in 2008 to $44.5 million in 2012, an increase of over 44%. The donor funding in the same period saw a mixed trend. Total per capita expenditure in health excluding private and out of pocket expenses increased from $54 in 2008 to $64 in 2012, out of which government’s contribution was $29 (52.7%) in 2008 and $37 (58.8%) in 2012. This was still lower than minimum $44 per capita health expenditure required by the government to provide essential health services in low income countries as estimated by the WHO Taskforce of Innovative International Financing for Health Systems (2009).

According to the World Health Statistics (WHO, 2013h), Timor-Leste’s per capita total health expenditure based on PPP international dollar value, in fact, increased very little from $83 in 2000 to $84 in 2010. Government expenditure on health based on the same method, in fact, has decreased from $65 in 2000 to $63 in 2010 (WHO, 2013h).

A closer scrutiny of the government health budget shows that the increase of government health expenditure was largely driven by the salary increase and infrastructure budget of the government and actual health services budget decreased
from $18.12 million (around 59% of the state health budget) in 2008 to $17.26 million (around 39% of the state health budget) in 2012 as it can be seen in the following table and chart:

Table 6.7: State budget by cost category for MoH of Timor-Leste (in thousand US dollars)

<table>
<thead>
<tr>
<th></th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total MoH Budget</td>
<td>30,799</td>
<td>32,893</td>
<td>35,692</td>
<td>38,198</td>
<td>44,480</td>
<td>62,058</td>
</tr>
<tr>
<td>(government</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>contribution)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Salaries &amp; wages</td>
<td>6,177</td>
<td>8,189</td>
<td>10,369</td>
<td>11,063</td>
<td>16,133</td>
<td>22,468</td>
</tr>
<tr>
<td>Goods and services</td>
<td>18,123</td>
<td>16,240</td>
<td>12,695</td>
<td>14,982</td>
<td>17,261</td>
<td>21,088</td>
</tr>
<tr>
<td>Transfer</td>
<td>250</td>
<td>0</td>
<td>1030</td>
<td>10345</td>
<td>9064</td>
<td>11213</td>
</tr>
<tr>
<td>Minor capital</td>
<td>3,481</td>
<td>985</td>
<td>1,277</td>
<td>1,758</td>
<td>1,950</td>
<td>7,289</td>
</tr>
<tr>
<td>Capital development</td>
<td>3,768</td>
<td>7,479</td>
<td>10,321</td>
<td>50</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Percentage of goods</td>
<td>58.84%</td>
<td>49.37%</td>
<td>35.57%</td>
<td>39.22%</td>
<td>38.81%</td>
<td>33.98%</td>
</tr>
<tr>
<td>and services budget of</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>the total MoH budget</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: MoH and Ministry of Finance of Timor-Leste, Budget Books 2008-2013

Figure 6.4: State budget by cost category for MoH of Timor-Leste (in thousand US dollars)

Data source: MoH and Ministry of Finance of Timor-Leste, Budget Books 2008-2013

Timor-Leste's share of public health expenditure as a percentage of general government expenditure (GGE) was around 10% in 2002-2003 (World Bank, 2013a). However, over the past years this percentage decreased and came down to around 6.4% in 2008, 3.5% in 2011 and 3.7% in 2012. This was due to the swelling general government expenditure driven by the infrastructure, institution building and social benefits budgets and considering these areas as higher priorities than health by the government. According to
the World Health Statistics (WHO, 2013h), Timor-Leste's total expenditure on health as percentage of GDP decreased from 8% in 2000 to 5.7% in 2010.

Fortunately, over the past years, the health sector was the top most recipient of donor funding in Timor-Leste to fill in the essential gaps. Donor funding in Timor-Leste’s health sector was estimated as $27.7 million in 2008 (Ministry of Health of Timor-Leste, 2011n). Actual disbursement of donor funding in the health sector was around $20 million in 2009, $30.3 million in 2010, $35.8 million in 2011, and $31.1 million in 2012 (Ministry of Finance of Timor-Leste, 2014a). While the share of donor funding in the country’s combined budget steadily decreased from around 27.6% in 2008 to 17.5% in 2012, share of donor funding in the health sector was as high as 47% and 41% in the same years respectively.

The charts below show the trends in total health expenditure, government contribution and donor contribution (in thousand US dollars) in the health sector and the percentage trend of government funding and donor funding in the health sector and the government’s total combined budget:

![Graph showing trends in total health expenditure, government and donor contribution to health financing in Timor-Leste 2008-2013](image)

**Figure 6.5: Total health expenditure, government and donor contribution to health financing in Timor-Leste 2008-2013**

Figure 6.6: Proportion of government contribution to health budget in Timor-Leste 2008-2013
Data Source: MoH and Ministry of Finance, Timor-Leste Budget Books 2008-2013

Timor-Leste still does not have any health insurance or social security programme in place. Private expenditure in health was 25.3% and out of pocket expenses was estimated as 14.2% of the total health expenditure in 2010 (WHO, 2013h).

Challenges for health financing include monitoring of health expenditure and strengthening PFM systems especially at the district level. Timor-Leste needs to find a longer-term sustainable means for health financing without depending on the Petroleum Fund and donor financing. To ensure equitable health services, the health systems need to reduce the private expenditure and out of pocket expenses for the hard to reach and poor who cannot afford these costs by themselves (Braveman & Gruskin, 2003).

6.4 Discussion

According to the World Development Report on Conflict, Security and Development (World Bank, 2011c), building institutions in a fragile state to help the country get out of fragility may take 20 or 30 years. The World Bank believes that Timor-Leste has a high risk of falling back into the perpetual trap of fragility if it fails to manage the process of transition from fragility to sustainable development (World Bank, 2013a). In the short run, it seems that Timor-Leste successfully used the Petroleum Fund income to address conflict and buy peace. However, the analysis of factors of fragility in Timor-Leste
indicates the remaining elements of state fragility and potential future conflicts if state building and longer-term development objectives of the country are not achieved.

Timor-Leste has developed a long-term national strategic development plan that aspires to create a diversified and socially inclusive economy, with improved quality of life, health and education for all Timorese by the year 2030 (Government of Timor-Leste, 2011). Emerging from the post-conflict fragility Timor-Leste had some commendable successes within a short period. The conflict in the country has been largely under control since 2008. Peaceful elections were held in 2012 and Timor-Leste has achieved the highest percentage (38.5%) of women parliamentarians in the Asia-Pacific region. The establishment and successful management of the Petroleum Fund is another success for Timor-Leste.

In the health sector Timor-Leste has also had some success. Most of the health infrastructures are now being built up. The country now has increased life expectancy, reduced child mortality and reduced incidence of infectious diseases compared to 2002 and 2008. Vaccination coverage has also been increased.

The strengths for Timor-Leste’s country context and health system include relative stability over the last five years, rapid economic growth and assurance of funding from the oil-based economy, assistance from international development partners, constitutional commitment for providing free health services in a decentralized manner to Timorese people, human resources cooperation agreement with Cuba to increase the supply of much needed healthcare workers, MoH’s commitment to bring health services closest to the village level, and having a long term vision of moving towards a “healthy Timorese people in a healthy Timor-Leste” (Ministry of Health of Timor-Leste, 2011 n).

However, the health systems of the country still have weaknesses and challenges along with the threats from its fragile state context. The country still needs to diversify its economy and find non-oil economic activities and growth. Jobs and equitable economic opportunities need to be created to cope up with the demographic dividend of the number of young people entering the labour market every year. Private sector and civil
society need to be strengthened to complement government’s activities and to create stable and accountable governance (World Bank, 2013a).

Timor-Leste's health and human development outcomes are still poor and many of the MDGs are unlikely to be met by the country by 2015. Maternal mortality is very high. The country ranks number one in the world in terms of the percentage of under-weight children under five years of age. More than half of the country’s children under five are stunted indicating chronic mal-nutrition. 31% of the population do not have access to safe drinking water and 61% do not have access to improved sanitation facilities (UNDP, 2014).

The overall disease burden in Timor-Leste is very high (WHO, 2014). The table below shows the major causes of morbidity and mortality in the country:

**Table 6.8: Main causes of morbidity and mortality in Timor-Leste, 2010**

<table>
<thead>
<tr>
<th>Cause of morbidity</th>
<th>Value per 1,000 population</th>
<th>Cause of mortality</th>
<th>Value per 1,000 population</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Acute respiratory tract infections (ARI)</td>
<td>433</td>
<td>1. Bronchopneumonia</td>
<td>0.15</td>
</tr>
<tr>
<td>2. Malaria</td>
<td>104</td>
<td>2. TB (all forms)</td>
<td>0.15</td>
</tr>
<tr>
<td>3. Diarrhoea</td>
<td>60</td>
<td>3. Diarrhoal disease</td>
<td>0.05</td>
</tr>
<tr>
<td>4. Pneumonia</td>
<td>46</td>
<td>4. Malaria</td>
<td>0.05</td>
</tr>
<tr>
<td>5. TB (all forms)</td>
<td>3.78</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Annual Health Statistics Report, 2010 (Ministry of Health of Timor-Leste, 2011c)

The high disease burden and poor achievements against MDGs, as we have seen in the previous chapter, indicate challenges and overall performance of the health systems in Timor-Leste in the context of state fragility.

Several gaps and challenges remain within the country's health systems. For example the governance structure, management capacity, planning and coordination of the health officials needs to be improved further for efficient and seamless functioning of the health system (Asante et al., 2011). A significant section of the country's population living in remote and hard to reach areas are still beyond the coverage of most of the health services. Community participation in health service delivery and bringing health services close to the hard to reach population groups is a challenge for the health services delivery (Martins & Trevena, 2014).
Quality assurance of pharmaceutical procurement and medical supplies and stock out of essential drugs have been frequent issues in Timor-Leste’s health sector that need to be addressed through holistic pharmaceutical systems strengthening. The country still needs to develop an efficient and functioning health management information system at various levels.

For the health workforce the country needs to ensure an equitable distribution of service providers and effective induction of new workforce (graduates returning from Cuba) into the health systems (Ministry of Health of Timor-Leste, 2011e). Finding longer-term sources of health financing without being dependent on the Petroleum Fund and donor support is another challenge for the country.

Timor-Leste’s fragility situation and health systems performance are comparable to issues identified in other fragile situations. It can be seen that conflict affected fragile states are often worse off than the non-fragile states in terms of key health indicators and social determinants of health (Ranson et al., 2007). State fragility is, therefore, one of the most significant factors for the performance of health systems in a fragile situation.

Kruk et al. (2010) pointed out that in the conflict and fragile situations the national health system is also a victim of conflict, with destruction of clinic and hospital infrastructure, the flight of health professionals, and the interruption of drugs and other medical supplies. Timor-Leste tried to address this situation by having an agreement with Cuba to get a high number of Timorese students medically trained to fill in the gaps in health workforce.

Cabral et al. (2013), however, discussed the challenge associated with the rapid deployment of a large number of Cuban trained physicians in Timor-Leste and suggested taking other complementary measures for effective use of this scaling up.

Scaling up and increasing coverage of healthcare services in Timor-Leste also need to be complemented by improving quality of services. Using the experience of South Sudan,
Berendes et al. (2014) suggested a “desperate need for improving the quality of care” in fragile situations.

Rapid economic growth, low human development outcomes and challenging achievements of MDGs, in fact, indicate Timor-Leste’s transition from a fragile state to a state of reconstruction and development. The country is overcoming its conflict affected past but is still in the transition to reach to a state of stability with both economic and human developments. Newbrander et al. (2011) argued:

“Although disease outbreaks or high mortality rates must be the short-term focus in many fragile states, it is essential to start planning at the same time for the transition to longer-term development of the health system. This dual focus requires addressing the elements of a dysfunctional health system and how it can be rebuilt while dealing with immediate health priorities. This two-pronged approach can be seen in several countries” (p.642).

Analysing the post-conflict reproductive health services situation in Timor-Leste, Wayte et al. (2008) argued that focusing primarily on saving lives during the emergency phase of a complex and protracted conflict may result in negligence of other major health problems such as unwanted pregnancies, abortion, STIs, HIV/AIDS, and gender-based violence. Reviewing recent literature on conflict and fragility Haar and Rubenstein (2012) also found that “health, both generally as well as in realms such as mental health, sexual and reproductive health and children’s health, is gravely affected by violent political circumstances” (p.308).

6.5 Conclusion

Timor-Leste’s state fragility context is vulnerable to poverty, lack of economic opportunities (especially in the districts and rural areas), weak institutional services, horizontal inequalities between the people from the east and people from the west of the country, a non-reconciled past, global economic shocks, and climate change. In this context, the health system of Timor-Leste has a daunting task of increasing capacity, coverage, quality, and equity. The health system gaps compounded by the state fragility, in fact, may have a confounding effect on the outcomes of any health interventions in Timor-Leste.
Since state fragility is closely related to the national health system’s performance, the country can better manage the risks and avoid unwanted consequences in the health sector by identifying and assessing the fragility and health sector risks and putting a preparedness and risk-mitigation plan in place.
7 Aid effectiveness and stakeholders’ perspectives: balanced scorecards

7.1 Introduction

A key question in a fragile situation is what management modality of external funding in the health sector is most effective in terms of achieving programmatic objectives and strengthening the country’s health systems. This chapter attempts to learn from the stakeholders to what extent the management modalities of the three selected interventions adhered to the Paris principles of aid effectiveness and how their performances on the achievement of programmatic objectives and health systems strengthening compared with each other. Following the OCED-DAC’s criteria for evaluation of development assistance (OECD-DAC, 2010), this chapter uses stakeholder views for analyzing and comparing the relevance, economy, efficiency, effectiveness, and likely sustainability of each intervention’s organization, processes and outcomes. A balanced scorecard approach is used for presenting the comparison and analysis.

7.2 Balanced scorecards and the research framework

As discussed in Chapter 2, a balanced scorecard approach identifies the long term vision and strategy of a business and links them to the short-term operational control in a causal system to see how the performance of the “soft assets” or non-financial perspectives leads to the achievement of strategic objectives aligned to the financial perspective of the business (Kaplan & Norton, 1997, p. 28). This hierarchical approach focusing on the business vision and strategy and a set of tangible and intangible factors in a cause and effect relationship ensure the identification of the major strategic issues of a business and link them to their relevant performance drivers (Figge et al., 2002).

Due to its ability to present and compare multiple perspectives at various level of hierarchy, the balanced scorecard approach has been modified and adapted for assessing the effectiveness of various application domains including healthcare and health systems (Chan et al., 2010; Commonwealth Fund Commission on a High Performance Health System, 2008; Fairchild, 2002; Gauld et al., 2011; Grebergen &
Like the balanced scorecards, a programme logic model or programme theory (Bickman, 1987, 1990; Bickman & Peterson, 1990; Weiss, 1997) also depicts the underlying assumptions of an intervention and logically links the resources needed to support the intervention activities to produce expected outputs. It then links the activities and outputs needed to realize the intended outcomes of an intervention (United Way of America, 1996; Wholey, 1994). Thus the balanced scorecards and underlying programme theory of an intervention provide an integrative framework for a constructivist and theory-driven evaluation of the three interventions under this research.

7.3 Methods used

As discussed in Chapter 5, following a constructivist approach (Guba & Lincoln, 1989), this part of the research attempted to collect, analyze and evaluate the perspectives of the stakeholders regarding design, process, results, efficiency, effectiveness, and likely sustainability of the three selected interventions by using mixed methods as described below.

7.3.1 Sources of information

The primary sources of information for this particular chapter were the findings from fieldwork, stakeholder interviews, and three focus group discussions conducted in Timor-Leste. In addition, this chapter used relevant project documents, donor documents, and published work including project concept notes, proposal, project appraisal document, work plan, results framework, monitoring and evaluation plan, budget, progress reports, project review reports, account statements, and audit reports as listed in Annex 5.

7.3.2 Fieldwork and stakeholder interviews

The researcher conducted fieldwork in Timor-Leste from April to September 2013, visiting the three selected intervention venues and project teams, talked to the project personnel and staff from different sections of the MoH, and interviewed a total 22 key
informants selected from different sections the stakeholders, and collected project documents and field notes.

The length of each interview ranged from 40 to 60 minutes and consisted of filling in a structured ranking questionnaire, and seeking additional opinions and clarifications from the interviewees. With informed consent from the participants, the interviews were electronically recorded and notes taken by the researcher during the interview process. However, taking an opt-out option in the informed consent form, all participants opted for remaining anonymous in the research documents.

7.3.3 Selection of stakeholders

The sampling and selection of stakeholders to take part in the interviews were purposive based on the relevance and participants’ knowledge about the interventions. Based on a preliminary discussion with and suggestions from the project teams and a senior policy level official from the MoH, a total of 26 stakeholders from donor representatives, policy makers, technical experts, project implementation team members, sector specialists, people with oversight function, and representatives from different sub-systems of the health sector were identified as people available in Timor-Leste with knowledge of one or more of the three selected projects.

All of these 26 stakeholders were approached for their availability and consent to take part in an interview. A total of 22 stakeholders gave informed consent and indicated their availability for an interview. A donor representative took part in an informal discussion with the researcher but preferred not to fill in the ranking questionnaire. Following is a list of different categories of participants that filled in the ranking questionnaire and formally took part in the interview:
Table 7.1: Sampling of stakeholder interviews

<table>
<thead>
<tr>
<th>Category</th>
<th>Number of participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sector coordination representative from Country Coordinating Mechanism</td>
<td>1</td>
</tr>
<tr>
<td>Health sector specialist from Academia</td>
<td>1</td>
</tr>
<tr>
<td>MoH senior management and policy makers</td>
<td>4</td>
</tr>
<tr>
<td>Members of project implementation teams</td>
<td>7</td>
</tr>
<tr>
<td>Counterpart representatives from relevant sub systems of the health system</td>
<td>4</td>
</tr>
<tr>
<td>Technical advisers and donor representatives</td>
<td>4</td>
</tr>
<tr>
<td>External NGO</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>22</td>
</tr>
</tbody>
</table>

Since all participants were not knowledgeable on all three interventions and all items of the questionnaire, participants were allowed to choose freely if they would like or not to respond to an item related to an intervention. Thus the number of responses received to an item varied between interventions.

7.3.4 Focus group discussions

Three focus group discussions were conducted with the three different intervention teams during the field visits. The focus group discussion with the HSSP-SP project team included three members. Three members from the Division of Global Fund took part in the focus group discussion for the NAP. However, only one member from the IPL project team could take part in the focus group discussion related to IPL.

The purpose of focus group discussions were to elicit in-depth information on the operations, programme logic, context, successes, and challenges of each intervention and validate factual correctness of any potential misinformed responses provided by any participant of the stakeholder interviews. The participants of the focus group discussion were asked open-ended questions about the design, programme strategy, implementation, outcomes, drivers, barriers, lessons learned, strengths and weaknesses of their project.

The participants of the focus group discussions also took part in the stakeholder interviews and completed the ranking questionnaire. However, the focus group discussions were conducted after completing all stakeholder interviews and screening
the responses and thus provided an opportunity to clarify and correct any widely varied perceptions or conflicting information from other participants of the interviews.

7.3.5 Questionnaire

Following the integrative framework combining the balanced scorecard and intervention logic model, the researcher developed a ranking questionnaire with a total of 22 questions to score and compare different aspects of the selected projects in terms of their compliance with the Paris principles, economy, efficiency, effectiveness, and sustainability. Table 7.2 provides the list of aspects reviewed for the projects whereas Annex 1 provides the full list of criteria used for ranking the aspects of the projects.

Table 7.2: Balanced scorecard domains and indicators used in the questionnaire

<table>
<thead>
<tr>
<th>Domains</th>
<th>Indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paris principles</td>
<td>1. Degree of ownership by recipient country</td>
</tr>
<tr>
<td>Paris principles</td>
<td>2. Alignment: degree to which there is a serious effort to connect aid programmes with country policies and processes</td>
</tr>
<tr>
<td>Paris principles</td>
<td>3. Degree of alignment of donor technical cooperation with the partner country’s capacity development objectives and strategies</td>
</tr>
<tr>
<td>Paris principles</td>
<td>4. Alignment: Avoid a dedicated management units designed to support development projects or programmes</td>
</tr>
<tr>
<td>Paris principles</td>
<td>5. Alignment: Use of country’s financial management, procurement and monitoring and evaluation systems</td>
</tr>
<tr>
<td>Paris principles</td>
<td>6. Harmonization: Takes part in joint mission and joint country analytic work</td>
</tr>
<tr>
<td>Paris principles</td>
<td>7. Managing for results: Extent of using a results-based M&amp;E system that, building on sound statistical data and open access to information, produces data on progress toward desired inputs, outputs, and outcomes of the project</td>
</tr>
<tr>
<td>Paris principles</td>
<td>8. Mutual accountability: undertake mutual assessments of progress in implementing agreed commitments</td>
</tr>
<tr>
<td>Economy</td>
<td>9. Economy: Unit costs</td>
</tr>
<tr>
<td>Economy</td>
<td>10. Economy: Management and transaction costs</td>
</tr>
<tr>
<td>Economy</td>
<td>11. Economy: Procurement</td>
</tr>
<tr>
<td>Efficiency</td>
<td>12. Project outputs: Percentage of achievement of different project outputs against their targets</td>
</tr>
<tr>
<td>Efficiency</td>
<td>13. Health system outputs: Percentage of achievement of health system outputs (related to service delivery, workforce, information, supplies and commodities, financing, and governance) against the targets established through participatory impact pathway analysis</td>
</tr>
<tr>
<td>Efficiency</td>
<td>14. Efficiency: Productivity measure</td>
</tr>
<tr>
<td>Efficiency</td>
<td>15. Efficiency: Business lead time—Regular activities</td>
</tr>
<tr>
<td>Efficiency</td>
<td>16. Efficiency: Business lead time—Exceptional activities</td>
</tr>
<tr>
<td>Effectiveness</td>
<td>17. Project outcome indicators: Percentage of achievement of strategic objectives of the intervention against the target</td>
</tr>
<tr>
<td>Effectiveness</td>
<td>18. Health system outcome: Extent of improvement of related health service delivery system (access, coverage, quality and equity); Extent of improvement of health system’s responsiveness on the related issue; Extent of improved efficiency of the health system on the related issue</td>
</tr>
<tr>
<td>Effectiveness</td>
<td>19. Effectiveness: Relevance and Robustness (Relevance= clear, planned, causally linked, and system-wide. Robust=data to support indicators are available, accessible, credible, and disaggregate-able)</td>
</tr>
<tr>
<td>Effectiveness</td>
<td>20. Effectiveness: Theory of Change</td>
</tr>
<tr>
<td>Sustainability</td>
<td>21. Sustainability: Leverage/Replication</td>
</tr>
<tr>
<td>Sustainability</td>
<td>22. Sustainability: Will the outputs delivered through the project be sustained by national ownership and capacities, after the end of the project duration</td>
</tr>
</tbody>
</table>
Each indicator of the structured questionnaire had defined ranking criteria on a scale of 1 to 5 drawn from the possible best practices for an indicator. Each rank represented the number of steps away from the best practice. That is, the rank 5 represented the highest score meaning the project was fully comparable to the best practice on that aspect, whereas the rank 1 represented the lowest score meaning the project was at the very elementary stage and would take the maximum number of steps to reach to the best practice on that aspect.

The researcher provided the questionnaire in advance to the respondents and explained the questions and ranks in a semi-structured interview process with the respondents. The respondents then indicated what rank they would give to a project against a specific indicator and ranking criteria. In addition to the ranking, some additional open ended probing questions were asked to the participants to confirm their views, clarify any confusion, and get their opinions on any aspect that seemed contradictory.

7.3.6 Importance weighting

In order to determine the relative importance of each item of the questionnaire and evaluate the scores obtained by each project, three senior policy makers from the MoH of Timor-Leste were asked to rank the importance of each item on a scale of 1 to 10 where 1 represented the lowest importance and 10 the highest importance. These three senior officials were selected from the study participants with the criteria of holding decision making authority on behalf of MoH. They first completed the 22 ranking questions on the scale of 1-5 during the regular semi-structured interview. But later they were asked individually to rank the importance of each item of the questionnaire.

While the stakeholder interview questionnaire had defined ranking criteria for each item in a scale of 1 to 5, the importance weighting used a wider and generalized score of 1 to 10 to be able to accommodate more subtle variation of the comparative importance of each item.

7.3.7 Data processing and analysis

As suggested by Jick (1979) and Creswell and Miller (2000), the researcher listened to each recorded interview and transcribed them in the form of written notes of opinions
and justifications of the participant’s ranking scores against each indicator. The researcher then cross-checked the scores obtained from a participants’ response against the interview transcription notes, additional notes taken by the researcher during the interview process, and notes from the focus group discussion to take note of any confusion or seeming factual error.

The scores from the stakeholder interviews and importance weighting were then tabulated and transferred to a spread-sheet for descriptive statistical analysis including standard deviation, weighted scores, total scores, and average scores on each item and group of items representing an aspect of the interventions.

It needs to be noted that scoring by the respondents was based on ordinal scales with possible inequal interval between the ranks. However, for the sake of simplicity of calculation and in order to derive an indicative collective view of each aspect of the projects, analysis was done based on average scores, importance weights, and weighted scores.

While the average value of the ranking of an item by the respondents gave an indication of the status of that indicator, for a holistic analysis this value was multiplied by the average importance weight for that item to obtain the weighted scores. The weighted scores were then amenable to both quantitative and qualitative analysis with statistical comparison of an average score of an intervention on a particular aspect or on a domain comprising several aspects and then looking for validity and possible explanations.

**7.3.8 Presentation and interpretation of scores**

The ‘Balanced scorecards’ method was used to present and interpret the findings from the data analysis to compare the performance of three interventions on five domains: the degree of adherence to the Paris principles, economy, efficiency, effectiveness, and likely sustainability. The balanced scorecards drawn for the purpose of this research consisted of these five domains and side-by-side presented the scores of the three interventions on each domain. In order to compare the weighted scores on a particular aspect and on an equal scale for comparing aspects, the weighted scores under the Paris principles domain were aggregated and converted to a total score of 100 with 20 points
for each aspect of the Paris principles: country ownership, alignment, harmonization, managing for results, and mutual accountability.

Similarly, weighted scores under economy, efficiency, effectiveness, and sustainability domains were aggregated and converted into a standard score of 100 for each domain. Thus, the balanced scorecard compared the design and performance of three interventions in a total score of 500 with a maximum 100 score for each domain.

**7.4 Results**

**7.4.1 Summary of overall response and importance weight**

Table 7.3 presents the number of responses received, the average score (out of maximum 5 points), and standard deviation against each item of the questionnaire. Standard deviations greater than 1 are highlighted to take note of wide variation of opinions.

**Table 7.3: Summary of stakeholder responses**

<table>
<thead>
<tr>
<th>Indicator</th>
<th>HSSP-SP</th>
<th>NAP</th>
<th>IPL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. of resp.</td>
<td>Ave. score</td>
<td>Std dev.</td>
</tr>
<tr>
<td>1  Degree of ownership by recipient country</td>
<td>13</td>
<td>2.88</td>
<td>0.845</td>
</tr>
<tr>
<td>2  Alignment: degree to which there is a serious effort to connect aid programmes with country policies and processes</td>
<td>13</td>
<td>3.38</td>
<td>0.740</td>
</tr>
<tr>
<td>3  Degree of alignment of donor technical cooperation with the partner country’s capacity development objectives and strategies</td>
<td>13</td>
<td>3.46</td>
<td>0.749</td>
</tr>
<tr>
<td>4  Alignment: Avoid a dedicated management units designed to support aid programme</td>
<td>13</td>
<td>2.5</td>
<td>0.913</td>
</tr>
<tr>
<td>5  Alignment: Use of country’s financial management, procurement and monitoring and evaluation systems</td>
<td>13</td>
<td>2.5</td>
<td>0.890</td>
</tr>
<tr>
<td>6  Harmonization: Take part in joint mission and joint country analytic work</td>
<td>13</td>
<td>3.46</td>
<td><strong>1.198</strong></td>
</tr>
<tr>
<td>7  Managing for results: Extent of using a results-based monitoring and evaluation (M&amp;E) system</td>
<td>13</td>
<td>2.62</td>
<td><strong>1.193</strong></td>
</tr>
<tr>
<td>8  Mutual accountability: undertake mutual assessments of progress in implementing agreed commitments</td>
<td>13</td>
<td>3.73</td>
<td>0.439</td>
</tr>
<tr>
<td>9  Economy: Unit costs</td>
<td>13</td>
<td>2.54</td>
<td>0.660</td>
</tr>
<tr>
<td>Indicator</td>
<td>HSSP-SP</td>
<td>NAP</td>
<td>IPL</td>
</tr>
<tr>
<td>-------------------------------------------------------------------------</td>
<td>---------</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td></td>
<td>No. of resp.</td>
<td>Ave. score</td>
<td>Std dev.</td>
</tr>
<tr>
<td>10 Economy: Management and transaction costs</td>
<td>13</td>
<td>2.69</td>
<td>0.630</td>
</tr>
<tr>
<td>11 Economy: Procurement</td>
<td>13</td>
<td>3.19</td>
<td>0.663</td>
</tr>
<tr>
<td>12 Project outputs: Percentage of achievement of different project</td>
<td>13</td>
<td>2.77</td>
<td>0.904</td>
</tr>
<tr>
<td>outputs against their targets</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13 Health system outputs: Percentage of achievement of health system</td>
<td>13</td>
<td>2.92</td>
<td>0.732</td>
</tr>
<tr>
<td>outputs (related to service delivery, workforce, information, supplies</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>and commodities, financing, and governance) against the targets</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>established through participatory impact pathway analysis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14 Efficiency: Productivity measure</td>
<td>13</td>
<td>2.92</td>
<td>0.641</td>
</tr>
<tr>
<td>15 Efficiency: Business lead time—Regular activities</td>
<td>13</td>
<td>2.60</td>
<td>0.733</td>
</tr>
<tr>
<td>16 Efficiency: Business lead time—Exceptional activities</td>
<td>13</td>
<td>2.58</td>
<td>0.932</td>
</tr>
<tr>
<td>17 Project outcome indicators: Percentage of achievement of strategic</td>
<td>13</td>
<td>2.58</td>
<td>0.954</td>
</tr>
<tr>
<td>objectives of the intervention against the targets</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18 Health system outcome: Extent of improvement of related health service delivery system (access, coverage, quality and equity); Extent of improvement of health system’s responsiveness on the related issue; Extent of improved efficiency of the health system on the related issue</td>
<td>13</td>
<td>3.08</td>
<td>0.760</td>
</tr>
<tr>
<td>19 Effectiveness: Relevance and Robustness (Relevance = clear, planned, causally linked, and system-wide. Robust=data to support indicators are available, accessible, credible, and disaggregate-able)</td>
<td>13</td>
<td>3.5</td>
<td>0.764</td>
</tr>
<tr>
<td>20 Effectiveness: Theory of Change</td>
<td>13</td>
<td>3.15</td>
<td>0.875</td>
</tr>
<tr>
<td>21 Sustainability: Leverage/Replication</td>
<td>13</td>
<td>3.27</td>
<td>0.904</td>
</tr>
<tr>
<td>22 Sustainability: Will the outputs delivered through the project be</td>
<td>13</td>
<td>2.52</td>
<td>1.120</td>
</tr>
<tr>
<td>sustained by national ownership and capacities, after the end of the project duration</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

It can be noted that the response rate for most of the questions related to the NAP was 68.18%, as a total 15 respondent out of 22 responded to the NAP related questions. This response rate was 59.09% for HSSP-SP (13 responses out 22 respondents) and 50% for IPL (maximum 11 responses out of 22 respondents) indicating that more respondents were knowledgeable on the NAP than the HSSP-SP and IPL.
With standard deviation of less than 1, it seems most of the respondents were pretty close to each other in their opinion about the different aspects of the three projects. However, responses varied on a few items with standard deviations of greater than 1. These items include the issues of HSSP-SP taking part in joint mission and country analytic work; using a result-based monitoring systems, and whether outputs from the HSSP-SP would be sustainable beyond the funding period. For the NAP, opinion varied on the issues of the Global Fund funded projects aligning to the country’s financial management, procurement, and monitoring and evaluation systems; whether the Global Fund took part in any joint mission and country analytic work; whether the NAP was replicable and leveraged additional support; and whether outputs from the NAP would be sustainable beyond the funding period. For IPL opinions varied on the questions if IPL used a result-based monitoring system, and productivity measure to achieve greater efficiency.

As mentioned earlier the average scores obtained from the responses were analyzed together with the importance weightage for each questions on a scale of 1 to 10. Table 7.4 presents the importance weightage calculated from individual assessment by three senior policy makers from the MoH:

<p>| Table 7.4: Importance weight of evaluation indicators (1= least important, 10= most important) |
|--------------------------------------------------|------------------|------------------|-------------------------------|------------------|</p>
<table>
<thead>
<tr>
<th>Question</th>
<th>Indicator</th>
<th>Res.1</th>
<th>Res.2</th>
<th>Res.3</th>
<th>Average weight</th>
<th>Adjusted weight</th>
<th>Std. dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Degree of ownership by recipient country</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>Alignment: degree to which there is a serious effort to connect aid programmes with country policies and processes</td>
<td>10</td>
<td>10</td>
<td>8</td>
<td>9.33</td>
<td>10</td>
<td>1.155</td>
</tr>
<tr>
<td>3</td>
<td>Degree of alignment of donor technical cooperation with the partner country's capacity development objectives and strategies</td>
<td>7</td>
<td>10</td>
<td>8</td>
<td>8.33</td>
<td>8</td>
<td>1.528</td>
</tr>
<tr>
<td>4</td>
<td>Alignment: Avoid a dedicated management units designed to support aid programme</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>1.732</td>
</tr>
<tr>
<td>5</td>
<td>Alignment: Use of country's financial management, procurement and monitoring and evaluation systems</td>
<td>6</td>
<td>5</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>Harmonization: Take part in joint mission and joint country analytic work</td>
<td>7.5</td>
<td>10</td>
<td>10</td>
<td>9.17</td>
<td>9</td>
<td>1.443</td>
</tr>
<tr>
<td>7</td>
<td>Managing for results: Extent of using a results-based monitoring and evaluation (M&amp;E) system</td>
<td>10</td>
<td>10</td>
<td>8</td>
<td>9.33</td>
<td>10</td>
<td>1.155</td>
</tr>
<tr>
<td>8</td>
<td>Mutual accountability: undertake mutual assessments of progress in implementing agreed commitments</td>
<td>10</td>
<td>8</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>Economy: Unit costs</td>
<td>10</td>
<td>10</td>
<td>8</td>
<td>9.33</td>
<td>10</td>
<td>1.155</td>
</tr>
<tr>
<td>Question</td>
<td>Indicator</td>
<td>Res.1</td>
<td>Res. 2</td>
<td>Res. 3</td>
<td>Average weight</td>
<td>Adjusted weight</td>
<td>Std. dev.</td>
</tr>
<tr>
<td>----------</td>
<td>---------------------------------------------------------------------------</td>
<td>-------</td>
<td>--------</td>
<td>--------</td>
<td>----------------</td>
<td>-----------------</td>
<td>-----------</td>
</tr>
<tr>
<td>10</td>
<td>Economy: Management and transaction costs</td>
<td>5</td>
<td>5</td>
<td>7.5</td>
<td>5.83</td>
<td>6</td>
<td>1.443</td>
</tr>
<tr>
<td>11</td>
<td>Economy: Procurement</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>9.67</td>
<td>10</td>
<td>0.577</td>
</tr>
<tr>
<td>12</td>
<td>Project outputs: Percentage of achievement of different project outputs against their targets</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>Health system outputs: Percentage of achievement of health system outputs (related to service delivery, workforce, information, supplies and commodities, financing, and governance) against the targets established through participatory impact pathway analysis</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>9.67</td>
<td>10</td>
<td>0.577</td>
</tr>
<tr>
<td>14</td>
<td>Efficiency: Productivity measure</td>
<td>6</td>
<td>10</td>
<td>10</td>
<td>8.67</td>
<td>9</td>
<td>2.309</td>
</tr>
<tr>
<td>15</td>
<td>Efficiency: Business lead time—Regular activities</td>
<td>8</td>
<td>8</td>
<td>8.5</td>
<td>8.17</td>
<td>8</td>
<td>0.289</td>
</tr>
<tr>
<td>16</td>
<td>Efficiency: Business lead time—Exceptional activities</td>
<td>8</td>
<td>2</td>
<td>9</td>
<td>6.33</td>
<td>8</td>
<td>3.786</td>
</tr>
<tr>
<td>17</td>
<td>Project outcome indicators: Percentage of achievement of strategic objectives of the intervention against the targets</td>
<td>7</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>1.732</td>
</tr>
<tr>
<td>18</td>
<td>Health system outcome: Extent of improvement of related health service delivery system (access, coverage, quality and equity); Extent of improvement of health system’s responsiveness on the related issue; Extent of improved efficiency of the health system on the related issue</td>
<td>8</td>
<td>8</td>
<td>9</td>
<td>8.33</td>
<td>8</td>
<td>0.577</td>
</tr>
<tr>
<td>19</td>
<td>Effectiveness: Relevance and Robustness (Relevance= clear, planned, causally linked, and system-wide. Robust=data to support indicators are available, accessible, credible, and disaggregate-able)</td>
<td>8</td>
<td>9</td>
<td>8</td>
<td>8.33</td>
<td>8</td>
<td>0.577</td>
</tr>
<tr>
<td>20</td>
<td>Effectiveness: Theory of Change</td>
<td>9</td>
<td>5</td>
<td>10</td>
<td>8</td>
<td>8</td>
<td>2.65</td>
</tr>
<tr>
<td>21</td>
<td>Sustainability: Leverage/Replication</td>
<td>8</td>
<td>9</td>
<td>8</td>
<td>8.33</td>
<td>8</td>
<td>0.577</td>
</tr>
<tr>
<td>22</td>
<td>Sustainability: Will the outputs delivered through the project be sustained by national ownership and capacities, after the end of the project duration</td>
<td>10</td>
<td>10</td>
<td>9.5</td>
<td>9.83</td>
<td>10</td>
<td>0.289</td>
</tr>
</tbody>
</table>

As it can be seen in the above table, except for Question16, the average importance weight calculated from the scores given by three policy makers were rounded to derive the adjusted importance weight. However, for Question 16 (business lead time taken by an intervention for getting necessary approvals and implementing an exceptional activity), while two policy makers thought this item is very important giving a score of 8 and 9 out of 10 for this item, the other policy maker thought this item is not very important for Timor-Leste and scored only 2 for this item. Due to the high variability and outlier effect of this opinion (standard deviation 3.78), the adjusted weight for this item considered the two closely matched opinions.
The following sections present further analysis and results related to the performance of each project on the five domains of the balanced scorecards: Compliance with the Paris principles, Economy, Efficiency, Effectiveness and Sustainability.

### 7.4.2 Compliance with the Paris principles

In order to assess to what extent the selected projects’ design and policies were complying with the Paris principles of aid effectiveness, respondents were asked 8 questions related to the degree of each project’s compliance with the country ownership, alignment to country systems, harmonization with other funding sources, managing for results, and mutual accountability. The following table provides a summary of weighted score received by each project on different aspects of the Paris principles of aid effectiveness:

#### Table 7.5: Weighted scores on compliance with the Paris principles

<table>
<thead>
<tr>
<th>Question</th>
<th>Category</th>
<th>Importance weight</th>
<th>Maximum weighted Score</th>
<th>Weighted Score HSSP-SP</th>
<th>Weighted Score NAP</th>
<th>Weighted Score IPL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ownership</td>
<td>10</td>
<td>50</td>
<td>28.85</td>
<td>38.83</td>
<td>30.00</td>
</tr>
<tr>
<td>2</td>
<td>Alignment</td>
<td>10</td>
<td>50</td>
<td>33.85</td>
<td>38.00</td>
<td>33.18</td>
</tr>
<tr>
<td>3</td>
<td>Alignment</td>
<td>8</td>
<td>40</td>
<td>27.69</td>
<td>28.53</td>
<td>25.45</td>
</tr>
<tr>
<td>4</td>
<td>Alignment</td>
<td>3</td>
<td>15</td>
<td>7.50</td>
<td>9.20</td>
<td>5.59</td>
</tr>
<tr>
<td>5</td>
<td>Alignment</td>
<td>6</td>
<td>30</td>
<td>15.00</td>
<td>18.40</td>
<td>9.82</td>
</tr>
<tr>
<td>6</td>
<td>Harmonization</td>
<td>9</td>
<td>45</td>
<td>31.15</td>
<td>31.20</td>
<td>36.20</td>
</tr>
<tr>
<td>7</td>
<td>Managing for results</td>
<td>10</td>
<td>50</td>
<td>26.15</td>
<td>39.00</td>
<td>36.36</td>
</tr>
<tr>
<td>8</td>
<td>Mutual accountability</td>
<td>9</td>
<td>45</td>
<td>33.58</td>
<td>37.50</td>
<td>25.98</td>
</tr>
<tr>
<td>Total</td>
<td>Paris principles</td>
<td></td>
<td>325</td>
<td>203.77</td>
<td>240.67</td>
<td>202.59</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Converted score out of 100</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100</td>
</tr>
</tbody>
</table>

|                             | 62.70          |
|                             | 74.05          |
|                             | 62.34          |

**Ownership:**

Although the SWAp through the HSSP-SP was meant to go through the government system, it was found that the feeling of ownership was highest (3.88 out of 5) with the NAP and lowest with the HSSP-SP (2.88 out of 5). This was because the Global Fund’s proposal was developed in country without the Global Fund’s involvement. It was the in-country stakeholders comprising the MoH, NGOs, development partners and people
living with diseases that planned the activities, developed the budget and submitted the proposal to the Global Fund, creating a very strong country ownership feeling. As one respondent commented,

"For Global Fund [programme development] although we hire external consultants for proposal writing, actual process of project design always take place in country and based on country strategy. But for HSSP-SP and IPL the design was done based on donor assessment" (a respondent from MoH).

For HSSP-SP, the project design was done by a Design Mission comprising of consultants from the World Bank and AusAID and involving the MoH. However, the MoH officials felt that they were not in control in designing HSSP-SP, rather it was mostly done by the consultants. One of the respondents said,

"For Global Fund [funded programme] we designed everything in the country and Global Fund agreed to fund them. But for HSSP-SP everything was done by the World Bank and imposed on us. Even it’s a pain to obtain no-objection letter from the World Bank if we want to vary from what was designed by them" (a senior manager from the MoH).

Similarly, for the USAID funded IPL, the project design was done by USAID and JSI. However, the MoH officials seemed happy with the way IPL team worked, as IPL worked closely together with the MoH and involved MoH in planning and organizing activities and making decisions. As another respondent commented,

"The EPI [Expanded Programme on Immunization] Working Group belongs to the Ministry of Health and for [IPL] project activities related decision making, we always involve the Ministry of Health" (a implementation team member from IPL).

As per the importance score, country ownership was thought as one of the most important aspects of the project with all three respondents assigning the highest importance weight (10 out of 10) to this item.

**Alignment:**
The alignment aspect of the projects was assessed from four different perspectives. They were the degree of serious efforts by the project to align with country’s policies and processes; aligning the project’s objectives with country’s development objectives and strategies; avoiding a dedicated project implementation unit; and using country’s procurement, financial management and monitoring and evaluation systems. It was found that the country assigned most importance (10 out of 10) with the principle of aligning development effort with the country's policies and processes and all three projects scored slightly higher than the average score in a scale of 1 to 5; the NAP scored
higher (3.80) than others. This was because the NAP followed the National HIV/AIDS Strategic Plan and all government policies in implementing the programme. Although HSSP-SP aimed to support the National Health Sector Strategic Plan, the project imposed conditions of following World Bank procedures in procurement. IPL, on the other hand, followed USAID procedure and guidelines. Commenting on HSSP-SP one respondent pointed out,

“World Bank’s procurement process is complex. It depends on the skills of the procurement staff and advisers, how they understand World Bank’s procurement policies, and the relationship between the Ministry of Health and World Bank in dealing with the procurement issues. I think it was a difficult relationship” (a respondent working with the donors).

For all three projects there were efforts to address the country’s capacity development gaps, but they were not quite aligned to the country’s capacity development objectives and strategies as outlined in the National Health Workforce Development Plan of Timor-Leste (Ministry of Health of Timor-Leste, 2005).

As mentioned in the previous chapter, all three projects used a dedicated project implementation unit. However, for the NAP, even if a Division of Global Fund within the Department of Partnership Management of the MoH facilitated grant management, implementation of most of the programmatic activities were led and managed by the NAP giving it a higher score (3.07) than HSSP-SP (2.50) and IPL (1.86) on this aspect. For the importance, the scorers did not find this aspect very important (3 out of 10), as for them ownership and implementation of programmes were more important than how implementation was organized.

For the use of country’s procurement, financial management and monitoring and evaluation (M&E) systems, all three projects had separate bank accounts and separate book keeping for project funded activities. However, for procurement the NAP used the country’s systems whereas the World Bank put conditions on HSSP-SP to use World Bank’s procurement policies and procedures and IPL had to follow the USAID policies and procedures for procurement.

For monitoring and evaluation, the NAP used a vertical stand-alone monitoring system adopted and owned by the MoH for the NAP. HSSP-SP tried to use the regular Health Management Information System (HMIS) of the MoH, which had been often burdened
and modified by donor requirements. IPL, on the other hand, used their own M&E system but relied on the HMIS for coverage indicators. The NAP received the highest score (3.07) and IPL the lowest score (1.64) on this indicator. The importance value for this indicator was considered 6 out of 10.

**Harmonization:**

The scorers assigned 9 out of 10 for the importance of harmonization process for the donors to harmonize their efforts with other funding sources and to avoid duplication and competition. In this aspect, USAID funded IPL scored higher (4.02) than the NAP (3.47) and HSSP-SP (3.46). It was acknowledged that IPL took part more in joint mission and joint country analytical work with government and other donors than the NAP or HSSP-SP did. As one respondent pointed out,

“For HSSP-SP the intent [to harmonize] was there but it was the implementation which was difficult. The ownership [of the process of harmonization] by the government was limited” (a respondent working with the donors).

**Managing for results:**

Scorers thought managing for results was one of the most important (10 out of 10) aspects of project implementation. HSSP-SP scored lowest (2.62) in this aspect, as the project did not have any agreed results framework for the first couple of years. Even the results framework the project developed later focused on general indicators and depended on results from other programmes not directly funded by HSSP-SP making it difficult to establish a results-based monitoring system. In fact, the results framework of HSSP-SP was later modified in 2012 to focus on more manageable process indicators.

The Global Fund, on the other hand, strictly implemented a results-based monitoring and funding system. The performance framework of the project was developed during the grant agreement signing. Tied with the fund disbursement request, the principal recipient (the MoH) had to report to the Global Fund and Country Coordinating Mechanism against the agreed performance framework every quarter, which was then externally verified by a Local Fund Agent appointed by the Global Fund. Thus, the NAP received a score of 3.90 out of 5 on this aspect. IPL also had an agreed performance framework and strong reporting system and received a score of 3.64 out of 5 on this
aspect. However, regarding the Global Fund’s results management process one respondent commented,

“The process [of managing for results by the Global Fund] is extremely complex. There could be a little modification from the Global Fund in terms of expectations from a country with low capacity. Timor-Leste is not Indonesia or Cambodia. You cannot expect the same [results] from Timor-Leste. For CCM [the country coordinating mechanism], I think there is very little ownership by the government” (a respondent working with the donors).

Mutual accountability:

Scorers assigned 9 out of 10 importance value for mutual accountability. Mutual accountability was measured in terms of the extent of undertaking mutual assessment of progress of the project by the donor and aid recipient. The World Bank conducted a number of joint assessments such as annual reviews and mid-term reviews with the MoH for HSSP-SP and obtained an average score of 3.73 out of 5. With the Local Fund Agent verifying progress updates and the Global Fund providing regular feedback on progress and performance in the form of Management Letters tied with fund disbursements, Global Fund seemed strongest on this aspect scoring 4.17 out of 5. However, IPL’s mutual accountability mechanism was questioned as they reported directly to JSI and USAID with hardly any joint assessment taking place between the MoH and IPL on the progress updates and use of funds. IPL scored 2.89 out of 5 on this aspect.

Converted scores on the Paris principles:

In order to compare the degree of each project’s compliance with the five principles of the Paris Declaration on an equal scale, all weighted scores for a specific aspect of a project (e.g. for Alignment questions) were aggregated. The aggregated scores were then converted to a standard scale of 20 for each aspect of the Paris principles: Ownership, Alignment, Harmonization, Managing results, and Mutual accountability. The results are presented in the following table:

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Weighted score: HSSP-SP</th>
<th>Weighted score: NAP</th>
<th>Weighted score: IPL</th>
<th>Weighted score: scale</th>
<th>Converted score: HSSP-SP</th>
<th>Converted score: NAP</th>
<th>Converted score: IPL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ownership</td>
<td>28.85</td>
<td>38.83</td>
<td>30.00</td>
<td>50</td>
<td>20</td>
<td>11.54</td>
<td>15.53</td>
</tr>
<tr>
<td>Alignment</td>
<td>84.04</td>
<td>94.13</td>
<td>74.05</td>
<td>135</td>
<td>20</td>
<td>12.45</td>
<td>13.95</td>
</tr>
<tr>
<td>Harmonization</td>
<td>31.15</td>
<td>31.20</td>
<td>36.20</td>
<td>45</td>
<td>20</td>
<td>13.85</td>
<td>13.87</td>
</tr>
<tr>
<td>Results</td>
<td>26.15</td>
<td>39.00</td>
<td>36.36</td>
<td>50</td>
<td>20</td>
<td>10.46</td>
<td>15.60</td>
</tr>
<tr>
<td>Accountability</td>
<td>33.58</td>
<td>37.50</td>
<td>25.98</td>
<td>45</td>
<td>20</td>
<td>14.92</td>
<td>16.67</td>
</tr>
</tbody>
</table>
The following graph compares results from each project’s compliance with the Paris principles:

![Graph comparing results from each project's compliance with the Paris principles](image)

**Figure 7.1: Compliance with the Paris principles by three projects**

It can be seen that the NAP had a high sense of country ownership, rigorous accountability mechanisms, and good managing for results. The NAP tried to align with the country’s systems or build country systems where existing systems were not adequate to meet international standards. However, the effort of the NAP to harmonize with other funding sources was relatively weak. This was probably because the Global Fund provided the bulk amount of funding to cover almost all activities related to the NAP creating the situation of no further need for additional funding from other sources.

For IPL, on the other hand, the sense of country ownership was weak, mutual accountability mechanisms were not very clear, and the project’s systems were poorly aligned to the country’s procurement, financial management and M&E systems.
However, IPL had good systems for managing for results and harmonized well with other funding sources including the government.

Compared to the NAP and IPL, the HSSP-SP did not align their systems well to the country systems or harmonize with other funding sources. Although mutual accountability mechanisms for HSSP-SP were relatively strong, its greatest weakness was the less comprehensive efforts for managing for results than the other results based funding mechanisms.

As can be seen from Table 7.5, overall the NAP scored 74.05, HSSP-SP scored 62.70, and IPL scored 62.34 on compliance with the Paris principles of aid effectiveness as per the total weighted scores out of 100.

**7.4.3 Economy**

Questions 9-11 of the questionnaire attempted to evaluate the project activities in terms of the economy for organizing them. This was measured in terms of the ratio between the resources used for organizing the activities and the cost of resources. The following table shows the weighted scores of the three projects according to their average scores multiplied by their corresponding importance weights.

**Table 7.7: Weighted scores on economy of three interventions**

<table>
<thead>
<tr>
<th>Question</th>
<th>Category</th>
<th>Importance weight</th>
<th>Maximum weighted Score</th>
<th>Weighted Score HSSP-SP</th>
<th>Weighted Score NAP</th>
<th>Weighted Score IPL</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>Economy: Unit costs</td>
<td>10</td>
<td>50</td>
<td>25.38</td>
<td>33.57</td>
<td>33.50</td>
</tr>
<tr>
<td>10</td>
<td>Economy: Management and</td>
<td>6</td>
<td>30</td>
<td>16.15</td>
<td>20.80</td>
<td>18.00</td>
</tr>
<tr>
<td></td>
<td>transaction costs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Economy: Procurement</td>
<td>10</td>
<td>50</td>
<td>31.92</td>
<td>33.33</td>
<td>36.25</td>
</tr>
<tr>
<td>Total</td>
<td>Economy</td>
<td></td>
<td>130</td>
<td>73.46</td>
<td>87.70</td>
<td>87.75</td>
</tr>
<tr>
<td></td>
<td>Converted score out of 100</td>
<td></td>
<td>100</td>
<td>56.51</td>
<td>67.47</td>
<td>67.50</td>
</tr>
</tbody>
</table>

The economy scores tried to measure the economy of inputs used for the projects. It was measured in terms of unit costs of inputs such as human resources and technical assistance, management and transaction costs, and procurement efforts for cost reduction and cost management. The scorers considered unit costs and procurement as
most important and assigned 10 out of 10 for them for the importance ranking whereas they assigned 6 out of 10 for the importance of management and transaction costs.

**Unit costs:**
It can be seen that the unit costs of HSSP-SP (average score 2.54 out of 5) were higher than those of the NAP (average score 3.36 out of 5) and IPL (average score 3.35 out of 5). This is because HSSP-SP used highly expensive technical assistance and specialized project implementation staff compared to other projects.

**Management and transaction costs:**
The management and transaction cost was measured as a ratio between programme activity costs and human resources and administrative costs in organizing the activities. For the ranking score, participants were asked to compare the management and transaction cost of each project with benchmarked costs of other comparable projects. It was found that the average management and transaction cost of the NAP (average score 3.47 out of 5) was lower than IPL (average score 3 out of 5) and HSSP-SP (average score 2.69 out of 5).

**Procurement efforts:**
All three projects demonstrated procurement efforts to reduce and manage costs with IPL scoring slightly higher (average score 3.63 out of 5) than the NAP (average score 3.33 out of 5) and HSSP-SP (average score 3.19 out of 5).

However, these scores were based on the impressions of the respondents. In an attempt to further verify the validity of these scores, Chapter 10 of this dissertation analyzes the economy of actual unit costs, management and transaction costs and procurement unit costs of these three projects in comparison to each other.

Figure 7.3 compares the average scores obtained by the three projects on the aspects of unit costs, management and transaction costs, and procurement efforts:
Figure 7.2: Perceived economy of three interventions

Figure 7.3 shows that the economy of HSSP-SP was perceived as lower than the economy of the NAP and IPL with the NAP scoring higher on the aspect of management and transaction costs than IPL and HSSP-SP. As can be seen in Table 7.7, overall IPL scored 67.50, the NAP scored 67.47, and HSSP-SP scored 56.51 on the aspects of economy as per their total weighted scores out of 100.

7.4.4 Efficiency

Questions 12-16 of the research questionnaire (Table 7.3) attempted to evaluate the production of outputs of the projects against their performance targets and the project’s efficiency in producing them. The efficiency was measured in terms of the ratio between the outputs produced and inputs used for the project activities. It also measured the operational efficiency of the projects in terms of their business lead time for regular and exceptional activities. Table 7.8 summarizes the weighted scores of the efficiency measure of three projects:
Table 7.8: Perceived efficiency of three interventions

<table>
<thead>
<tr>
<th>Question</th>
<th>Category</th>
<th>Importance weight</th>
<th>Maximum weighted Score</th>
<th>Weighted Score HSSP-SP</th>
<th>Weighted Score NAP</th>
<th>Weighted Score IPL</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>Project Outputs</td>
<td>10</td>
<td>50</td>
<td>27.69</td>
<td>36.67</td>
<td>35.45</td>
</tr>
<tr>
<td>13</td>
<td>Health system Outputs</td>
<td>10</td>
<td>50</td>
<td>29.23</td>
<td>35.83</td>
<td>31.82</td>
</tr>
<tr>
<td>14</td>
<td>Productivity measure</td>
<td>9</td>
<td>45</td>
<td>26.31</td>
<td>34.50</td>
<td>32.32</td>
</tr>
<tr>
<td>15</td>
<td>Business lead time for regular activities</td>
<td>8</td>
<td>40</td>
<td>20.77</td>
<td>26.40</td>
<td>30.18</td>
</tr>
<tr>
<td>16</td>
<td>Business lead time for exceptional activities</td>
<td>8</td>
<td>40</td>
<td>20.62</td>
<td>22.93</td>
<td>22.00</td>
</tr>
<tr>
<td>Total</td>
<td>Efficiency</td>
<td></td>
<td></td>
<td>125</td>
<td>156.33</td>
<td>151.77</td>
</tr>
<tr>
<td></td>
<td>Converted score out of 100</td>
<td></td>
<td></td>
<td>100</td>
<td>55.38</td>
<td>69.48</td>
</tr>
</tbody>
</table>

**Project outputs:**
From January 2012 to December 2013, the NAP achieved most of its performance targets related to the project outputs and ended with a “B1” (Adequate performance) performance rating given by the Global Fund (2014). IPL did not put any output targets on their performance framework. However, the rate of achievement of project outputs against their work plan for IPL was satisfactory against their baseline values. Thus the NAP and IPL received an average score of 3.67 and 3.55 respectively for the rate of achievement of project outputs against their targets (Table 7.3).

HSSP-SP, on the other hand, scored only 2.77 out of 5 on this indicator indicating a slow production of outputs by this project. This is consistent with the “Unsatisfactory” rating given by the World Bank for the overall implementation progress of HSSP-SP until June 2013 (World Bank, 2013k).

The scorers considered achievement of project outputs as one of the most important indicators assigning 10 out of 10 to this indicator.

**Health systems outputs:**
As analyzed in the previous chapter, all three projects had some activities and planned outputs related to strengthening health systems for service delivery, workforce
development, information, supplies and commodities, financing, and governance and leadership. Achievement of health systems strengthening outputs of the NAP was considered higher (score 3.58 out of 5) than that of IPL (score 3.18) and HSSP-SP (score 2.92) by the respondents (Table 7.3). This indicator was also found as one of the most important indicators and was assigned a 10 out 10 importance score by the scorers.

**Productivity measure:**

Scorers assigned 9 out 10 for the importance of the productivity measure of the projects. Comparing the cost of producing outputs with benchmark cost ratio of comparable projects, the NAP (score 3.83 out of 5) and IPL (score 3.59 out of 5) were found more cost efficient compared to the cost efficiency of the HSSP-SP (score 2.92 out of 5). This can be explained by a perception of high input costs used and low outputs produced by HSSP-SP compared to other projects.

**Operational efficiency:**

Scorers assigned 8 out of 10 points for the importance of the business lead time (i.e., time taken from initiation to execution of an activity including approval process) both for regular and exceptional activities. IPL had the shortest business lead time (score 3.77 out of 5) for organizing regular activities, as the project had to go through fewer bureaucratic processes associated with the MoH structure and processes. The NAP scored moderately (score 3.30 out of 5) on this aspect, whereas HSSP-SP scored the lowest (score 2.60 out of 5) for the time they required to organize and implement project activities. This was because compared to the vertical nature of IPL and the NAP, activities of HSSP-SP were more horizontally linked to multiple units of the MoH and thus required more time in planning, coordinating and organizing them.

All three projects scored poorly (the NAP 2.87, IPL 2.75, and HSSP-SP 2.58 out of 5) on the indicator for business lead-time for organizing exceptional activities. This was because all three projects required additional clearance from their donors to be able to organize them and obtaining clearance for exceptional activities often involved a lengthy process with multiple exchange of communication and supporting documents. The figure below compares these different aspects of efficiency of the three projects:
It can be seen that except for the business lead time required for regular activities the NAP was perceived as more efficient than IPL and HSSP-SP on the aspects of project outputs, health systems outputs, productivity measure, and business lead time for exceptional activities. For business lead time required for regular activities, IPL was perceived as more efficient than the NAP and HSSP-SP. HSSP-SP was clearly perceived as less efficient than the NAP and IPL on all five aspects of efficiency considered. Overall, on the aspects of efficiency the total weighted score of the NAP was 69.48 out of 100. For IPL this score was 67.45, and for HSSP-SP it was 55.38.

### 7.4.5 Effectiveness

Questions 17 to 20 of the research questionnaire attempted to evaluate the outcomes and effectiveness of the selected projects against their strategic objectives and results framework as well as their effect on the health systems. This set of criteria attempted to measure the effectiveness in terms of the ratio of outcomes achieved and inputs used by the project. This set of criteria further assessed the relevance and robustness of the project outcomes in the country context and the effectiveness of the theory of change.
underlying the projects. The following table summarizes the weighted scores obtained by the projects.

**Table 7.9: Perceived effectiveness of three interventions**

<table>
<thead>
<tr>
<th>Question</th>
<th>Category</th>
<th>Importance weight</th>
<th>Maximum weighted Score</th>
<th>Weighted Score HSSP-SP</th>
<th>Weighted Score NAP</th>
<th>Weighted Score IPL</th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td>Project outcome</td>
<td>9</td>
<td>45</td>
<td>23.19</td>
<td>33.38</td>
<td>33.55</td>
</tr>
<tr>
<td>18</td>
<td>Health systems outcome</td>
<td>8</td>
<td>40</td>
<td>24.62</td>
<td>28.80</td>
<td>27.64</td>
</tr>
<tr>
<td>19</td>
<td>Relevance &amp; robustness of outcomes</td>
<td>8</td>
<td>40</td>
<td>28.00</td>
<td>34.67</td>
<td>34.18</td>
</tr>
<tr>
<td>20</td>
<td>Effectiveness of theory of change</td>
<td>8</td>
<td>40</td>
<td>25.23</td>
<td>32.00</td>
<td>30.91</td>
</tr>
<tr>
<td></td>
<td>Effectiveness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Converted score out of 100</td>
<td>100</td>
<td>61.24</td>
<td>78.09</td>
<td>76.53</td>
<td></td>
</tr>
</tbody>
</table>

**Project outcomes:**

Scorers assigned an importance rating of 9 out of 10 for achievement of project outcomes. Not assigning the highest importance score (10) to this item was due to one scorer who believed that in a fragile situation achieving expected project outcomes might not be realistic. Rather, projects should focus on getting the process right to build local capacity and strengthening the health system even if the outcomes were not readily visible. In terms of the rate of achieving the strategic objectives of the project, IPL scored highest among three projects getting 3.73 out of 5. The NAP was also considered moderately progressing towards achieving its strategic objectives and received an average score of 3.71 out of 5 (Table 7.3). This was consistent with the programme’s coverage and outcome performance.

Unlike IPL and the NAP, achievement of strategic objectives by HSSP-SP was considered low receiving an average score of 2.58 out of 5. This score is consistent with the World Bank’s internal evaluation of the project’s progress towards achievement of the primary development objective that considered HSSP-SP’s progress as “moderately satisfactory” until December 2012 (World Bank, 2013k).
Health system outcomes:

All three projects had strategic objectives related to health systems strengthening. For example, HSSP-SP had objectives for strengthening support services, human resource development, and management capacity in the health sector. The NAP had objectives for strengthening health systems related to HIV/AIDS and STI and strengthening community systems related to HIV/AIDS and STI. IPL also had objectives related to improving health service delivery, improving management capacity of health centres and improving M&E systems. Scorers ranked this indicator as 8 out of 10 for its importance.

All three projects showed moderate contributions towards health systems strengthening with the NAP receiving an average score of 3.60, IPL 3.45 and HSSP-SP 3.08 out of 5. While HSSP-SP contributed to improved governance and leadership with increased sector-wide coordination efforts, IPL’s micro-planning and strengthening district level management of health centres were appreciated by the respondents. The MoH also found the tools and processes used by the NAP useful for improving overall management capacity and learning for the Ministry. As one respondent said,

“Although Global Fund [funded programme] uses vertical system for the programme itself, we learn much from its management and implementation. For example, the service delivery area based budgeting tool used by the Global Fund programme can also be useful for preparing a results-based budget for the MoH by linking and tracking resource allocations with the service delivery and outputs” (a senior official from the MoH of Timor-Leste).

However, another respondent from the MoH believed all three projects could do better by integrating their activities and services at the district level for better coordination, resource allocation, and human resources management:

“We need to integrate those programmes with other disease programmes and government staffing. Otherwise it will be very difficult [for the government] to put one officer for each disease at the district level and coordinate those activities” (a senior policy maker from MoH).

Relevance and robustness of outcomes:

The questionnaire asked if project and health system outcomes claimed by the projects were clear, planned, causally linked, and system-wide and if there were available, accessible, credible, and disaggregable data to support the claims. It was found that the NAP was more advanced than the other projects on this aspect with their regularly documented and verified M&E system. IPL also produced good evidence to support their
claim. Although the World Bank was producing regular status reports for HSSP-SP, the results reported by the project were not clearly linked with the scope and resources allocated by HSSP-SP.

The NAP scored 4.33 out of 5 on an average on this aspect. IPL’s average score for this indicator was 4.27 whereas HSSP-SP scored 3.50.

**Theory of change:**

With clear linkages between resource allocations, activities, outputs, and outcomes, the respondents considered that the effectiveness of theory of change underlying the NAP was very strong (average score 4 out of 5). This was closely followed by IPL with an average score of 3.86 out of 5. However, the linkages between resource allocations, activity planning and intended outcomes of HSSP-SP were not perceived as strong as those of the other two projects. One respondent commented,

> “The design of HSSP-SP was flexible and it was agreed that it would remain flexible... But the whole project [HSSP-SP] focused on a procurement plan. The Ministry of Health’s ownership of the health sector strategic plan was minimum, as the plan was developed by the previous government. Without a work plan it was difficult to plan and implement the activities” (a respondent from the health sector specialization).

For theory of change, HSSP-SP received an average score of 3.15 out of 5. The scorers ranked theory of change as 8 out 10 for its importance.

The figure below compares the average scores of the three projects on different aspects of their effectiveness:
It can be seen that the NAP and IPL were clearly ahead of HSSP-SP in terms of effectiveness as perceived by the respondents on their scores on project outcomes, health systems outcomes, relevance and robustness of outcomes, and effectiveness of underlying theory of change. As shown in Table 9, on overall effectiveness, the total weighted score of the NAP was 78.09 out of 100, which was closely followed by IPL with 76.53. HSSP-SP’s overall weighted score on effectiveness was 61.24.

### 7.4.6 Sustainability

Questions 21 and 22 evaluated likely leverage and replication influenced by the project and expected lasting effect of the project outcomes beyond the external funding period. Table 7.10 summarizes the weighted scores obtained by each project on these aspects:
Table 7.10: Perceived sustainability of three interventions

<table>
<thead>
<tr>
<th>Question</th>
<th>Category</th>
<th>Importance weight</th>
<th>Maximum weighted Score</th>
<th>Weighted Score HSSP-SP</th>
<th>Weighted Score NAP</th>
<th>Weighted Score IPL</th>
</tr>
</thead>
<tbody>
<tr>
<td>21</td>
<td>Leverage &amp; replication</td>
<td>8</td>
<td>40</td>
<td>26.15</td>
<td>20.80</td>
<td>26.18</td>
</tr>
<tr>
<td>22</td>
<td>Expected lasting effect</td>
<td>10</td>
<td>50</td>
<td>25.19</td>
<td>25.33</td>
<td>33.64</td>
</tr>
<tr>
<td>Total</td>
<td>Sustainability</td>
<td>90</td>
<td>100</td>
<td>57.05</td>
<td>46.13</td>
<td>59.82</td>
</tr>
</tbody>
</table>

Replication and leverage:
The scorers ranked 8 out of 10 on the importance of the aspect of leverage and replication of a project. It was found that the likely leverage and replication influenced by the NAP was the lowest with an average score of 2.60 out of 5. Although the NAP worked closely with NGOs and in-country development partners (such as USAID, WHO, UNFPA and UNICEF), contributions from other funding sources to the NAP were very little. IPL’s project model, on the other hand, was considered as easily replicable to other districts not covered by the IPL and this project leveraged funding and resource allocations by the Government. IPL’s average score for leverage and replication was 3.27 out of 5. With a MDTF model, HSSP-SP also leveraged resource allocations from the government, AusAID, World Bank and European Union. HSSP-SP also received an average score of 3.27 out of 5 for likely leverage and replication.

Expected lasting effect:
The scorers considered the expected lasting effect and likely sustainability as one of the most important aspects and ranked it as 10 out of 10 for importance. The NAP scored lower than HSSP-SP and IPL. With relatively high amount of funding and complex activities, both HSSP-SP and the NAP scored low on the question of likely sustainability of the project outputs beyond their external funding period. As commented by one respondent,

“If the responsibility of funding HSSP-SP were wholly left with the country, the country might not prioritize the activities in the same way and might not undertake the same health systems reform initiatives” (An interviewee from the NGO sector in Timor-Leste).

Question 22 of the stakeholder interviews considered whether the country would own the projects, have necessary capacity and allocate necessary funding to carry on the projects without any external funding. HSSP-SP scored 2.52 out of 5 on this aspect.
whereas the NAP scored 2.53. This indicated that the MoH would probably be not in a position to allocate the current level of funding to carry on the activities currently funded by HSSP-SP and the Global Fund. However, one counter-argument from a respondent was:

“At the end of external funding period, HSSP-SP and Global Fund funded programmes will have achieved their objectives already. So, the MoH will not need to continue the same level of activities, rather need to allocate much lower budget to maintain the achievements. This lower amount of budget should be within the means and capability of the MoH” (A senior policy maker from the MoH).

With relatively lower budget and simpler programme than HSSP-SP and the NAP, IPL was considered to be likely sustainable beyond the funding period. In fact, IPL was designed to be handed over to the MoH at the end of the project period. IPL scored 3.36 out of 5 on the sustainability aspect of the project. One respondent was very confident about IPL’s contribution to sustainable results saying that,

“I don’t believe Timorese are lazy. I believe Timorese, if they receive proper guidance and support, they can do things and achieve results” (a respondent from technical assistance providers).

The figure below compares the average scores of three projects on the aspects of leverage and replication, and expected lasting effects of sustainability:

![Figure 7.5: Perceived sustainability of three interventions](image)

It can be seen that with higher leverage, replication, and likely continuation of project activities beyond the external funding period, IPL was clearly perceived as more sustainable than the NAP or HSSP-SP. The NAP, on the other hand, was perceived as
least sustainable by the respondents. Overall, the total weighted score for IPL on the aspect of likely sustainability was 66.46 out of 100, followed by 57.05 scored by HSSP-SP, and 51.26 scored by the NAP on this domain.

7.4.7 Summary scores on five domains and triangulation of results

Putting together the weighted scores on all five domains for the three projects, we can further compare their overall performance. The following table summarizes key findings from the balance scorecards:

<table>
<thead>
<tr>
<th>Table 7.11: Summary of balanced scorecards for three interventions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compliance with Paris principles</td>
</tr>
<tr>
<td>----------------------------------</td>
</tr>
<tr>
<td>Economy</td>
</tr>
<tr>
<td>Efficiency</td>
</tr>
<tr>
<td>Effectiveness</td>
</tr>
<tr>
<td>Sustainability</td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td>Overall converted score out of 100</td>
</tr>
</tbody>
</table>

It appears that the overall score of the NAP (68.07) closely matches with the overall score of the IPL (68.06) and both of them are placed ahead of the HSSP-SP, which has an overall score of 58.58. However, scores vary between the NAP and IPL, especially on individual domains of compliance with the Paris principles, and sustainability. The figure below compares the scores of the three projects on their individual domains:
We can see from the above chart that apart from the sustainability domain, HSSP-SP scored lower than the NAP and IPL in terms of its compliance with the Paris principles, economy, efficiency, and effectiveness.

The scores for the NAP showed that its compliance with the Paris principles and effectiveness of the project were very high. The NAP also received good scores on economy and efficiency. In fact, the programme’s score on the domains of compliance with the Paris Principles, efficiency, and effectiveness were the highest among three projects. However, the NAP’s score on sustainability was the lowest among three projects.

IPL scored closely with the NAP on the aspects economy, efficiency, and effectiveness. But IPL’s score on the compliance with the Paris principles was the lowest, although this
score was pretty close to the score of HSSP-SP on this aspect. However, IPL’s score on the aspect of sustainability was the highest among three projects.

The stakeholder views as found in the interviews and ranking are quite consistent with the project results. For example, stakeholders identified weak theory of change (i.e. the programme logic) behind the HSSP-SP design. To summarize its achievements, a World Bank document reported that till December 2012, HSSP-SP had:

“(i) supported direct delivery of health services in approximately 450 communities per month through the MoH’s outreach services, known as the Integrated Community Health Services, or Sistema Integrado de Saúde Comunitária (SISCa); (ii) improved health infrastructure by rehabilitating and upgrading the Central Medical Stores warehouse, or Serviço Autónomo de Medicamentos e Equipamentos de Saúde (SAMEs), the National Health Institute or Instituto Nacional de Saúde (INS), and three health posts; (iii) provided information and communication equipment and solar panels for a range of health facilities; (iv) increased health workforce capacity through a variety of short- and long-term training and professional development opportunities, including the provision of over 100 international scholarships, training of over 120 nurses and midwives, and reintegration training of over 500 returning medical students from Cuba and other places where they have received training; and (v) improved storage and distribution of pharmaceutical and medical supplies through direct funding for these medical inputs and installation of a new warehouse management system.” (World Bank, 2013k, pp. 4-5).

However, overall implementation of HSSP-SP was slow and in June 2013, HSSP-SP amended it’s primary development objective from: “improving the quality and coverage of preventive and curative health services, particularly for women and children, in order to accelerate progress toward the health Millennium Development Goals (MDGs)” to: “supporting the Government of Timor-Leste to get more resources to where they are needed most to improve the delivery of health services in districts and sub-districts” (World Bank, 2013k). Thus HSSP-SP received low scores on its performance.

Likewise, the scores related to the NAP’s performance were consistent with the project results and performance reports. As reported in the Grant Performance Report (The Global Fund, 2014), from January 2012 to December 2013, the NAP reached 1,095 men who have sex with men (MSM) against the target of 1,050 with outreach and prevention services indicating 88% coverage against the size of MSM population estimated in 2011. The Programme reached 815 female sex workers (FSW) with outreach and prevention services during the same period against the target of 800 achieving 97% coverage of total FSWs as estimated in 2011.
The outcome indicators of the NAP showed that HIV prevalence among FSW population decreased from 2.8% in 2010 to 2.0% in 2012. HIV prevalence also decreased among MSM population from 1.3% in 2010 to 1.0% in 2012. Condom use in the last sex act increased among MSM population from 51.6% in 2010 to 67.0% in 2012. Among men in uniform (e.g. Police and Military), condom use in last sex with a non-regular female partner increased from 21.8% in 2010 to 53.0% in 2012. However, condom use in the last sex act among FSWs decreased from 68.5% in 2010 to 63.0% in 2012. Knowledge level for correctly identifying modes of HIV transmission and rejecting common misconceptions decreased among MSM from 48.7% in 2010 to 11.0% in 2012; and among FSW from 56.5% in 2010 to 38.0% in 2012. The programme also fell short in achieving anti-retroviral treatment (ART) coverage and till June 2013, treated only 121 patients against the target of 249 achieving only 49% of the target (The Global Fund, 2014).

Like the NAP, IPL also received high scores on the achievement of its targets. IPL’s performance was strongly focused on the project’s goal of increasing immunization coverage among the children in Timor-Leste and they made good progress toward that goal. As pointed out in the final evaluation report of IPL —

“IPL did not achieve its extremely ambitious goal of raising the national average of DPT3 and measles vaccination in infants from 67.5% to 81.5%. However, it might have achieved that coverage target in its seven focus districts and come close nationally had there not been a national stock-out of measles vaccine in the first half of 2013. Various analyzes of vaccination coverage, including comparisons between IPL and non-IPL focus districts and analyzes based solely on numbers of children vaccinated (a logical approach, given the unreliability and yearly ups and downs in target populations), show significantly better coverage in IPL districts” (MCHIP, 2013g, p. 2).

It was also noted that IPL’s performance in 2013 was affected by a nation-wide stock out of measles vaccine and readiness of the health system to provide increased services. One respondent commented,

“We are creating demand [for immunization services]. But sometimes health system is not ready to provide the services. So, it is sometimes difficult to engage the communities if the [increased] demands for services are not met” (a respondent from the IPL team).

Still, immunization coverage in the districts with IPL intervention increased from 65% in 2011 to 79% in 2012 to 76% in 2013, whereas immunization coverage in non-IPL focus districts were 68% in 2011, 76% in 2012 and again dropped to 68% in 2013.
Overall Timor-Leste’s immunization coverage increased from 66% in 2011 to 78% in 2012 to 73% in 2013 (MCHIP, 2013a).

7.5 Discussion

The approach adopted in this study led to a composite performance score for each domain of evaluation. For the health care setting, many researchers have found it useful and convenient to use the balanced scorecard method to evaluate and compare performance of health care systems (Chen et al., 2006; Gauld et al., 2011; Inamdar et al., 2001; Peters et al., 2007; Pink et al., 2001; Voelker et al., 2001; Weir et al., 2009). However, none of these balanced scorecard approaches attempted to see the interplay of the Paris principles of aid effectiveness with the economy, efficiency, effectiveness, and sustainability of an aid funded intervention.

The stakeholder interviews highlighted similar backgrounds of need for coordination, harmonization and alignment. However, despite variations in individual project scores on various aspects of the Paris principles, overall all three projects scored above 60 out of 100 on the compliance indicating a good progress for Timor-Leste in implementing the Paris Principles of aid effectiveness. This is consistent with the findings from the OECD’s report for Timor-Leste for monitoring the principles of aid effectiveness (OECD, 2010) as summarized in the following table:
Table 7.12: Implementation of Paris Declaration in Timor-Leste

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Achievements or Challenges</th>
<th>Lessons or Priority Actions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ownership</td>
<td>Timor-Leste has without a doubt exercised increased ownership over its development agenda in recent years.</td>
<td>The priority is to translate this increased ownership into the writing and implementation of its own Strategic Development Plan (SDP) with development partners moving to a supporting role.</td>
</tr>
<tr>
<td>Alignment</td>
<td>To a large extent development partners have been able to align with the priorities outlined in the National Priorities Process though use of local systems has been very low.</td>
<td>With the SDP there will be even greater scope for development partners to align with government plans. Taking alignment forward, the Government is close to finalising a new aid policy and has plans in place to upgrade physical and human capacity of the NDAE [National Directorate of Aid Effectiveness], which will further contribute to alignment in the near future. Development partners must be more willing to provide information on their activities recognising this information is fundamental to policy development.</td>
</tr>
<tr>
<td>Harmonization</td>
<td>Co-ordination between development partners is making mixed progress. Recently there has been some progress towards harmonising aid by calling for joint efforts to prevent aid fragmentation, and increasing attention to joint analytic work.</td>
<td>DPs [Development Partners] need to engage in more joint mission planning, and joint analytic and programming work to reduce costs, duplication and promote alignment with government policy.</td>
</tr>
<tr>
<td>Managing for results</td>
<td>The National Directorate for Aid Effectiveness (NDAE) has been making progress in strengthening data collection and analysis capacity to allow for more results focused development programming.</td>
<td>Development partners need to align with the current results framework at each stage of the development of Timor-Leste. The new SDP should provide an improved results framework to drive the work of both the DPs and government.</td>
</tr>
<tr>
<td>Mutual accountability</td>
<td>The government has launched its transparency portal (<a href="http://www.transparency.gov.tl">www.transparency.gov.tl</a>) to promote transparency and mutual accountability from the government side. The government plans to launch the Aid Management Platform as part of its effort to strengthen mutual accountability between it and DPs.</td>
<td>With a stronger focus on managing for results will come an increased focus on the need for mutual accountability. Steps towards the use of Timor-Leste country systems as they become stronger will no doubt help sustain the focus on mutual accountability. Development partners can contribute to development of the government’s transparency portal and align with developing systems of government, work with government on the shared management of risks.</td>
</tr>
</tbody>
</table>

Source: (OECD, 2010, p. 21)

While Timor-Leste has made good progress on the aspects of aid coordination, harmonization, and alignment, it may appear contradictory that HSSP-SP employing a sector-wide approach scored lower than the vertical projects such as IPL and the NAP on the aspects of economy and efficiency. Although there have been persuasive arguments that a SWAp in health sector increases health sector coordination, stronger national leadership and ownership, and strengthened countrywide management and delivery systems (Hill, 2002), a strong sense of country ownership was not found for HSSP-SP.
A SWAp espouses increased reliance on budget support based on a government-owned comprehensive sector policy; greater ownership of the policy by key stakeholders; enhanced co-ordination and harmonization among major donors; and use of government systems (Hill, 2002). It is believed that a SWAp in the health sector reduces transaction costs of aid and has a systemic effect to increase efficiency of the government’s own resource management (Walford, 2003). However, studies by McNee (2012), Hutton (2004), Martinez (2008), Chansa et al. (2008) and Barakat (2009) indicate mixed results for SWAp in the health sector with no significant evidence for reduced transaction costs.

A DFID (2004) document, however, reckoned that in the short term the transaction costs for a sector-wide or budget support approach would be higher but over the time this cost would be reduced giving this approach more economy and efficiency:

“... the transaction costs of budget support are likely to be higher in the short term, as new ways of working between governments and donors are established, net transaction costs should be lowered over the medium term. This is because donors should be using the government’s own reporting and accounting systems to monitor progress, rather than negotiating, managing and monitoring literally thousands of projects. ... Preliminary evidence from Uganda indicates that transaction costs may have increased in the short run, as donors have yet to embed new ways of interacting with each other and with Government. However, there are reasonable prospects of a reduction in the medium term” (DFID, 2004, p. 7).

These reviews and findings seemed comparable to the findings about HSSP-SP in Timor-Leste. In fact the key strategic objective of HSSP-SP related to the strengthening health sector capacity and coordination for achieving health related MDGs was very ambitious and a slow process. Considering the fragile situation and low capacity of the government, it was fair to expect that HSSP-SP would require more sector investments and time before it could demonstrate any clearly attributable tangible results. The balanced scorecard, however, pointed out that HSSP-SP, still, could improve in the areas of compliance with the Paris principals, economy, operational efficiency and linking results with the inputs.

Unlike HSSP-SP, the NAP scored high on most of the aspects of the balanced scorecards. However, the overall score of the NAP was pulled down by its low score on the sustainability domain. According to its framework, the Global Fund funding was intended to be “additional” and not to replace any other funding sources that would
otherwise be available for the same disease programme in the country (The Global Fund, 2001). However, the huge amount of vertical funding from the Global Fund for a single disease programme, in fact, made it difficult for the country to sustain the same level of programming if external funding would be withdrawn.

The Global Fund worked in partnership with other development partners who also participated in a Country Coordinating Mechanism for the Global Fund grants. But it seemed that Global Fund’s partnership with other development partners did not leverage any additional funding for increased sustainability of the programme. This issue was also captured in a five-year evaluation of the Global Fund saying that:

“The Global Fund partnership model has opened spaces for the participation of a broad range of stakeholders. This progress notwithstanding, existing partnerships are largely based on good will and shared impact-level objectives rather than negotiated commitments or clearly articulated roles and responsibilities, and do not yet comprise a well-functioning system for the delivery of global public goods” (Macro International Inc, 2009, p. 33).

This low leverage and high volume of vertical funding explain the low score for the NAP on the sustainability domain of the balanced scorecard.

A review of the effects of global health initiatives including the Global Fund on country health systems found both positive and negative effects:

“Positive effects have included a rapid scale-up in HIV/AIDS service delivery, greater stakeholder participation, and channeling of funds to non-governmental stakeholders, mainly NGOs and faith-based bodies. Negative effects include distortion of recipient countries’ national policies, notably through distracting governments from coordinated efforts to strengthen health systems and re-verticalization of planning, management and monitoring and evaluation systems” (Biesma et al., 2009, p. 239).

This finding was equally applicable for the NAP in Timor-Leste. With the Global Fund funding, the NAP could scale up the intervention and involve NGOs in useful outreach work. However, Timor-Leste’s proposal to the Global Fund for Round 10 HIV/AIDS funding noted the following unintended negative effects of the Global Fund grant implementation:

“From its experience of implementing Global Fund funded National HIV/AIDS, TB and Malaria programs in the past, Timor-Leste has identified certain unintended consequences on the health systems such as: Uneven allocation/availability of resources for HIV/AIDS, TB and Malaria compared to other diseases creating destabilization of country priorities, inequality within health systems and jealousy among health workforce. Overtaxing of growing capacity of the management and workforce of the health systems due to additional efforts for meeting rigorous requirements of Global Fund grant management and
performance based funding, which often leave inadequate time and management resources to address other priorities of the health systems. Vertical implementation of Global Fund funded programs often creates additional demands on health systems for collecting, keeping and reporting data outside the regular reporting systems; facing verifications; having different salary and compensation packages (Govt funded and Grant funded) making it difficult to move towards a sector-wide approach with ‘one plan, one budget’ principle” (Timor-Leste Country Coordinating Mechanism for the Global Fund, 2010, p. 33).

The Multilateral Aid Review of DFID (2011) also noted the strengths of the Global funds in the areas of “focus on poor countries, critical role in meeting development objectives, climate and environment, cost and value consciousness and transparency” but weaknesses in the areas of their work related to “gender and fragile contexts” (p.21).

The balanced scorecard showed USAID-MCC funded IPL as comparatively more balanced in its scores in all five domains. IPL scored moderately on its compliance with the Paris principles but outperformed the other projects on economy and sustainability aspects. IPL’s scores on efficiency and effectiveness were also quite close to those of the NAP. This was probably because compared to HSSP-SP and the NAP, IPL had a simpler project design, easy to understand programme logic model, less and easily manageable funding, and time-bound clear exit strategy with agreed terms of reference to hand over the programme to the government at the end of the funding period. The small amount of funding requirement and the presence of the same programme for some districts made it easy for the MoH to understand, take over and own the programme at the end of the funding period.

IPL also had a number of other advantages over HSSP-SP and the NAP, which might have contributed to its effectiveness and sustainability scores. For example, IPL had comparatively fewer activities and a smaller budget and definite deadlines. For HSSP-SP and the NAP there were implied expectations of continuing funding beyond the agreement period. However, IPL had a definite project end date and exit strategy. IPL’s main activities and programme planning also took part at the district level in a ‘bottom-up manner’ whereas for HSSP-SP and the NAP, most programme planning took place at the central level.
IPL had clear counterparts with a client-vendor relationship with the MoH (client) and JSI (vendor). For HSSP-SP and the NAP, since project implementation teams were within the MoH, this relationship and accountability mechanisms were not very clear. As a client and approving authority, the MoH felt more empowered in dealing with the small IPL team than dealing with the large organizations such as the World Bank or the Global Fund.

Related studies, however, question the accountability mechanism and the effect on the state institution building of the contracted out service delivery mechanism as adopted in IPL’s case. Batley and Mcloughlin (2010) argued,

“in the context of most developing countries, non-state actors have an important role in service delivery. The questions are whether, particularly in the case of fragile states, governments have the capacity to perform the indirect roles that only they can perform, and whether inadequate performance risks bringing about damage to service delivery” (p.135).

IPL, in fact, provides an example of adapting a humanitarian aid approach of funding outside the government mechanisms for addressing a developmental need of service delivery and capacity building. The short project cycle and inadequate time for capacity building are among the concerns of this approach. As Canavan et al. (2008) pointed out:

“A criticism of adapting humanitarian instruments to development contexts is that they generally still retain inappropriately short project cycles, which include exit plans that can result in health service delivery gaps, while having limited long-term capacity building components” (p.29).

A DFID review (Cox & Hemon, 2009) also acknowledges that in fragile situations restoring services require a significant element of non-state service delivery. However, along with the benefits of not-state service delivery, the policy makers need to be aware of the unintended effect of such mechanism on policy coherence and state-building in fragile states and situations. A joint guideline from the United Nations Development Group and World Bank (2007) identified that:

“Non-state parallel service delivery channels may initially be necessary, particularly to meet immediate needs, but these can detract from building state capacity in the longer run if there is no explicit transition or ‘exit’ strategy. Perhaps more importantly, while service delivery through NGOs, private sector, donor or international agencies can play a vital role in protecting the welfare of vulnerable groups, efforts are also needed to build the capacity, accountability and credibility of the state in the eyes of the population. If all the positive recovery projects are ‘branded’ with the logos of donor or international agencies rather than being provided under the auspices of the state, the population will associate these services with international partners rather than building a sense of trust and connection with their own institutions and leadership” (p.15).
One can further compare the findings from the balanced scorecards with the scores from the Quality of Official Development report (Birdsall & Kharas, 2014). Developed by the Center for Global Development, the Quality of Official Development Assistance (QuODA) measures “donors’ performance on 31 indicators of aid quality to which donors have made commitments. The indicators are grouped into four dimensions associated with effective aid: maximizing efficiency, fostering institutions, reducing the burden on partner countries, and transparency and learning” (Hashmi, Birdsall, & Kharas, 2014).

The following is a comparison of aid performance of the World Bank (IDA), the Global Fund (GF), and the US Government (USA) as accessed from the QuODA 2014 interactive database

![QuODA assessment of the World Bank, Global Fund, and US Government aid](http://www.cgdev.org/page/quality-oda-quoda)

**Figure 7.7: QuODA assessment of the World Bank, Global Fund, and US Government aid**


In the above chart, zero (0) represents the mean score. Scores within the gray diamond are below the mean and scores outside this area represent above the mean scores. It can be seen that according to QuODA assessment, the Global Fund scored very high in the dimensions of maximizing efficiency and transparency and learning. The Global Fund’s score on reducing burden for the recipient country was slightly higher than the average. However, consistent with the findings from the balanced scorecards on sustainability score, Global Fund’s performance on fostering institutions was below the average.
In the QuODA assessment, the World Bank (IDA) scored higher than the average scores on all four dimensions with very strong performance in fostering institutions and reducing recipient country’s burden. However, the World Bank’s performance on maximizing efficiency was lower than that of the Global Fund.

According to this database, the US Government aid performed very poorly on reducing burden with a score far below the average level. US Government aid also scored lower than the average scores on maximizing efficiency and fostering institutions but scored slightly higher than the average score on the aspect of transparency and learning. However, the performance of the Millennium Challenge Corporation (MCC) that funded the IPL was much different from the US Government aid performance showed here. This was clarified by QuODA in their brief saying that:

“The Millennium Challenge Corporation (MCC), which allocates less than 7 percent of overall ODA from the United States, performs very well in our assessment. MCC outperforms USAID on each of QuODA’s four dimensions of aid quality. The agency’s legislative mandate is to support only low and lower-middle income countries whose governments score well on such indicators as ‘ruling justly’; its high score on the share of its aid allocated to well-governed countries is therefore no surprise. It also scores well on providing high shares of untied aid and aid to partners with good operational strategies” (Hashmi et al., 2014, p. 5).

It is, therefore, no surprise that IPL’s overall score on the balanced scorecard was high compared to the QuODA scores on general US Government aid.

Although findings from the stakeholder interviews were triangulated with the focus group discussions and actual project results, it is important to be aware of the limitation of this chapter that it predominantly presents the perceptions of the stakeholders. The perceptions can be biased by the sample size, availability, and selection of the interviewees as well as by the individual beliefs, experience, and opinions of the selected stakeholders. Therefore, it is important to look into other empirical evidence to support the programme theories and verify the stakeholder views before forming any definitive informed opinions about the effectiveness of aid management modalities of the three interventions.
7.6 Conclusion

Using the Paris principles of aid effectiveness and evaluation criteria for economy, efficiency, effectiveness, and sustainability this chapter developed a set of ranking questions around these five domains to evaluate and compare three selected projects from Timor-Leste’s health sector. This chapter used stakeholder interviews with the structured ranking questionnaire, focus group discussions, and available project data to present the findings in the form of balanced scorecards.

Contrary to the belief that a sector-wide approach would be more compliant to the Paris principles, it was found that the NAP was more compliant to the Paris principles than the MDTF funded HSSP-SP and USAID funded IPL project. However, the NAP’s score on the aspect of sustainability was the lowest among three projects and IPL scored the highest on this aspect. Overall, the total score of IPL (340.28 out of 500) closely matched with the total score of the NAP (340.34 out of 500) and both of them surpassed the scores of HSSP-SP on the aspects of economy, efficiency, effectiveness, and total score (292.88 out of 500).

These scores led to the analysis of appropriateness of different aid management modalities in a fragile state context. Although it is a common practice to use multiple aid management modalities in fragile situations, each modality has their own set of advantages and disadvantages. Donors and recipient countries need to analyze the context and carefully choose the appropriate mix or innovative approaches to select an aid management modality to help a fragile state achieve the dual goals of service delivery and state-building.
8 Cost-Effectiveness and technical efficiency

8.1 Introduction

Chapter 7 compared and analyzed stakeholders’ perspectives in the form of balanced scorecards on the aspects of compliance with the Paris principles, economy, efficiency, effectiveness, and likely sustainability of three projects selected from the Timor-Leste’s health sector. This chapter provides empirical evidence through cost-effectiveness and technical efficiency analysis of each project. After introducing and explaining the concepts of cost effectiveness and technical efficiency, analyses are presented in three different sections. Each section explains the methods used for the cost effectiveness analysis of one particular project, and presents the results found, and sensitivity around the results. These sections were then followed by a discussion.

8.2 Key concepts and cost-effectiveness analysis framework

Cost effectiveness analysis is useful for comparative effectiveness analysis of different health programmes or strategies for informed resource allocation decisions (Drummond et al., 2005; Gold et al., 1996; Jamison et al., 2006a, 2006d; Jamison et al., 1993; Levin & McEwan, 2001; Miguel & Kremer, 2004).

In order to relate the outcomes of an intervention to its costs, a cost effectiveness ratio (CER) is calculated by dividing the costs by its benefits. In a cost effectiveness analysis (CEA), while costs are measured in a common monetary value (such as US Dollars), the effectiveness of options are measured in different physical units according to the outcomes such as the number of cases treated, number of infections prevented, number of lives saved and so on. Because these two units of measurement of costs and effects are incommensurable, the ratio of costs to effectiveness is calculated in the following way:

\[ \text{Cost effectiveness ratio (CER)} = \frac{\text{Costs of an option (C)}}{\text{Effects of that option (E)}} \]

The CER that is most often used in economic analysis of healthcare services is an incremental cost effectiveness ratio, or ICER. An ICER is calculated by comparing the incremental costs and benefits of an option over the costs and benefits of its next best alternative:
That means, the incremental cost effectiveness ratio of Programme 1 over Programme 2 equals to difference of costs of Programme 1 and Programme 2 divided by the difference of effects of Programme 1 and Programme 2. The ICER shows the incremental cost of obtaining an incremental unit of benefit.

However, one problem of such measurement is related to the different outcomes of an intervention making it difficult to aggregate all benefits. While a different method of economic analysis such as cost-benefit analysis attempts to convert different types of benefits into monetary values, it is often difficult to assign a monetary value against all the benefits (McCabe, 2009).

One way of dealing with this problem of evaluating multiple outcomes of a health intervention is by using a common health index such as ‘quality adjusted life years’ (QALYs) or ‘disability adjusted life years’ (DALYs) that combines both the quantity of lives saved and the quality of life improved (McCabe, 2009). Both DALYs and QALYs express health in time (life years) and give a weight to years lived with a disease or disability. As DALYs measure health loss and QALYs health gain, they express an inverse value (Gold, Stevenson, & Fryback, 2002). The National Institute of Health and Care Excellence of the UK also acknowledges the appropriateness of such generic measure of health benefits saying that: “Currently, the QALY is considered to be the most appropriate generic measure of health benefit that reflects both mortality and health-related quality of life effects” (National Institute for Health and Care Excellence, 2013, p. 37).

At the population level, DALYs measure disease and injury burden in terms of total number of years lost from the lives of a population due to disability and death from an illness, compared to the total number of years the population would live in a perfect health condition (Musgrove & Fox-Rushby, 2006). DALYs for a specific cause are calculated as the sum of the Years of Life Lost (YLL) from that cause and the Years Lived in Disability (YLD) for people living in states of less than good health resulting from the specific cause (Musgrove & Fox-Rushby, 2006):
DALY = YLL + YLD

The YLLs for a cause are calculated as the number of cause-specific deaths multiplied by a life loss function specifying the years of life lost due to deaths (that is life expectancy at the age of death):

\[ YLL = N \times L \]

Where, \( N \) is the number of deaths due to the cause and \( L \) is the life expectancy at the age of death.

To estimate YLDs for a particular cause in a particular time period, the number of incident cases in that period is multiplied by the average duration of the disease and a disability weight factor that reflects the severity of the disease on a scale from 0 (perfect health) to 1 (dead):

\[ YLD = I \times DW \times L \]

Where, \( I \) is the number of incident cases for the given cause; \( DW \) is the disability weight for the given cause; and \( L \) is the average duration of the case (in years) until remission or death.

As the method evolved, the calculation of DALYs additional factors were considered, such as time discounting and age weights to estimate the net present value of the DALYs in terms of opportunity costs of economic productivity (Murray & Lopez, 1996). In such calculations, a 3% time discounting was often considered and less weight was given to years lost at young and older ages. However, following consultations with relevant WHO programmes, collaborators and expert advisory groups in late 2012, WHO decided to adopt a simplified calculation methods for DALYs and dropped the time discounting and the non-uniform age weights in the its calculation of DALYs for Global Health Estimates (WHO, 2013a). As WHO (2013a) explained:

"The original GBD 1990 study and subsequent WHO updates published DALYs computed with a 3% discount rate for future lost years of healthy life and an alternative set with a 0% discount rate. The arguments for discounting future health were couched mainly in terms of avoiding various decisionmaking paradoxes when future costs of health interventions are discounted (Murray & Acharya, 2002). Critics have argued that there is no intrinsic reason to value a year of health as less important simply because it is in the future (Tsuchiya, 2002) and the experts consulted for the GBD 2010 study also advised against discounting, particularly in the context where the DALY has been more explicitly defined as quantifying loss of health, rather than the social value of loss of health. This also avoids the inconsistency in the original DALY method,"
where the start time for discounting future stream of YLDs was the year of incidence, whereas the start time for discounting YLLs was the year of death rather than the year of incidence” (p.7).

Unlike the DALY calculations used in the Global Burden of Disease (Institute for Health Metrics and Evaluation, 2013) studies, Global Health Estimates (GHE) uses a standard life table for both males and females with a life expectancy of about 92 years at birth without discounting and age-weighting. Consistent with the DALY measures used by WHO in Global Health Estimates (WHO, 2014), this chapter uses the simplified methods of calculating DALYs without time discounting and with uniform age weights compared to the highest level of life expectancies projected for the year 2050 by the World Population Prospects 2012 (United Nations Department of Economic and Social Affairs Population Division, 2013).

In order to measure the effects of a programme such as DALYs saved by an intervention, cost effectiveness analysis method presupposes “a well-specified intervention and a no-intervention condition, or control group, against which the intervention is compared” (McEwan, 2012, p. 190). The contribution of an intervention to the expected outcome can be measured by the difference between the outcome observed with the intervention and the likely outcome in a hypothetical scenario without the intervention. The cost effectiveness analysis in this chapter attempts to calculate and compare the cost per DALY saved by each selected project. Since all three projects were implemented in addition to the government’s efforts towards the same objectives, the observed outcomes of each project was compared with the likely outcomes in two other hypothetical scenarios: 1) the null or control scenario without the presence of the intervention; and 2) the optimum scenario if all the project targets were achieved and all the allocated resources were fully utilized.

The achievements of each project were measured by calculating DALYs saved by each project from the projected difference of outcomes between the current and null scenario. This chapter then attempts to measure the technical efficiency of each project by comparing the actual achievements of a project with the projected achievements by the project in the optimum scenario. It has been assumed that each project’s budget and performance targets were designed to achieve the maximum results by utilizing minimum resources. Therefore, in the optimum scenario a project’s performance would
be 100% efficient if the project achieved all its targets fully by utilizing 100% of the allocated budget.

Based on this assumption, the technical efficiency of each project was measured by comparing the actual cost per DALY saved by the project with the likely cost per DALY saved by the project in the optimum scenario:

\[
\text{Technical Efficiency} = \frac{\text{Cost per DALY saved in optimum scenario}}{\text{Cost per DALY saved in actual scenario}}
\]

### 8.3 Cost effectiveness and technical efficiency of HSSP-SP

#### 8.3.1 Methods

**Sources of data:**

Since the HSSP-SP aimed to strengthen the MoH’s Health Sector Strategic Plan: 2007-12 (Ministry of Health of Timor-Leste, 2007) with the overall objective to improve the quality and coverage of preventive and curative health services, it was difficult to measure and attribute particular cause-specific DALYs saved by the project. Following the economic analysis conducted by the World Bank (2007a) during project appraisal, this chapter attempted to calculate DALYs averted by HSSP-SP from the DALYs estimates for Timor-Leste from the global databases such as the Global Burden of Disease (GBD) estimates of the Institute for Health Metrics and Evaluation (Institute for Health Metrics and Evaluation, 2013) and the Global Health Estimates (GHE) of the WHO (2014).

Data related to the project design, costs and intervention results were collected from various project documents of HSSP-SP including project agreement, budget, project appraisal documents, project status reports, and audit reports collected from the researcher’s fieldwork as well as from the Website of the World Bank (http://www.worldbank.org/projects/P104794/health-sector-strategic-plan-support-project?lang=en).
Estimating DALYs for Timor-Leste:
The total DALYs estimates for Timor-Leste for the years 1990, 1995, 2000, 2005 and 2010 were collected from the website (http://www.healthdata.org/) of the Institute of Health Metrics and Evaluation that led the GBD Study of 2010 (Institute for Health Metrics and Evaluation, 2013). These estimates were compared with the recent GHE for Timor-Leste published by WHO (2014) for the years 2000 and 2012.

As noted in the WHO methods and data sources for the global burden of disease estimates (WHO, 2013a), there are significant differences in the methods followed in GBD and GHE making it difficult to compare the DALYs across years. For example, although both GBD and GHE used a simpler form of calculating DALYs without age-weighting and time discounting, GHE used prevalence estimates adjusted for independent comorbidity for calculating years lived in disability (YLD) instead of incidence rates used in the GBD. GBD used a standard life table with about 86 years life expectancy at birth based on the highest life expectancy at birth by Japanese women. However, in GHE the standard life table used for calculation of years of life lost for a death at a given age is based on the projected frontier life expectancy for 2050, with a life expectancy at birth of 92 years. Moreover, GHE calculated the years of life lost from mortality (YLLs) by using recently revised WHO estimates of deaths by region, cause, age and sex for years 2000-2011 (WHO, 2013a, 2013b). In addition to these differences, the total number of disease sequelae and disability weights of some disease sequelae were also different in GHE from GBD.

In order to derive a comparable indicative estimate of Timor-Leste's DALYs in the GHE methods from the GBD estimates, the GBD estimates for Timor-Leste for the year 2000 were compared with the corresponding estimates for 2000 in the GHE to derive the percentage of change in the DALYs estimates per 100,000 population between GBD and GHE for all clusters of causes: communicable, maternal, neo-natal, perinatal, and nutritional conditions; non-communicable diseases; injuries. The percentage of change of DALYs per 100,000 was then applied to the GBD estimates for the year 2005 and 2010 to derive the indicative DALYs estimates for those years in GHE methods. The indicative DALYs rates for the years from 2000 to 2005 and from 2005 to 2010 and finally from 2010 to 2012 were then interpolated to derive the indicative DALY rates for
the years from 2008 to 2012. These DALYs rates were then applied to the Timor-Leste’s population estimates as per 2012 population data from the UN Population Department (United Nations Department of Economic and Social Affairs Population Division, 2013) to derive the total estimated DALYs for Timor-Leste for each single year from 2008 to 2012.

**Estimating DALYs for counterfactual and target scenarios:**

The economic analysis conducted by the World Bank during the project appraisal of HSSP-SP (World Bank, 2007a) assumed that without HSSP-SP intervention, there would be 5% reduction of DALYs per 100,000 population in Timor-Leste from 2008 to 2009, 4% reduction from 2009 to 2010, 3% reduction from 2010 to 2011 and 2% reduction from 2011 to 2012 for all clusters of causes: i) Communicable, maternal, neo-natal, perinatal, and nutritional conditions; ii) Non-communicable diseases; and iii) Injuries. This analysis estimated two different target scenarios—the pessimistic and the optimistic scenario for total DALYs in Timor-Leste with the HSSP-SP intervention.

For the pessimistic target scenario, the World Bank project appraisal document (World Bank, 2007a) estimated that with HSSP-SP intervention there would be 15% reduction of DALYs in 2012 from that of 2008 in the cluster of communicable, maternal, neo-natal, peri-natal and nutritional conditions; 5% reduction of DALYs in the cluster of non-communicable diseases; and 10% reduction of DALYs in the cluster of injuries. For the optimistic targets scenario, these reductions of DALYs were estimated as 25% in the cluster of communicable, maternal, neo-natal, peri-natal and nutritional conditions; 8% in the cluster of non-communicable diseases; and 15% in the cluster of injuries in 2012 from that of 2008.

These two scenarios were totally speculative for the sake of economic analysis conducted by the World Bank analyst and did not claim to represent any empirical evidence. The optimistic scenario was speculated with the assumption that the performance and health sector’s response for the HSSP-SP would be most favourable. The pessimistic scenario, on the other hand, was speculated with the assumption of a not so favourable performance and the health sector’s response. While an optimistic scenario indicates some ambitious targets, the pessimistic scenario provides a more
conservative target. Both these scenarios have been taken from the economic analysis section of the World Bank’s project appraisal document (World Bank, 2007a). The World Bank document, however, used WHO Global Burden of Disease estimates for 2002 (WHO, 2004a) as the baseline for its calculations. As derived from the DALYs estimates of the GBD and GHE described earlier, this chapter used Timor-Leste’s DALYs estimates for 2008 (in GHE methods) as baseline. Estimated DALYs in each cluster of causes were then calculated in the counterfactual (without HSSP-SP intervention), pessimistic targets, and optimistic targets scenarios by applying the assumptions and estimates from the World Bank project appraisal document mentioned above. For calculating the yearly rates of reduction of DALYs in the pessimistic and optimistic target scenarios, the estimated total reduction rates from 2008 to 2012 were evenly distributed for the years 2009, 2010 and 2011 by applying one-quarter of the reductions of each cluster to each interim year.

Calculating DALYs saved, cost effectiveness ratio and technical efficiency:

DALYs saved by the HSSP-SP intervention were calculated from the difference between the estimated DALYs in the counterfactual scenario (i.e. without HSSP-SP intervention) and actual DALYs estimates for each cluster of causes for each year from 2008 to 2012. Similarly, the total DALYs that would be likely saved in the pessimistic and optimistic target scenarios were calculated from the difference between the actual DALYs estimates and the expected DALYs in the pessimistic and optimistic targets scenario for each cluster of causes for each year from 2008 to 2012.

The cost effectiveness ratio in terms of cost per DALY saved was calculated by dividing the total expenditure of HSSP-SP by total DALYs saved by HSSP-SP from 2008 to 2012. Similarly, the likely costs per DALY saved in the pessimistic and optimistic target scenarios were calculated by dividing the total allocated cost (i.e. the total budget) by the total estimated DALYs saved from 2008 to 2012 in these scenarios. As explained earlier, the technical efficiency of HSSP-SP against its pessimistic and optimistic targets were calculated by separately dividing the expected costs per DALY saved in the pessimistic and optimistic scenarios by the actual cost per DALY saved by HSSP-SP.
8.3.2 Results for HSSP-SP

In order to posit the results in their own context, Table 8.1 provides an overview of the key parameters of HSSP-SP such as evaluation period, goal, objectives, total input value, reported outputs, and reported outcomes of each project:

**Table 8.1: Overview of HSSP-SP**

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Overview</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evaluation period</td>
<td>June 2008 to December 2012</td>
</tr>
<tr>
<td>Goal</td>
<td>To improve the quality and coverage of preventive and curative health services, particularly for women and children, in order to accelerate overall progress toward the health MDGs</td>
</tr>
</tbody>
</table>

| Objectives | HSSP-SP had two parts (a) direct financial support based on the Health Sector Strategic Plan and Medium Term Expenditure Framework (MTEF) and (b) addressing remaining challenges and innovations needed for the health sector to be prepared for the next generation of issues. Specific objectives for the first part: (i) improve accessibility, demand and quality of health services; (ii) strengthen support services, human resource development, and management; and (iii) strengthen coordination, planning and monitoring. The second objective was to promote innovation and programme development (linked to priorities identified in the HSSP) by providing resources to pilot and evaluate priority health sector innovations. |

| Inputs | Total allocation: $21 million USD |
|        | Total estimated expenditure: $16.24 million USD |

| Reported outputs | (i) supported direct delivery of health services in approximately 450 communities per month through the MoH’s outreach services (SISCa); (ii) improved health infrastructure by rehabilitating and upgrading the Central Medical Stores warehouse (SAMES), the National Health Institute (INS), and three health posts; (iii) provided information and communication equipment and solar panels for a range of health facilities; (iv) increased health workforce capacity through a variety of short- and long-term training and professional development opportunities, including the provision of over 100 international scholarships, training of over 120 nurses and midwives, and reintegration training of over 500 returning medical students from Cuba and other places where they have received training; and (v) improved storage and distribution of pharmaceutical and medical supplies through direct funding for these medical inputs and installation of a new warehouse management system. |

| Reported outcomes | Percentage of children under 1 year vaccinated against (i) DPT3; and (ii) measles increased from 63% and 61% (Dec 2006) to 83% and 73% (Dec 2012) respectively; Percentage of births attended by skilled health personnel increased from 27 (Dec 2006) to 59 (Dec 2012); Percentage of pregnant women receiving four or more prenatal visits increased from 36% (Dec 2006) to 49% (Dec 2012); Percentage of children (6-59 months) receiving vitamin A supplements increased from 36% (Dec 2006) to 57% (Dec 2012); Contraceptive Prevalence Rate increased from 10% (Dec 2006) to 26% (Dec 2012). |

Source: Author’s review of project appraisal, project status reports, and audit reports of HSSP-SP

Table 8.2 presents the estimated DALYs per 100,000 population for Timor-Leste for the years from 2008 to 2012 for four different scenarios: a) without HSSP-SP intervention (counterfactual scenario), b) with HSSP-SP intervention (actual scenario), c) if pessimistic targets were achieved (pessimistic target scenario), and d) if optimistic targets were achieved (optimistic target scenario):
### Table 8.2: Estimated DALYs per 100,000 in Timor-Leste, 2008-2012

<table>
<thead>
<tr>
<th></th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>Rate of change 2008-2012</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>a. Without intervention (Counterfactual scenario)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Communicable, neo-natal, peri-natal and nutritional diseases</td>
<td>29,738</td>
<td>28,251</td>
<td>27,121</td>
<td>26,307</td>
<td>25,781</td>
<td>13.3%</td>
</tr>
<tr>
<td>Non-Communicable diseases</td>
<td>20,592</td>
<td>19,563</td>
<td>18,780</td>
<td>18,217</td>
<td>17,852</td>
<td>13.3%</td>
</tr>
<tr>
<td>Injuries</td>
<td>3,518</td>
<td>3,342</td>
<td>3,208</td>
<td>3,112</td>
<td>3,050</td>
<td>13.3%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>53,848</td>
<td>51,156</td>
<td>49,109</td>
<td>47,636</td>
<td>46,683</td>
<td>13.3%</td>
</tr>
<tr>
<td><strong>b. With intervention (actual scenario)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Communicable, neo-natal, peri-natal and nutritional diseases</td>
<td>29,738</td>
<td>27,514</td>
<td>25,291</td>
<td>24,774</td>
<td>24,257</td>
<td>18.4%</td>
</tr>
<tr>
<td>Non-Communicable diseases</td>
<td>20,592</td>
<td>20,349</td>
<td>20,105</td>
<td>17,819</td>
<td>15,533</td>
<td>24.6%</td>
</tr>
<tr>
<td>Injuries</td>
<td>3,518</td>
<td>3,437</td>
<td>3,356</td>
<td>3,825</td>
<td>4,294</td>
<td>-22.1%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>53,848</td>
<td>51,300</td>
<td>48,752</td>
<td>46,418</td>
<td>44,084</td>
<td>18.1%</td>
</tr>
<tr>
<td><strong>c. With pessimistic targets (pessimistic target scenario)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Communicable, neo-natal, peri-natal and nutritional diseases</td>
<td>29,738</td>
<td>27,782</td>
<td>25,826</td>
<td>23,870</td>
<td>21,914</td>
<td>26.3%</td>
</tr>
<tr>
<td>Non-Communicable diseases</td>
<td>20,592</td>
<td>19,684</td>
<td>18,776</td>
<td>17,868</td>
<td>16,960</td>
<td>17.6%</td>
</tr>
<tr>
<td>Injuries</td>
<td>3,518</td>
<td>3,325</td>
<td>3,131</td>
<td>2,938</td>
<td>2,745</td>
<td>22.0%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>53,848</td>
<td>50,791</td>
<td>47,733</td>
<td>44,676</td>
<td>41,619</td>
<td>22.7%</td>
</tr>
<tr>
<td><strong>d. With optimistic targets (optimistic target scenario)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Communicable, neo-natal, peri-natal and nutritional diseases</td>
<td>29,738</td>
<td>27,137</td>
<td>24,537</td>
<td>21,936</td>
<td>19,336</td>
<td>35.0%</td>
</tr>
<tr>
<td>Non-Communicable diseases</td>
<td>20,592</td>
<td>19,550</td>
<td>18,508</td>
<td>17,466</td>
<td>16,424</td>
<td>20.2%</td>
</tr>
<tr>
<td>Injuries</td>
<td>3,518</td>
<td>3,287</td>
<td>3,055</td>
<td>2,824</td>
<td>2,592</td>
<td>26.3%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>53,848</td>
<td>49,974</td>
<td>46,100</td>
<td>42,226</td>
<td>38,352</td>
<td>28.8%</td>
</tr>
</tbody>
</table>

Source: Calculated from (WHO, 2014) and (Institute for Health Metrics and Evaluation, 2013)

It can be seen that with HSSP-SP intervention Timor-Leste had an estimated overall reduction of DALYs per 100,000 population by 18.1% between the years 2008 and 2012. Without HSSP-SP intervention, this reduction would be 13.3%. However, this estimate is lower than the pessimistic (22.7%) and optimistic targets (28.8%) of the project.

The low achievement of targets can be explained by a very slow achievement of targets in 2009, in which year the total DALYs were, in fact, slightly higher than the projected DALYs without intervention. This was probably because of a slow start of the project activities, initial confusion, challenging sector-wide coordination issues, and the state fragility situation followed by an attempted assassination on the life of the President of Timor-Leste in 2008.
Another remarkable observation is related to the DALYs for Injuries. While DALYs for Communicable and Non-communicable disease clusters decreased more steadily, DALYs for Injuries increased in 2011 and 2012 giving an overall increase for Injuries related DALYs by 22.1% between the year 2008 and 2012. This coincides with the recent developments in the roads and transport sector in Timor-Leste with increased number of vehicles and road accidents and indicates an unintended effect of development on the health systems.

The above DALYs rates were applied to the total population size of Timor-Leste for corresponding years to estimate the total DALYs in four different scenarios and to calculate DALYs saved by the HSSP-SP intervention and DALYs expected to be saved in the pessimistic target scenario and the optimistic target scenario. This also allowed the calculation of crude technical efficiency (i.e. without considering the costs) of HSSP-SP in terms of DALYs saved in comparison with the DALYs expected to be saved in the pessimistic and optimistic target scenarios.

Table 8.3 presents the results related to the calculation of total DALYs in different scenarios, DALYs saved by HSSP-SP, and crude technical efficiency of HSSP-SP against the pessimistic and optimistic targets:
Table 8.3: Total DALYs in different scenarios and DALYs saved by HSSP-SP in Timor-Leste, 2008-2012

<table>
<thead>
<tr>
<th></th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>1,050,244</td>
<td>1,064,191</td>
<td>1,079,450</td>
<td>1,096,293</td>
<td>1,114,106</td>
<td></td>
</tr>
<tr>
<td>Total DALYs without</td>
<td>565,535</td>
<td>544,393</td>
<td>530,111</td>
<td>522,231</td>
<td>520,102</td>
<td></td>
</tr>
<tr>
<td>intervention</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total DALYs with current</td>
<td>565,535</td>
<td>545,930</td>
<td>526,255</td>
<td>508,879</td>
<td>491,144</td>
<td></td>
</tr>
<tr>
<td>results</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total DALYs with optimistic</td>
<td>565,535</td>
<td>540,509</td>
<td>515,257</td>
<td>489,779</td>
<td>463,675</td>
<td></td>
</tr>
<tr>
<td>targets</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total DALYs averted by</td>
<td>0</td>
<td>-1,538</td>
<td>3,856</td>
<td>13,352</td>
<td>28,958</td>
<td>44,628</td>
</tr>
<tr>
<td>HSSP-SP: Current results</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DALY averted by</td>
<td>0</td>
<td>3,884</td>
<td>14,854</td>
<td>32,452</td>
<td>56,427</td>
<td>107,616</td>
</tr>
<tr>
<td>HSSP-SP: Pessimistic targets</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DALY averted by</td>
<td>0</td>
<td>12,573</td>
<td>32,482</td>
<td>59,307</td>
<td>92,815</td>
<td>197,178</td>
</tr>
<tr>
<td>HSSP-SP: Optimistic targets</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crude Technical Efficiency</td>
<td>n/a</td>
<td>-39.59%</td>
<td>25.96%</td>
<td>41.14%</td>
<td>51.32%</td>
<td>41.47%</td>
</tr>
<tr>
<td>against pessimistic targets</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crude Technical Efficiency</td>
<td>n/a</td>
<td>-12.23%</td>
<td>11.87%</td>
<td>22.51%</td>
<td>31.20%</td>
<td>22.63%</td>
</tr>
<tr>
<td>against optimistic targets</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Calculated from (WHO, 2014) and (Institute for Health Metrics and Evaluation, 2013)

The following chart provides a graphical representation of DALYs saved by the HSSP-SP intervention and DALYs expected to be saved in its pessimistic and optimistic targets:

![Figure 8.1: DALYs saved by HSSP-SP against pessimistic and optimistic target scenarios](image)

It can be seen that from 2008 to 2012, HSSP-SP saved a total 44,628 DALYs against its pessimistic targets of saving 107,616 DALYs and optimistic targets of saving 197,178 DALYs. This gives the intervention a crude technical efficiency of 41.47% against the pessimistic targets and 22.63% against the optimistic targets. However, in order to calculate the cost effectiveness ratio and technical efficiency considering costs per DALY
saved in different scenarios, the budget and expenditure related data were considered for HSSP-SP.

Table 8.4 presents the cost per DALY saved and technical efficiency of HSSP-SP in different scenarios by comparing the expenditure and budget with the DALYs saved by the current results and target scenarios:

<table>
<thead>
<tr>
<th>Table 8.4: Cost effectiveness ratio and technical efficiency of HSSP-SP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Programme Expenditure (2008-2012)</td>
</tr>
<tr>
<td>Total Programme Budget (2008-2012)</td>
</tr>
<tr>
<td>Efficiency of budget utilization</td>
</tr>
<tr>
<td>Cost per DALY saved for current results</td>
</tr>
<tr>
<td>Cost per DALY saved for pessimistic targets</td>
</tr>
<tr>
<td>Cost per DALY saved for optimistic targets</td>
</tr>
<tr>
<td>Technical efficiency of cost per DALY saved against pessimistic targets</td>
</tr>
<tr>
<td>Technical efficiency of cost per DALY saved against optimistic targets</td>
</tr>
</tbody>
</table>

Source: Author’s calculations

It can be seen that from June 2008 to December 2012, HSSP-SP obtained a total disbursement of 16.24 million US Dollars against the total approved budget of 21.00 million US Dollars achieving an efficiency of 77.33% for budget utilization. Overall HSSP-SP had $363.90 cost per DALY saved. But the cost per DALY saved would be reduced to $195.14 if HSSP-SP could achieve its pessimistic targets. This cost would be further reduced to $106.50 if HSSP-SP could achieve its optimistic targets. This gives the intervention a technical efficiency of 53.62% against its pessimistic targets and 29.27% against its optimistic targets.

8.4 Cost effectiveness and technical efficiency of the NAP

8.4.1 Methods

Sources of data:
Both the GBD database of the Institute of Health Metrics and Evaluation and the GHE database of the WHO did not provide any estimates of HIV/AIDS related DALYs in Timor-Leste. This was probably because of the lack of reliable data, small population size and the low prevalence rate of HIV/AIDS in Timor-Leste. This chapter, therefore,
estimated the number of deaths and new infections averted by the NAP of Timor-Leste by synthesizing available information and using epidemiological modeling tools such as Modes of Transmission (UNAIDS, 2012) and Spectrum (Futures Institute, 2014a).

Along with the project related documents such as Grant Agreement, Work Plan and Budget, Performance Framework, Grant Performance Report, and Financial Reports, this chapter used information from a number of surveys and reviews including the review of the NAP (Chan et al., 2013), epidemiological modeling exercise conducted by AusAID (HEMI Study Team, 2006), HIV and behavioral survey conducted by Pissany and team (2004), behavioral surveillance survey conducted by the University of New South Wales (2008), integrated biological and behavioral surveillance survey conducted by the MoH in 2010 (MoH, 2010) and 2011 (MoH, 2011d), national demographic and health survey of Timor-Leste for 2009-2010 (National Statistics Directorate of Timor-Leste, 2010), and Global AIDS Progress Report for Timor-Leste (UNAIDS, 2012b). In addition to these sources, data from the GBD database of the Institute of Health Metrics and Evaluation and the GHE database of the WHO were used to estimate the DALYs related to sexually transmitted infections (STIs) in Timor-Leste.

**Estimating HIV/AIDS related infections and deaths averted:**
In order to estimate the number of HIV infections and AIDS related deaths averted by the NAP and compare them with the likely scenario without the intervention and the scenario if the full targets of the intervention were achieved, this chapter used UNAIDS endorsed epidemiological modeling tools-- the Modes of Transmission Tool (UNAIDS, 2012) and Spectrum (Futures Institute, 2014a).

The Modes of Transmission (MoT) is a Visual Basic enabled Microsoft Excel based spreadsheet that helps to calculate the expected number of new HIV infections per year. MoT uses a mathematical modeling on the basis of a description of the current distribution of HIV and STI infections, at risk sub-populations, and patterns of risk behaviours within and across the sub-populations including the number of sex acts, use of condoms, and ratio of male circumcision.
Downloaded from the UNAIDS website*, the Mode of Transmission tool was used to estimate the number of HIV infections in 2005, and 2006 in Timor-Leste. This estimation used the 2005 baseline data that was used in the HIV Epidemiological Modeling and Impact (HEMI) study (HEMI Study Team, 2006) for Timor-Leste and calculated HIV prevalence in 2006 and 2007 from the estimated new infections. For calculating HIV/AIDS prevalence in 2006 and 2007, it was assumed that 20% of people living with HIV/AIDS would be AIDS cases and 50% of the previous year’s AIDS cases without Anti-retroviral Treatment (ART) would die. It was also assumed that 5% of the high risk sub-populations such as female sex workers (FSW), clients of female sex workers (CSW), and men who have sex with men (MSM) aged between 15 to 49 years would change their risk groups due to graduation from the age group (80% of who changed the risk group) or due to the change of their risk behavior (20% of who changed the risk group) to be low risk hetero-sexual.

HIV prevalence estimates for different risk groups derived from the HEMI (2006) baseline data for 2005 and from the MoT exercise mentioned above for 2006 and 2007 were used for epidemiological curve fitting and calibrating the Spectrum software for the modeling exercise. Estimations derived for 2007 were used as the baseline in the modeling exercise with the Spectrum software.

Spectrum is Future Institute's analytical tool to support analysis of impact and costing of health interventions along with underlying demographics to support decision-making process. Available at the Future Institute’s website†, Spectrum software has several modules including DemProj (Demography), FamPlan (Family Planning), LiST (Lives Saved Tool for Child Survival), AIM (AIDS Impact Model), Goals (Cost and impact of HIV Intervention), Resource Needs Module (Costs of implementing a national HIV/AIDS programme), and RAPID (Resources for the Awareness of Population Impacts on Development).

---

† http://www.futuresinstitute.org/spectrum.aspx
For the purpose of analyzing the impact of the NAP, the DemProj, AIM, and Goals modules of the Spectrum software (version 5.07 released on 14 August 2014) were activated. This version of Spectrum software used the Country Data Pack for Timor-Leste as released on 13 December 2013 on the Future Institute’s website. However, the country data pack for Timor-Leste only included demographic projections for Timor-Leste based on the United Nations Population Division’s latest demographic projections (United Nations Department of Economic and Social Affairs Population Division, 2013) and did not have any country data for the AIM and Goals modules. Therefore, for the AIM and Goals modules the researcher had to insert all relevant available data and use triangulation, interpolation and best estimates where data were not available.

Both the Modes of Transmission (MoT) tool and the Goals module of the Spectrum software assign the 15-49 years old male and female population into different risk groups such as low, medium and high risk heterosexual; men who have sex with men; and injection drug users and consider the extent of risk behavior interactions within and across different risk groups.

The Goals Module uses epidemiological, surveillance, and programmatic information and helps efforts to respond to the HIV/AIDS epidemic by showing how the amount and allocation of funding is related to the achievement of national goals, such as reduction of HIV prevalence and expansion of care and support. Along with the population data projected by the UN Population Division (United Nations Department of Economic and Social Affairs Population Division, 2013), the modeling exercise with the Spectrum software used data from multiple surveys and reports mentioned earlier to triangulate and synthesize the information and develop an informed and best-guess scenario for the HIV/AIDS projections in Timor-Leste. The results and discussions section of this chapter provide detailed assumptions, which were derived from the synthesis of these surveys and reports used in the epidemiological modeling exercise.

Unlike the HIV prevalence, risk behaviour, and other programmatic assumptions, which were derived from the available survey data, the average default values of the Goal Module for other key parameters and impact matrix were used for the modeling exercise. As explained in the Spectrum Manual (Futures Institute, 2014f), these default
values are based on available research findings. Table 8.5 presents the key parameters used in the epidemiological modeling exercise with Spectrum:

**Table 8.5: Key parameters of epidemiological modeling used in Spectrum for Timor-Leste**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population projections</td>
<td>UN Population Projection 2012</td>
</tr>
<tr>
<td>Transmission of HIV per act (female to male)</td>
<td>0.0011</td>
</tr>
<tr>
<td>Transmission multiplier for male to female</td>
<td>1</td>
</tr>
<tr>
<td>Transmission multiplier for STI</td>
<td>4</td>
</tr>
<tr>
<td>Transmission multiplier for MSM contacts</td>
<td>2.6</td>
</tr>
<tr>
<td>Relative infectiousness during Primary Infection</td>
<td>8</td>
</tr>
<tr>
<td>Months in Primary Stage</td>
<td>3</td>
</tr>
<tr>
<td>Relative infectiousness during Asymptomatic stage</td>
<td>1</td>
</tr>
<tr>
<td>Relative infectiousness during Symptomatic stage (no ART)</td>
<td>4</td>
</tr>
<tr>
<td>Relative infectiousness during Symptomatic stage (with ART)</td>
<td>0.6</td>
</tr>
<tr>
<td>Transmission reduction (0 - 100%)</td>
<td></td>
</tr>
<tr>
<td>Condom efficacy for transmission reduction (%)</td>
<td>80</td>
</tr>
<tr>
<td>Reduction in male susceptibility when circumcised (%)</td>
<td>60</td>
</tr>
<tr>
<td>Reduction in male infectiousness when circumcised</td>
<td>0</td>
</tr>
<tr>
<td>Size of initial pulse of infection (0 - 0.01)</td>
<td>0.00012</td>
</tr>
<tr>
<td>Epidemic start year</td>
<td>1990</td>
</tr>
</tbody>
</table>

Source: (Futures Institute, 2014f), (UNAIDS, 2012) and Author’s analysis

Table 8.6 presents the average default values for impact matrix used in the Goals module:

**Table 8.6: Impact matrix for programme coverage**

<table>
<thead>
<tr>
<th>Programme components</th>
<th>Reduction in condom non-use</th>
<th>Reduction in number of partners</th>
<th>Increase in age at first sex</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High risk heterosexual</td>
<td>Medium risk heterosexual</td>
<td>Low risk heterosexual</td>
</tr>
<tr>
<td>Community mobilization</td>
<td>0</td>
<td>-18.6</td>
<td>0</td>
</tr>
<tr>
<td>Mass media</td>
<td>0</td>
<td>-11.5</td>
<td>-17</td>
</tr>
<tr>
<td>VCT</td>
<td>-34</td>
<td>-16</td>
<td>-18.3</td>
</tr>
<tr>
<td>Condoms</td>
<td>-76</td>
<td>-18.9</td>
<td>-24.2</td>
</tr>
<tr>
<td>Youth: in-school</td>
<td>0</td>
<td>-13.4</td>
<td>0</td>
</tr>
<tr>
<td>Youth: out-of-school</td>
<td>0</td>
<td>-33.3</td>
<td>0</td>
</tr>
<tr>
<td>Workplace programmes</td>
<td>0</td>
<td>-23.2</td>
<td>-22.6</td>
</tr>
<tr>
<td>Sex workers</td>
<td>-36.5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>MSM: outreach</td>
<td>0</td>
<td>0</td>
<td>-25.3</td>
</tr>
<tr>
<td>MSM: lubricants</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Source: Spectrum default values (Futures Institute, 2014a)
Modeling counterfactual and optimum scenarios:
With the use of AIM and Goal modules, the Spectrum software examined the HIV/AIDS projection scenarios in Timor-Leste for the years from 2007 to 2013 for the intervention results as well as for a hypothetical “no-intervention” (counterfactual) scenario and a hypothetical “best case” (optimum) scenario with the assumption that all programmatic targets had been achieved by the NAP by using 100% of allocated resources. In order to compare the impact of the NAP with a “no-intervention” (counterfactual) scenario, it was assumed that treatment and prevention services coverage would remain at the baseline (2007) level for the years from 2007 to 2013.

The AIM and Goal modules of the Spectrum software produced projections for HIV prevalence, annual new infections, number of people living with HIV/AIDS in different risk groups, and estimated number of deaths due to HIV/AIDS each year for the programme outcomes, hypothetical counterfactual scenario and the hypothetical optimum scenario. The modules also calculated the number of HIV cases and number of HIV/AIDS related deaths averted each year by the NAP and the projected number of HIV/AIDS cases and HIV/AIDS related deaths expected to be averted in the best-case scenario.

Calculating DALYs averted by the NAP:
In order to calculate DALYs averted by the NAP, the Years of Life Lost (YLL) and the Years Lived in Disability (YLD) were calculated based on the results from the modeling exercise with the Spectrum software for the number of cases and deaths averted by the NAP. Being consistent with the DALY methods followed by the GHE (WHO, 2014), YLLs were calculated by multiplying the number of deaths (assuming a median age group of age at death as 35-39) by the standard loss function (54.67 years) as used in GHE (WHO, 2013a).

The YLDs were calculated as the product of number of cases averted and the disability weight (DW) for a particular disease sequelae (Vos et al., 2012). Following the assumptions used in the Spectrum software and the Global Health Estimates (WHO, 2014), it was assumed that a person infected with HIV (assumed at the median age of 26.6) would live in pre-AIDS asymptomatic sequelae for 8 years with a DW of 0.051; and
a person with AIDS without anti-retroviral treatment (ART) would live in that state for 2 years with a DW of 0.547 before dying. Following the methods suggested by Strive Learning Lab (Terris-Prestholt, 2013), three different outcomes for the people living with AIDS on anti-retroviral treatment (ART) were considered: 1) a person on ART would live for 16.5 years beyond the pre-AIDS asymptomatic period; 2) a person on ART would continue to live up to the full life expectancy beyond the pre-AIDS asymptomatic period; and 3) a person on ART would live for 11 years beyond the pre-AIDS asymptomatic period. In each case the disability weight for AIDS with ART would be 0.053. It was assumed that 40% of the total averted cases would be on anti-retroviral treatment, and half of the cases on ART would live up to full life expectancy beyond the pre-AIDS asymptomatic period, while one-quarter of the cases on ART would live 16.5 years and one-quarter would live 11 years beyond the pre-AIDS asymptomatic period.

Since the NAP had prevention, treatment and care, and sexually transmitted infection (STI) programme components, DALYs averted by each programme component were considered. DALYs averted by the prevention components of the programme were calculated by adding the DALYs averted by number of deaths prevented and DALYs averted by number of cases prevented with all sequelae such as sequelae without ART (60%), sequelae with ART for 16.5 years (10%), sequelae with ART for life-long period (20%), and sequelae with ART for 11 years (10%).

DALYs averted by the treatment components (PMTCT and ART) of the programme were calculated by the AIM module of the Spectrum software for the current results against the counterfactual and optimum targets scenario.

In order to calculate DALYs averted by the STI components of the programme, the exercise used the GHE data for sexually transmitted diseases (STD) excluding HIV/AIDS for Timor-Leste for 2012 and compared them with STD related DALYs for Timor-Leste for the year 2005, as derived from the GBD data adjusted for the GHE methods. For this exercise, 50% of the reduction of STD DALYs was attributed to the NAP. This was based on the opinion presented in the NAP review (Chan et al., 2013). Assuming 50% efficiency of the NAP for reducing STD related DALYs, it was estimated that for the
optimum scenario the estimated DALYs saved by NAP for STD were expected to be double.

Finally, total DALYs saved by the NAP were calculated by adding DALYs saved by prevention components, DALYs saved by treatment components (ART and PMTCT), and DALYs saved by STI programme components. This calculation was conducted both for the current results and optimum scenarios.

**Calculating cost effectiveness ratio and technical efficiency:**
Considering 10% direct contribution by the government to the NAP, 90% of the total DALYs saved by the NAP was attributed to the Global Fund funding. The cost per DALY saved by the NAP activities for the period from June 2007 to December 2013 was then calculated by dividing the total expenditure of the Global Fund’s HIV/AIDS grant for that period by the total DALYs saved attributed to the Global Fund grant for that period.

Similarly the expected cost per DALY saved in the optimum scenario was calculated by dividing the total budget of the Global Fund’s HIV/AIDS grant for the same period by the attributed total DALYs expected to be saved by the NAP if the Programme achieved all the targets.
8.4.2 Results for the NAP

Table 8.7 presents an overview of the key parameters such as evaluation period, goal, objectives, total input value, reported outputs, and reported outcomes of the NAP:

Table 8.7: Overview of design, outputs, and outcomes of the NAP

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Overview</th>
</tr>
</thead>
</table>
| **Evaluation time period** | Round 5: 1 June 2007 to 31 December 2011  
Round 10 Phase 1: 1 January 2012 to 31 December 2013 |
| **Goal**            | To reduce STI and HIV/AIDS mortality and morbidity in Timor-Leste by enhancing related prevention and treatment services                                                                                       |
| **Objectives**      | 1) Improve and increase coverage of strategic HIV and STI prevention services; 2) Strengthen and expand HIV-AIDS treatment and care; 3) Strengthen health systems relating to STIs and HIV; 4) Strengthen community systems relating to STIs and HIV; 5) Build enabling environment for implementation of STI and HIV programmes   |
| **Inputs**          | Total funding: $14.4 million; Estimated expenditure: $11.84 million                                                                                                                                         |
| **Reported outputs**| 1329 MSM & 1134 FSW reached through outreach and prevention services; around 450,000 condoms distributed per year; 31,434 cases of VCT performed; 35 HIV+ women received prophylaxis to avoid PTCT of HIV; 112,579 STI cases treated; 100% of blood units (around 1,800 units) screened for HIV; HIV and STI reporting system established in 100% (13 districts and 66 health facilities) health facilities up to sub-district level; 140 people currently receiving ART; 47% (172/367) of PLHIV who know their HIV status reached with care & support services; 265 health professionals trained on HIV and STI; 10 NGOs received capacity building support; 5 different sectors developed HIV/AIDS related work plan and budget |
| **Reported outcomes** | Outreach coverage increased from 5% to 94% for self-identified MSM and from 5% to 87% for FSW population; 47% ART coverage achieved; 61% (43/71) health facilities prepared to provide STI services as per national guidelines; HIV prevalence reduced from 3% (2003) to 2.76% (2010) to 2% (2012) among FSW; HIV prevalence contained from 1% (2003) to 1.33% (2010) to 1% (2012) among MSM; 63% FSW and 67% MSM reported condom use with most recent partner; 53% uniformed personnel reported condom use with non-regular partner |

Source: Author’s review of programme planning, budget, performance framework, results and expenditure documents of the projects

It appeared from the available studies, programmatic results and reviews (P.-L. Chan et al., 2013; HEMI Study Team, 2006; Ministry of Health of Timor-Leste, 2010, 2011; Pissany & Dili survey team, 2004; University of New South Wales, 2008) that starting from as early as 1990, Timor-Leste had been experiencing a low-level concentrated epidemic where sexual interactions among most at risk groups such as men who have sex with men (MSM), female sex workers (FSW), clients of female sex workers (CSW), female partners of MSM and regular partners of CSW would result in maximum transmission of HIV. The exercise with UNAIDS Modes of Transmission (MoT) tool also found that the maximum transmission of HIV occurred among the MSM group in Timor-Leste in 2005 and 2006. Table 8.8 presents the estimated HIV prevalence among different risk groups as obtained from the HEMI (2006) study MoT exercise:
Table 8.8: HIV prevalence estimates in Timor-Leste, 2005-2007

<table>
<thead>
<tr>
<th></th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIV prevalence total population</td>
<td>0.00013672</td>
<td>0.000241302</td>
<td>0.000394384</td>
</tr>
<tr>
<td>HIV prevalence FSW</td>
<td>0.0317</td>
<td>0.027540535</td>
<td>0.028224687</td>
</tr>
<tr>
<td>HIV prevalence Clients of FSW</td>
<td>0.0004</td>
<td>0.001574744</td>
<td>0.002558903</td>
</tr>
<tr>
<td>HIV prevalence regular and casual partners of FSW</td>
<td>0.0004</td>
<td>0.001976583</td>
<td>0.003057552</td>
</tr>
<tr>
<td>HIV prevalence regular and casual partner of clients of FSW</td>
<td>0.0003</td>
<td>0.00310406</td>
<td>0.000411638</td>
</tr>
<tr>
<td>HIV prevalence MSM</td>
<td>0.0105</td>
<td>0.016809346</td>
<td>0.026875313</td>
</tr>
<tr>
<td>HIV prevalence female partners of MSM</td>
<td>0.0003</td>
<td>0.000957555</td>
<td>0.002005188</td>
</tr>
</tbody>
</table>

Source: HEMI 2006 estimation for 2005 and projections based on MoT exercise

As mentioned earlier, these estimates were used in the Spectrum software for calibrating the model for Timor-Leste and for setting up the baseline data for 2007. The following sections provide details of assumptions, estimations and results from the Spectrum exercise.

Size estimates:

Table 8.9 presents the size estimation of different risk sub-populations in Timor-Leste from 2007 to 2013 as projected by the Spectrum software:

Table 8.9: Estimated size of risk sub-populations in Timor-Leste, 2007-2013

<table>
<thead>
<tr>
<th></th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total adult population (15-49 male and female)</td>
<td>431,132</td>
<td>440,530</td>
<td>452,880</td>
<td>467,927</td>
<td>483,815</td>
<td>498,192</td>
<td>510,364</td>
</tr>
<tr>
<td>Male- Not sexually active</td>
<td>24,285</td>
<td>25,904</td>
<td>27,997</td>
<td>30,074</td>
<td>31,416</td>
<td>31,599</td>
<td>30,847</td>
</tr>
<tr>
<td>Male- Low risk heterosexual</td>
<td>118,048</td>
<td>119,556</td>
<td>121,687</td>
<td>124,714</td>
<td>128,578</td>
<td>132,892</td>
<td>137,365</td>
</tr>
<tr>
<td>Male- Medium risk heterosexual</td>
<td>46,158</td>
<td>47,034</td>
<td>48,262</td>
<td>49,921</td>
<td>51,882</td>
<td>53,847</td>
<td>55,606</td>
</tr>
<tr>
<td>Male- High risk heterosexual</td>
<td>21,862</td>
<td>22,280</td>
<td>22,865</td>
<td>23,657</td>
<td>24,595</td>
<td>25,536</td>
<td>26,378</td>
</tr>
<tr>
<td>Male- Men who have sex with men (MSM)</td>
<td>5,117</td>
<td>5,234</td>
<td>5,397</td>
<td>5,619</td>
<td>5,888</td>
<td>6,176</td>
<td>6,448</td>
</tr>
<tr>
<td>Total Male adult population (15-49)</td>
<td>215,471</td>
<td>220,007</td>
<td>226,209</td>
<td>233,985</td>
<td>242,359</td>
<td>250,050</td>
<td>256,645</td>
</tr>
<tr>
<td>Female- Not sexually active</td>
<td>23,739</td>
<td>25,263</td>
<td>27,210</td>
<td>29,081</td>
<td>30,230</td>
<td>30,308</td>
<td>29,534</td>
</tr>
<tr>
<td>Female- Low risk heterosexual</td>
<td>127,388</td>
<td>129,317</td>
<td>131,731</td>
<td>134,873</td>
<td>138,684</td>
<td>142,813</td>
<td>147,012</td>
</tr>
<tr>
<td>Female- Medium risk heterosexual</td>
<td>63,498</td>
<td>64,884</td>
<td>66,643</td>
<td>68,864</td>
<td>71,378</td>
<td>73,816</td>
<td>75,933</td>
</tr>
<tr>
<td>Female- High risk heterosexual</td>
<td>1,037</td>
<td>1,059</td>
<td>1,088</td>
<td>1,124</td>
<td>1,165</td>
<td>1,205</td>
<td>1,240</td>
</tr>
<tr>
<td>Total Female adult population</td>
<td>215,661</td>
<td>220,523</td>
<td>226,672</td>
<td>233,942</td>
<td>241,457</td>
<td>248,142</td>
<td>253,719</td>
</tr>
</tbody>
</table>

Source: Modeling exercise with Spectrum
In the Spectrum exercise, it was estimated that 17% of the male and 21% of the female population aged 15-49 years were not sexually active. This estimation is consistent with Timor-Leste’s Demographic and Health Survey 2009-10 (National Statistics Directorate of Timor-Leste, 2010) that found 17.6% of women and 14% of men in the 20-49 age group had never had sexual intercourse.

In the modeling exercise, the clients of female sex workers were considered as high-risk hetero-sexual male population. Pissany and Dili Survey Team (2004) estimated 48% of taxi drivers and people in uniform, and the HEMI (2006) projection estimated 28% of the adult male population as clients of sex workers. However, being consistent with Timor-Leste’s Demographic and Health Survey (2010) findings that found 10.7% of 20-24 year old males paid for sex in the last 12 months, the Spectrum exercise estimated around 10.4% of the 15-49 year old male population as clients of sex workers. This percentage was validated by calculating and balancing the total number of high risk male-female sex acts based on the average number of partners and annual number of sex acts per partner as found in the behavioral surveillance surveys (2008, 2010 and 2011) for female sex workers and clients of sex workers.

Likewise, the female sex workers were considered as a high-risk female sub-population. The size of the female sex worker sub-population was estimated as only 450 in 2005 in the HEMI (2006) projection. However, the programmatic coverage data of the NAP showed that 1,134 female sex workers were reached by the Programme in 2011 and findings from Integrated Biological and Behavioral Surveillance (Ministry of Health of Timor-Leste, 2011) indicated that 87.22% of female sex workers were reached by the NAP. This gave a total number of female sex workers sub-population of approximately 1,300 in 2011. In the Spectrum exercise this number was adjusted considering a possible effect of over-reporting by outreach NGO. The Spectrum software was also allowed to balance and interpolate the number of female sex workers for each year based on other programme parameters. With those balancing, female sex workers were estimated to be 0.5% of the 15-49 years old female population in Timor-Leste.
The HEMI (2006) projection used an estimate of 5% of the urban male population and 2.5% of the rural male population as MSM without any definitive basis for these assumptions in Timor-Leste. In order to estimate the sub-population of MSM, this Spectrum exercise considered the programmatic coverage in 2011 from the programmatic data of the NAP and coverage data from IBBS 2011. These data showed 1329 MSM were reached in 2011, which was 94% of the MSM sub-population. However, the findings from the behavioral surveillance (University of New South Wales, 2008) indicated that the outreach programme mostly reached the receptive type of MSM who identified themselves as MSM. According the BSS 2008 (University of New South Wales, 2008), the receptive type of MSM represents only 29.4% of the MSM sub-population, whereas the insertive male partners of MSM, which was 64.5% of the MSM sub-population, were often bi-sexual having both regular and casual female partners. Since the MSM sub-population was considered as only one group in this Spectrum projection exercise, it attempted to estimate the size of the broader MSM sub-population including receptive, insertive, male sex workers and other types of MSM from the available data and estimates. This gave an estimate of 5.8% of 15-49 years old male population as MSM. This estimate of the MSM sub-population closely matches with the estimate done by Caceres et al. (2006) that estimated 6% to 12% of the adult male population as MSM in south and south-east Asia.

The hetero-sexual male and female who had more than one non-paid casual sexual partner were considered medium risk sub populations and their percentages (21.4% male and 30.6% female) were estimated from the Demographic and Health Survey (2010). The rest of the population were considered low-risk hetero-sexual with one steady partner as assumed in the HEMI (2006) projection.

Both the Pissany (2004) survey and the HEMI (2006) projection did not consider any injection drug user (IDU) sub-population in Timor-Leste. Although the subsequent behavioral surveillance (2008, 2010 and 2011) found some self reported injection drug use behavior among female sex workers and MSM sub-population, their number was very low. Since these people were covered under female sex worker and MSM sub-populations anyway and since the NAP did not specifically have any programme
component targeting the injection drug user, this Spectrum exercise did not consider any IDU sub-population in its projections.

**STI prevalence:**
Table 8.10 presents the estimates used in the Spectrum exercise for the STI prevalence for the intervention, no-intervention, and target scenarios:

<table>
<thead>
<tr>
<th></th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Male</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low risk hetero-sexual</td>
<td>Intervention</td>
<td>10%</td>
<td>10%</td>
<td>9%</td>
<td>7%</td>
<td>7%</td>
<td>6%</td>
</tr>
<tr>
<td></td>
<td>Targets</td>
<td>8%</td>
<td>5%</td>
<td>4%</td>
<td>3%</td>
<td>3%</td>
<td>2%</td>
</tr>
<tr>
<td></td>
<td>No-intervention</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
</tr>
<tr>
<td>Medium risk hetero-sexual</td>
<td>Intervention</td>
<td>16%</td>
<td>16%</td>
<td>15%</td>
<td>10%</td>
<td>10%</td>
<td>9%</td>
</tr>
<tr>
<td></td>
<td>Targets</td>
<td>12%</td>
<td>12%</td>
<td>8%</td>
<td>5%</td>
<td>5%</td>
<td>4%</td>
</tr>
<tr>
<td></td>
<td>No-intervention</td>
<td>16%</td>
<td>16%</td>
<td>16%</td>
<td>16%</td>
<td>16%</td>
<td>16%</td>
</tr>
<tr>
<td>High risk hetero-sexual</td>
<td>Intervention</td>
<td>16%</td>
<td>16%</td>
<td>15%</td>
<td>10%</td>
<td>10%</td>
<td>9%</td>
</tr>
<tr>
<td></td>
<td>Targets</td>
<td>12%</td>
<td>10%</td>
<td>8%</td>
<td>5%</td>
<td>5%</td>
<td>4%</td>
</tr>
<tr>
<td></td>
<td>No-intervention</td>
<td>16%</td>
<td>16%</td>
<td>16%</td>
<td>16%</td>
<td>16%</td>
<td>16%</td>
</tr>
<tr>
<td>MSM</td>
<td>Intervention</td>
<td>18%</td>
<td>18%</td>
<td>16%</td>
<td>12%</td>
<td>10%</td>
<td>9%</td>
</tr>
<tr>
<td></td>
<td>Targets</td>
<td>14%</td>
<td>12%</td>
<td>10%</td>
<td>9%</td>
<td>6%</td>
<td>4%</td>
</tr>
<tr>
<td></td>
<td>No-intervention</td>
<td>18%</td>
<td>18%</td>
<td>18%</td>
<td>18%</td>
<td>18%</td>
<td>18%</td>
</tr>
<tr>
<td><strong>Female</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low risk hetero-sexual</td>
<td>Intervention</td>
<td>10%</td>
<td>10%</td>
<td>9%</td>
<td>7%</td>
<td>7%</td>
<td>6%</td>
</tr>
<tr>
<td></td>
<td>Targets</td>
<td>8%</td>
<td>7%</td>
<td>6%</td>
<td>3%</td>
<td>3%</td>
<td>3%</td>
</tr>
<tr>
<td></td>
<td>No-intervention</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
</tr>
<tr>
<td>Medium risk hetero-sexual</td>
<td>Intervention</td>
<td>16%</td>
<td>16%</td>
<td>15%</td>
<td>10%</td>
<td>10%</td>
<td>9%</td>
</tr>
<tr>
<td></td>
<td>Targets</td>
<td>12%</td>
<td>8%</td>
<td>7%</td>
<td>5%</td>
<td>5%</td>
<td>4%</td>
</tr>
<tr>
<td></td>
<td>No-intervention</td>
<td>16%</td>
<td>16%</td>
<td>16%</td>
<td>16%</td>
<td>16%</td>
<td>16%</td>
</tr>
<tr>
<td>High risk hetero-sexual</td>
<td>Intervention</td>
<td>16%</td>
<td>16%</td>
<td>15%</td>
<td>10%</td>
<td>10%</td>
<td>9%</td>
</tr>
<tr>
<td></td>
<td>Targets</td>
<td>12%</td>
<td>8%</td>
<td>7%</td>
<td>5%</td>
<td>5%</td>
<td>4%</td>
</tr>
<tr>
<td></td>
<td>No-intervention</td>
<td>16%</td>
<td>16%</td>
<td>16%</td>
<td>16%</td>
<td>16%</td>
<td>16%</td>
</tr>
</tbody>
</table>

Source: Author’s estimations

These estimates considered the fact that a high level of STI prevalence in Timor-Leste, especially among most at risk groups, had been documented in the Pissany (2004) survey and Integrated Biological and Behavioral Surveys (2010, 2011). It also considered the assumption that almost 50% of the STI cases might be asymptomatic and remain untreated (Chan et al., 2013).

**Number of partners:**
Based on the data from behavioral surveillance surveys (Ministry of Health of Timor-Leste, 2010, 2011; Pissany & Dili survey team, 2004; University of New South Wales, 2008) and Demographic and Health Survey (National Statistics Directorate of Timor-
Leste, 2010), the following estimates were used for number of partners for each risk group in the actual, optimum and counterfactual scenarios:

Table 8.11: Estimated number of sexual partners for HIV risk groups in Timor-Leste, 2007-2013

<table>
<thead>
<tr>
<th></th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Male Population</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low risk hetero-sexual</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actual</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Optimum</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Counterfactual</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Medium risk hetero-sexual</td>
<td>2.4</td>
<td>2.39</td>
<td>2.38</td>
<td>2.37</td>
<td>2.36</td>
<td>2.35</td>
<td>2.34</td>
</tr>
<tr>
<td>Optimum</td>
<td>2.4</td>
<td>2.37</td>
<td>2.34</td>
<td>2.31</td>
<td>2.28</td>
<td>2.25</td>
<td>2.22</td>
</tr>
<tr>
<td>Counterfactual</td>
<td>2.4</td>
<td>2.4</td>
<td>2.4</td>
<td>2.4</td>
<td>2.4</td>
<td>2.4</td>
<td>2.4</td>
</tr>
<tr>
<td>High risk hetero-sexual</td>
<td>3.8</td>
<td>3.58</td>
<td>3.37</td>
<td>3.16</td>
<td>2.96</td>
<td>2.76</td>
<td>2.57</td>
</tr>
<tr>
<td>Optimum</td>
<td>3.8</td>
<td>3.50</td>
<td>3.21</td>
<td>2.93</td>
<td>2.66</td>
<td>2.41</td>
<td>2.16</td>
</tr>
<tr>
<td>Counterfactual</td>
<td>3.8</td>
<td>3.8</td>
<td>3.8</td>
<td>3.8</td>
<td>3.8</td>
<td>3.8</td>
<td>3.8</td>
</tr>
<tr>
<td>MSM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actual</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Optimum</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Counterfactual</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td><strong>Female Population</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low risk hetero-sexual</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actual</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Optimum</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Counterfactual</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Medium risk hetero-sexual</td>
<td>2.2</td>
<td>2.19</td>
<td>2.18</td>
<td>2.17</td>
<td>2.16</td>
<td>2.16</td>
<td>2.15</td>
</tr>
<tr>
<td>Optimum</td>
<td>2.2</td>
<td>2.17</td>
<td>2.14</td>
<td>2.12</td>
<td>2.09</td>
<td>2.06</td>
<td>2.03</td>
</tr>
<tr>
<td>Counterfactual</td>
<td>2.2</td>
<td>2.2</td>
<td>2.2</td>
<td>2.2</td>
<td>2.2</td>
<td>2.2</td>
<td>2.2</td>
</tr>
<tr>
<td>High risk hetero-sexual</td>
<td>160</td>
<td>150.8</td>
<td>141.9</td>
<td>133.1</td>
<td>124.6</td>
<td>116.3</td>
<td>108.2</td>
</tr>
<tr>
<td>Optimum</td>
<td>160</td>
<td>147.4</td>
<td>135.2</td>
<td>123.5</td>
<td>112.2</td>
<td>101.4</td>
<td>91.0</td>
</tr>
<tr>
<td>Counterfactual</td>
<td>160</td>
<td>160</td>
<td>160</td>
<td>160</td>
<td>160</td>
<td>160</td>
<td>160</td>
</tr>
</tbody>
</table>

Source: Behavioural Surveillance findings and modeling exercise with Spectrum

It can be seen that while there is no change for the number of partners between the actual, optimum, and counterfactual scenarios for the low-risk hetero-sexual male and female and MSM groups, for other groups there is a decrease in the number of partners over the years. This is due to the projected impact of the behaviour change intervention resulting in a slight reduction in the number of partners over the years. For the counterfactual (no-intervention) scenario, there is no change in the number of partners from the baseline (2007) year data.

**Number of sex acts per partner and condom use:**

The following are the estimates used in the Spectrum projection exercise for number of sex acts per partner and percentage of total sex acts covered by condom use in the actual, optimum, and counterfactual scenarios:
Table 8.12: Sex acts per partner and condom use by HIV risk groups in Timor-Leste, 2007-2013

<table>
<thead>
<tr>
<th></th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number of sex acts per partner</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Low risk heterosexual</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actual</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Optimum</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Counterfactual</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td><strong>Medium risk heterosexual</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actual</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>Optimum</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>Counterfactual</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td><strong>High risk heterosexual</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actual</td>
<td>4.85</td>
<td>5.19</td>
<td>5.07</td>
<td>4.81</td>
<td>4.73</td>
<td>4.73</td>
<td>4.73</td>
</tr>
<tr>
<td>Optimum</td>
<td>4.85</td>
<td>5.19</td>
<td>5.07</td>
<td>4.81</td>
<td>4.73</td>
<td>4.73</td>
<td>4.73</td>
</tr>
<tr>
<td>Counterfactual</td>
<td>3.8</td>
<td>3.8</td>
<td>3.8</td>
<td>3.8</td>
<td>3.8</td>
<td>3.8</td>
<td>3.8</td>
</tr>
<tr>
<td><strong>MSM</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actual</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Optimum</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Counterfactual</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td><strong>Condom use</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Low risk heterosexual</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actual</td>
<td>0%</td>
<td>1.23%</td>
<td>2.45%</td>
<td>3.65%</td>
<td>4.85%</td>
<td>6.04%</td>
<td>7.22%</td>
</tr>
<tr>
<td>Optimum</td>
<td>0%</td>
<td>3.01%</td>
<td>5.95%</td>
<td>8.84%</td>
<td>11.67%</td>
<td>14.43%</td>
<td>17.15%</td>
</tr>
<tr>
<td>Counterfactual</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td><strong>Medium risk heterosexual</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actual</td>
<td>15%</td>
<td>16.05%</td>
<td>17.09%</td>
<td>18.12%</td>
<td>19.14%</td>
<td>20.15%</td>
<td>21.15%</td>
</tr>
<tr>
<td>Optimum</td>
<td>15%</td>
<td>17.88%</td>
<td>20.60%</td>
<td>23.30%</td>
<td>25.93%</td>
<td>28.49%</td>
<td>30.98%</td>
</tr>
<tr>
<td>Counterfactual</td>
<td>15%</td>
<td>15%</td>
<td>15%</td>
<td>15%</td>
<td>15%</td>
<td>15%</td>
<td>15%</td>
</tr>
<tr>
<td><strong>High risk heterosexual</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actual</td>
<td>12%</td>
<td>17.17%</td>
<td>22.19%</td>
<td>27.04%</td>
<td>31.75%</td>
<td>36.30%</td>
<td>40.71%</td>
</tr>
<tr>
<td>Optimum</td>
<td>12%</td>
<td>20.48%</td>
<td>28.46%</td>
<td>35.95%</td>
<td>42.96%</td>
<td>49.51%</td>
<td>55.60%</td>
</tr>
<tr>
<td>Counterfactual</td>
<td>12%</td>
<td>12%</td>
<td>12%</td>
<td>12%</td>
<td>12%</td>
<td>12%</td>
<td>12%</td>
</tr>
<tr>
<td><strong>MSM</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actual</td>
<td>14%</td>
<td>15.40%</td>
<td>16.80%</td>
<td>18.20%</td>
<td>19.59%</td>
<td>20.99%</td>
<td>22.39%</td>
</tr>
<tr>
<td>Optimum</td>
<td>14%</td>
<td>16.49%</td>
<td>18.97%</td>
<td>21.46%</td>
<td>23.95%</td>
<td>26.43%</td>
<td>28.92%</td>
</tr>
<tr>
<td>Counterfactual</td>
<td>14%</td>
<td>14%</td>
<td>14%</td>
<td>14%</td>
<td>14%</td>
<td>14%</td>
<td>14%</td>
</tr>
</tbody>
</table>

Source: Behavioural Surveillance findings and modeling exercise with Spectrum

Unlike the number of partners that changed over time, it can be seen that there were no changes in the number of sex acts per partner in the actual, optimum, and counterfactual scenario. However, as expected the percentage of condom use increased over time in the actual (intervention results) and optimum (targets) scenario. In the above table, the actual increases in condom use were estimated from the behavioural surveillance data (2004, 2008, 2010, 2011) and DHS 2010 data while targets were estimated from the outcome indicators of the programme’s results framework.

**Programmatic coverage data:**

The following are the programmatic coverage data and estimations for the baseline, results and targets scenarios for the general population, most at risk groups and medical services. The baseline assumed that there would be no change in the coverage from the baseline year of 2007 to the target year of 2014.
Table 8.13: Estimated coverage by the NAP

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Community mobilization: Reached by intervention per year</td>
<td>0%</td>
<td>10%</td>
<td>40%</td>
</tr>
<tr>
<td>Mass media: Reached by campaigns per year</td>
<td>0%</td>
<td>15%</td>
<td>50%</td>
</tr>
<tr>
<td>Adult population receiving VCT each year</td>
<td>0%</td>
<td>20%</td>
<td>30%</td>
</tr>
<tr>
<td>Condom coverage</td>
<td>0%</td>
<td>10%</td>
<td>30%</td>
</tr>
<tr>
<td>Workforce receiving peer education</td>
<td>0%</td>
<td>60%</td>
<td>80%</td>
</tr>
<tr>
<td>Workforce receiving STI treatment</td>
<td>0%</td>
<td>50%</td>
<td>80%</td>
</tr>
<tr>
<td>Most at risk population:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female sex workers reached by intervention per year</td>
<td>0%</td>
<td>75%</td>
<td>100%</td>
</tr>
<tr>
<td>Male sex workers reached by intervention per year</td>
<td>0%</td>
<td>90%</td>
<td>100%</td>
</tr>
<tr>
<td>MSMs reached by intervention per year</td>
<td>0%</td>
<td>45%</td>
<td>80%</td>
</tr>
<tr>
<td>MSMs receiving lubricants</td>
<td>0%</td>
<td>45%</td>
<td>80%</td>
</tr>
<tr>
<td>Medical services:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males with STI receiving treatment</td>
<td>10%</td>
<td>50%</td>
<td>80%</td>
</tr>
<tr>
<td>Females with STI receiving treatment</td>
<td>10%</td>
<td>50%</td>
<td>80%</td>
</tr>
<tr>
<td>Units of blood for transfusion tested</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Post-exposure prophylaxis need that is met</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Unsafe medical injections replaced with AD syringes</td>
<td>0%</td>
<td>98%</td>
<td>100%</td>
</tr>
<tr>
<td>Hospital beds covered for universal precautions</td>
<td>0%</td>
<td>70%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Source: Author’s estimations based on programmatic reports

PMTCT and ART coverage:

The following data were used in the AIM module of the Spectrum software for prevention of mother to child transmission (PMTCT) and anti-retroviral treatment (ART) coverage for the actual (intervention results), optimum (targets) and counterfactual (no-intervention) scenarios:

Table 8.14: ART and PMTCT services coverage in Timor-Leste, 2007-2013

<table>
<thead>
<tr>
<th>Pregnant women receiving PMTCT services</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>15</td>
<td>7</td>
</tr>
<tr>
<td>Optimum</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>7</td>
<td>40</td>
<td>70</td>
</tr>
<tr>
<td>Counterfactual</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Adults receiving ART</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual</td>
<td>8</td>
<td>22</td>
<td>29</td>
<td>36</td>
<td>89</td>
<td>79</td>
<td>129</td>
</tr>
<tr>
<td>Optimum</td>
<td>8</td>
<td>24</td>
<td>25</td>
<td>36</td>
<td>200</td>
<td>201</td>
<td>298</td>
</tr>
<tr>
<td>Counterfactual</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
</tbody>
</table>


It can be seen that PMTCT targets and coverage in Timor-Leste were very low and although targets increased significantly in 2012 and 2013, the actual coverage did not match with the targets indicating difficulties in programme implementation. The ART targets also increased steeply in 2011 from those of the previous years. This was due to the adoption by the country of new WHO guidelines for ART (WHO, 2010) that
suggested a significant increase of CD4 threshold count from less than 200 to less than 350 for anti-retroviral treatment eligibility. However, actual treatment performance could not quite match with the ART targets in the years from 2011 to 2013.

**AIM results:**

The following is the summary of results produced by AIM module of the Spectrum software based on the PMTCT and ART inputs for the counterfactual (no-intervention), actual (intervention results), and optimum (targets) scenarios:

**Table 8.15: Impact of PMTCT and ART programme components in Timor-Leste, 2007-2013**

<table>
<thead>
<tr>
<th></th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Counterfactual scenario:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deaths averted by ART</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td>Infections averted by PMTCT</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Life years gained by ART and PMTCT</td>
<td>6</td>
<td>8</td>
<td>10</td>
<td>12</td>
<td>14</td>
<td>16</td>
<td>19</td>
<td>85</td>
</tr>
<tr>
<td><strong>Actual scenario:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deaths averted by ART</td>
<td>3</td>
<td>2</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>15</td>
<td>16</td>
<td>57</td>
</tr>
<tr>
<td>Infections averted by PMTCT</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Life years gained by ART and PMTCT</td>
<td>6</td>
<td>9</td>
<td>16</td>
<td>24</td>
<td>34</td>
<td>52</td>
<td>72</td>
<td>213</td>
</tr>
<tr>
<td>Incremental life year gained with intervention (Actual – Counterfactual)</td>
<td>0</td>
<td>1</td>
<td>6</td>
<td>12</td>
<td>20</td>
<td>36</td>
<td>53</td>
<td>128</td>
</tr>
<tr>
<td><strong>Optimum scenario:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deaths averted by ART</td>
<td>3</td>
<td>2</td>
<td>7</td>
<td>6</td>
<td>10</td>
<td>27</td>
<td>32</td>
<td>87</td>
</tr>
<tr>
<td>Infections averted by PMTCT</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>Life years gained by ART and PMTCT</td>
<td>6</td>
<td>9</td>
<td>16</td>
<td>24</td>
<td>36</td>
<td>67</td>
<td>105</td>
<td>263</td>
</tr>
<tr>
<td>Incremental life year gained if full targets achieved (Optimum – Counterfactual)</td>
<td>0</td>
<td>1</td>
<td>6</td>
<td>12</td>
<td>22</td>
<td>51</td>
<td>86</td>
<td>178</td>
</tr>
<tr>
<td>Treatment services efficiency</td>
<td>n/a</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>90.9%</td>
<td>70.6%</td>
<td>61.6%</td>
<td>71.9%</td>
</tr>
</tbody>
</table>

Source: Modeling exercise with Spectrum
It can be seen that anti-retroviral treatment had effects on preventing deaths, reducing infectiousness of HIV transmission, and improving quality of life for the people living with HIV/AIDS. From 2007 to 2013, the total incremental life years gained by the ART and PMTCT components of the NAP 128 against the a target situation of 178 total incremental life years gained if the programme could achieve all its targets fully. This gives a 71.9% efficiency of the treatment services in terms of incremental life years gained.

**DALYs saved by ART and PMTCT programme components:**

Total life years gained by the ART and PMTCT programme components were converted into DALYs saved by using the formula:

\[ \text{Total DALY} = \text{Total Life Year Gained} \times (1 - \text{DW}) \]

Where, \( \text{DW} = \text{Disability Weight for AIDS with ART} \), which is 0.053. Table 8.16 provides the calculation of total DALYs saved by the ART and PMTCT programme components in Timor-Leste and DALYs expected to be saved if those components could achieve their full targets:
### Table 8.16: DALYs saved by ART and PMTCT programme components in Timor-Leste

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Total incremental life years gained</th>
<th>Conversion factor (1 - DW)</th>
<th>Total DALYs saved</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual scenario</td>
<td>128</td>
<td>0.947</td>
<td>121.216</td>
</tr>
<tr>
<td>Optimum scenario</td>
<td>178</td>
<td>0.947</td>
<td>168.566</td>
</tr>
</tbody>
</table>

Source: Author’s calculations based on modeling exercise with Spectrum

This calculation gave an estimate of 121.22 DALYs saved by the PMTCT and ART programme components in Timor-Leste from 2007 to 2013. The Programme would have saved 168.6 DALYs if full targets related to PMTCT and ART were achieved by the Programme during this period. This gives an outcome efficiency of 71.9% to the treatment (ART and PMTCT) components of the NAP of Timor-Leste for the period from 2007 to 2013.

### Goal results:

Figure 8.3 provides projections for the number of new infections for the actual (intervention), counterfactual (no-intervention) and optimum (targets) scenarios based on the Goal module exercise:

![New HIV Infections](image)

Figure 8.3: Number of new HIV infections in counterfactual, actual, and optimum scenario

It can be seen that without the NAP, annual new infections for HIV were expected to rise from 236 in 2007 to 713 in 2013. However, the target of the NAP was to contain the
annual new HIV infections within the range of 212 to 257 per year during the period from 2007 to 2013.

Figure 8.4 provides projections for the number of deaths for the actual (intervention), counterfactual (no-intervention) and optimum (targets) scenarios:

![Graph showing AIDS deaths (Male + Female) for different scenarios from 2007 to 2014.]

Figure 8.4: Number of AIDS related deaths in counterfactual, actual, and optimum scenario

It can be seen that in the optimum scenario (if full targets of the programme had been achieved), the number of deaths would have dropped sharply in 2012 from those in 2011. This is because of the planned scaling up of the anti-retroviral treatment programme based on the adopted new guidelines of WHO for ART (WHO, 2010).

Table 8.17 presents the number of new infections and the number of deaths averted by the prevention components of the NAP for the actual and optimum scenarios:
Table 8.17: New infections and deaths averted by NAP, 2007-2013

<table>
<thead>
<tr>
<th></th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>New infections averted by</td>
<td>0</td>
<td>7</td>
<td>42</td>
<td>102</td>
<td>157</td>
<td>227</td>
<td>291</td>
<td>826</td>
</tr>
<tr>
<td>current intervention</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New infections expected</td>
<td>24</td>
<td>65</td>
<td>127</td>
<td>181</td>
<td>260</td>
<td>370</td>
<td>456</td>
<td>1483</td>
</tr>
<tr>
<td>to be averted if full</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>targets achieved</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Efficiency in averting new</td>
<td>0.00%</td>
<td>10.77%</td>
<td>33.07%</td>
<td>56.35%</td>
<td>60.38%</td>
<td>61.35%</td>
<td>63.82%</td>
<td>55.70%</td>
</tr>
<tr>
<td>infections</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deaths averted by current</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>5</td>
<td>8</td>
<td>16</td>
<td>18</td>
<td>52</td>
</tr>
<tr>
<td>intervention</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deaths expected to be averted</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td>6</td>
<td>12</td>
<td>31</td>
<td>40</td>
<td>95</td>
</tr>
<tr>
<td>if full targets achieved</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Efficiency in averting deaths</td>
<td>n/a</td>
<td>100%</td>
<td>80%</td>
<td>83.33%</td>
<td>66.67%</td>
<td>51.61%</td>
<td>45.00%</td>
<td>54.74%</td>
</tr>
</tbody>
</table>

Source: Modeling exercise with Spectrum

It can be seen that the prevention programme components of the NAP of Timor-Leste averted 826 new HIV infections from 2007 to 2013 against the target scenario of averting 1483 new infections. This gives the prevention programme an efficiency of 55.7% in terms of averting new infections. The prevention components of the programme also contributed to averting 52 AIDS related deaths from 2007 to 2013 against the target scenario of averting 95 AIDS related deaths in the same period. This gives the prevention programme an efficiency of 54.74% of averting AIDS related deaths through prevention programming.

**DALYs saved by the prevention programme:**

Table 8.18 presents the calculation of DALYs saved by the prevention components of the NAP. This calculation is based on the total number of deaths and new infections averted by the prevention components of the NAP from 2007 to 2013.

Table 8.18: DALYs saved by prevention components of NAP

<table>
<thead>
<tr>
<th></th>
<th>Calculation parameters</th>
<th>Actual scenario</th>
<th>Optimum scenario</th>
<th>Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Year of Life Lost (YLL)</td>
<td>Age at death: 36.6</td>
<td>2,842.84</td>
<td>5,193.65</td>
<td>54.74%</td>
</tr>
<tr>
<td>(YLL) averted for the deaths</td>
<td>Life expectancy at death: 54.67</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>prevented during the intervention</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B. Without ART:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B1. Averted Years Lived in</td>
<td>Proportion of cases without ART: 60%</td>
<td>202.20</td>
<td>363.04</td>
<td></td>
</tr>
<tr>
<td>Disability (YLD) during Pre-AIDS asymptomatic sequelae without ART</td>
<td>Average duration of Pre-AIDS asymptomatic sequelae: 8 years</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Disability weight (DW) of Pre-AIDS asymptomatic sequelae: 0.051</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B2. Averted YLD for cases</td>
<td>Average duration of AIDS sequelae without ART: 2 years</td>
<td>542.19</td>
<td>973.44</td>
<td></td>
</tr>
<tr>
<td>without ART</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Component</td>
<td>Description</td>
<td>Proportion or Duration</td>
<td>Calculations</td>
<td>DALYs Saved</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------</td>
<td>------------------------</td>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>B3. YLL averted for cases without ART</td>
<td>Age at death: 36.6  Life expectancy at death: 54.67</td>
<td></td>
<td></td>
<td>27,094.45  48,645.37</td>
</tr>
<tr>
<td>B. DALYs averted for cases without ART</td>
<td>B1 + B2 + B3</td>
<td></td>
<td></td>
<td>27,838.84  49,981.85  55.70%</td>
</tr>
<tr>
<td>C. AIDS with ART for 16.5 years:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C1. Averted Years Lived in Disability (YLD) during Pre-AIDS asymptomatic sequelae</td>
<td>Proportion of cases for AIDS with ART for 16.5 years: 10%  Average duration of Pre-AIDS asymptomatic sequelae: 8 years  Disability weight (DW) of Pre-AIDS asymptomatic sequelae: 0.051</td>
<td></td>
<td>33.70  60.51  55.70%</td>
<td></td>
</tr>
<tr>
<td>C2. Averted YLD for AIDS with ART for 16.5 years cases</td>
<td>Average duration of AIDS with ART: 16.5 years  Disability weight for AIDS with ART: 0.053</td>
<td></td>
<td>72.23  129.69</td>
<td></td>
</tr>
<tr>
<td>C3. YLL for AIDS with ART for 16.5 years cases</td>
<td>Age at death: 51.1 years  Life expectancy at death: 39.92 years</td>
<td></td>
<td>3,297.39  5,920.14</td>
<td></td>
</tr>
<tr>
<td>C. DALYs averted for AIDS with ART for 16.5 years</td>
<td>C1 + C2 + C3</td>
<td></td>
<td>3,403.33  6,110.33  55.70%</td>
<td></td>
</tr>
<tr>
<td>D. AIDS with life long ART:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D1. Averted Years Lived in Disability (YLD) during Pre-AIDS asymptomatic sequelae</td>
<td>Proportion of cases for AIDS with life long ART: 20%  Average duration of Pre-AIDS asymptomatic sequelae: 8 years  Disability weight (DW) of Pre-AIDS asymptomatic sequelae: 0.051</td>
<td></td>
<td>67.40  121.01</td>
<td></td>
</tr>
<tr>
<td>D2. Averted YLD for AIDS with life long ART</td>
<td>Average duration of AIDS with life long ART: 54.67 years  Disability weight for AIDS with ART: 0.053</td>
<td></td>
<td>478.67  859.40</td>
<td></td>
</tr>
<tr>
<td>D3. YLL for AIDS with life long ART</td>
<td>Age at death: 92 years  Life expectancy at death: 0 years</td>
<td></td>
<td>0.00  0.00</td>
<td></td>
</tr>
<tr>
<td>D. DALYs averted for AIDS with life long ART</td>
<td>D1 + D2 + D3</td>
<td></td>
<td>546.07  980.41  55.70%</td>
<td></td>
</tr>
<tr>
<td>E. AIDS with ART for 11 years:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E1. Averted Years Lived in Disability (YLD) during Pre-AIDS asymptomatic sequelae</td>
<td>Proportion of cases for AIDS with life long ART: 10%  Average duration of Pre-AIDS asymptomatic sequelae: 8 years  Disability weight (DW) of Pre-AIDS asymptomatic sequelae: 0.051</td>
<td></td>
<td>33.70  60.51</td>
<td></td>
</tr>
<tr>
<td>E2. Averted YLD for AIDS with ART for 11 years cases</td>
<td>Average duration of AIDS with ART: 11 years  Disability weight for AIDS with ART: 0.053</td>
<td></td>
<td>48.16  86.46</td>
<td></td>
</tr>
<tr>
<td>E3. YLL averted for AIDS with ART for 11 years cases</td>
<td>Age at death: 45.6 years  Life expectancy at death: 44.81 years</td>
<td></td>
<td>3,701.31  6,645.32</td>
<td></td>
</tr>
<tr>
<td>E. DALYs averted for AIDS with ART for 11 years</td>
<td>E1 + E2 + E3</td>
<td></td>
<td>3,783.16  6,792.29  55.70%</td>
<td></td>
</tr>
<tr>
<td>Total DALYs averted by Prevention Programme</td>
<td>A+B+C+D+E</td>
<td></td>
<td>38,414.24  69,058.53  55.63%</td>
<td></td>
</tr>
</tbody>
</table>

Source: Author’s calculations based on modeling exercise with Spectrum

It can be seen that the prevention components of the NAP in Timor-Leste saved a total 38,414.24 DALYs from 2007 to 2013. If all the targets of the Programme had been
achieved, the Programme would have saved 69,058.53 DALYs during this period. This gives an outcome efficiency of 55.63% to the prevention components of the Programme.

**DALYs saved by STI programme components:**

Table 8.19 presents the calculation for DALYs saved by the sexually transmitted infections treatment and prevention components of the NAP:

<table>
<thead>
<tr>
<th>Table 8.19: DALYs saved by STI programme components</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baseline</strong></td>
</tr>
<tr>
<td><strong>Total DALYs- STD excluding HIV</strong></td>
</tr>
<tr>
<td><strong>Estimated DALYs without intervention</strong></td>
</tr>
<tr>
<td><strong>Estimated DALYs if full targets achieved</strong></td>
</tr>
<tr>
<td><strong>DALYs averted by intervention</strong></td>
</tr>
<tr>
<td><strong>DALYs averted if full targets achieved</strong></td>
</tr>
<tr>
<td><strong>Efficiency in reducing STD related DALYs</strong></td>
</tr>
</tbody>
</table>

Source: Author’s calculations

As mentioned earlier, this calculation was based on the DALY estimates for the cause STD excluding HIV/AIDS as estimated by the GHE for 2000 and 2012. The GBD (Institute for Health Metrics and Evaluation, 2013) estimate for the same cause in 2005 was adjusted to match with the variation in the GHE methods.

Based on the available review of the NAP (Chan et al., 2013) and STI reporting trends, it was estimated that without the NAP the reduction of DALYs related to the STD would be only 50% of the difference of DALYs between 2005 and 2012 from that level of 2005. The Programme would have reduced 50% more DALYs if full targets of the Programme had been achieved.

**DALYs saved attributable to the Global Fund funded NAP activities:**

In order to calculate the total DALYs saved by the HIV/AIDS treatment, prevention and STI programme components in Timor-Leste, the total DALYs saved by PMTCT and ART programme components, total DALYs saved by prevention programme components, and total DALYs saved by the STI programme components were added up. It was assumed that 10% of the total achievements were attributable to the government funded health
systems and other ongoing efforts in Timor-Leste and therefore, 90% of the total achievements were attributed to the Global Fund funded NAP activities.

Table 8.20 presents the calculation of total DALYs saved by the NAP and DALYs saved attributable to the Global Fund funded NAP activities:

### Table 8.20: DALYs saved attributable to Global Fund funded NAP activities

<table>
<thead>
<tr>
<th></th>
<th>Results</th>
<th>Targets</th>
<th>Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>DALYs saved by ART and PMTCT programme components</td>
<td>121.22</td>
<td>168.57</td>
<td>71.91%</td>
</tr>
<tr>
<td>DALYs saved by Prevention Programme components</td>
<td>38,414.24</td>
<td>69,058.53</td>
<td>55.63%</td>
</tr>
<tr>
<td>DALYs saved by STI Programme components</td>
<td>714.75</td>
<td>1,429.50</td>
<td>50.00%</td>
</tr>
<tr>
<td>Total DALYs saved by NAP</td>
<td>39,250.21</td>
<td>70,656.59</td>
<td>55.55%</td>
</tr>
<tr>
<td>Proportion of achievement attributable to Global Fund funded NAP activities</td>
<td>90%</td>
<td>90%</td>
<td>55.55%</td>
</tr>
<tr>
<td>DALYs averted attributable to Global Fund funded NAP activities</td>
<td>35,325.19</td>
<td>63,590.94</td>
<td>55.55%</td>
</tr>
</tbody>
</table>

Source: Author’s calculations

It can be seen from 2007 to 2013, the Global Fund funded NAP activities in Timor-Leste saved a total of 35,325 DALYs against an optimum target scenario of saving 63,591 DALYs. This gives the NAP an efficiency of 55.55% in terms of saving DALYs.

### Calculating cost effectiveness ratio and technical efficiency:

In order to calculate the cost effectiveness ratio of the Global Fund funded NAP activities, total DALYs saved attributed to this intervention in the actual and optimum scenarios were compared with the total programme expenditure and budgets from 2007 to 2013 as funded by the Global Fund. Table 8.21 summarizes the calculations for cost-effectiveness ratio and technical efficiency of the NAP:

### Table 8.21: Cost effectiveness ratio and technical efficiency of the NAP

<table>
<thead>
<tr>
<th></th>
<th>Results</th>
<th>Targets</th>
<th>Technical Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>DALYs averted attributable to Global Fund funded NAP activities</td>
<td>35,325.19</td>
<td>63,590.94</td>
<td>55.55%</td>
</tr>
<tr>
<td>Total Cost of Global Fund funded NAP activities</td>
<td>$11,842,703</td>
<td>$14,446,916</td>
<td>81.97%</td>
</tr>
<tr>
<td>Cost per DALYs saved</td>
<td>$335.25</td>
<td>$227.19</td>
<td>67.77%</td>
</tr>
</tbody>
</table>

Source: Author’s calculations

It can be seen that the cost per DALY saved by the NAP during 2007 to 2013 was $335.25 against $227.19 cost per DALY saved in the optimum scenario with full targets achieved. Although overall efficiency of the NAP in terms of DALYs saved was 55.55%, it
utilized 81.97% of its allocated budget. In terms of cost per DALYs saved, the technical efficiency of the NAP was 67.77%.

8.5 Cost effectiveness and technical efficiency of IPL

8.5.1 Method

Sources of data:
Although the GBD studies and GHE provided DALYs estimates for Timor-Leste for Childhood cluster diseases such as Diphtheria, Pertussis, Tetanus, and Measles, it was not possible to segregate them for 7 districts covered by IPL and 6 other districts that were not covered by IPL intervention. This chapter, therefore, attempted to assess the cost effectiveness and technical efficiency of the IPL intervention from the IPL programmatic reports on the actual number of children vaccinated, targets, and expenditure. Such information was collected from the Programme Monitoring Plan of IPL (Immunizasaun Proteje Labarik, 2012), periodical progress reports (MCHIP, 2013a), and programme review of IPL (MCHIP, 2013g). The final cost information of IPL was collected from the Aid Transparency Portal of the Government of Timor-Leste (Ministry of Finance of Timor-Leste, 2014a).

Estimating number of new infections and deaths prevented by vaccination:
The number of new infections and deaths averted by the vaccination were estimated by using the programmatic coverage data, case-fatality ratio, vaccine efficacy, and susceptibility of the diseases among the unvaccinated cases during the period of immunization provided by vaccination. Following the suggestions of the Department of Epidemiology of the School of Public Health of the University of California, Los Angeles (http://www.ph.ucla.edu/EPI/41508/415cmat/lect15_41508.pdf accessed on 17 December 2014), this chapter developed a Microsoft Excel based mathematical model to calculate the number of cases and number of deaths prevented by the vaccination programme. The following formula was used to calculate the number of prevented cases:

Number of prevented cases = Number of vaccinated children X Likely susceptibility of infection during the period of immunity if children were not vaccinated x Vaccine effectiveness
The likely susceptibility of infection during the period of immunity if the children were not vaccinated was estimated based on the susceptibility of the disease during pre-vaccination era. For calculating the number of prevented deaths, the following formula was used:

\[
\frac{\text{Number of prevented deaths}}{\text{Number of prevented cases}} \times \text{Case fatality ratio}
\]

Table 8.22 presents the parameters used in calculating the number of cases and deaths prevented by vaccination: the calculations and uncertainty analysis:

**Table 8.22: Parameters for calculating number of cases and deaths prevented by vaccination**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of children vaccinated for DPT3 and Measles</td>
<td>Programmatic data</td>
</tr>
<tr>
<td>Diphtheria:</td>
<td></td>
</tr>
<tr>
<td>Vaccine effectiveness</td>
<td>80%</td>
</tr>
<tr>
<td>Proportion of cumulative background incidence among unvaccinated during the period of immunity provided by vaccination</td>
<td>0.075% 10%</td>
</tr>
<tr>
<td>CFR</td>
<td></td>
</tr>
<tr>
<td>Pertussis:</td>
<td></td>
</tr>
<tr>
<td>Vaccine effectiveness</td>
<td>80%</td>
</tr>
<tr>
<td>Proportion of cumulative background incidence among unvaccinated during the period of immunity provided by vaccination</td>
<td>80% 2%</td>
</tr>
<tr>
<td>Case fatality ratio</td>
<td></td>
</tr>
<tr>
<td>Tetanus:</td>
<td></td>
</tr>
<tr>
<td>Vaccine effectiveness</td>
<td>90%</td>
</tr>
<tr>
<td>Proportion of cumulative background incidence among unvaccinated during the period of immunity provided by vaccination</td>
<td>65% 35%</td>
</tr>
<tr>
<td>Case fatality ratio</td>
<td></td>
</tr>
<tr>
<td>Measles:</td>
<td></td>
</tr>
<tr>
<td>Vaccine effectiveness</td>
<td>90%</td>
</tr>
<tr>
<td>Proportion of cumulative background incidence among unvaccinated during the period of immunity provided by vaccination</td>
<td>100% 3%</td>
</tr>
<tr>
<td>Case fatality ratio</td>
<td></td>
</tr>
</tbody>
</table>

Source: (Birmingham & Stein, 2003; Brenzel et al., 2006; Murray, Lopez, & Mathers, 2004; WHO, 2002, 2004c)

**Calculating DALYs saved by vaccination:**

For calculating DALYs averted by the vaccination programme, the natural history of diseases for Diphtheria, Pertussis, Tetanus and Measles were analyzed to estimate the duration of illness and duration of different outcomes of illness. DALYs averted by vaccination were then calculated by summing up the product of number of deaths averted and life expectancy at the age of death; and years lived in disability as calculated from the number of cases averted, distribution of the cases in different disease sequelae, disability weights of different sequelae, and duration of each sequelae. The parameters assumed for calculating DALYs are presented in Table 8.23:
### Table 8.23: Parameters for calculating DALYs averted by vaccination

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value used</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Diphtheria:</strong></td>
<td></td>
</tr>
<tr>
<td>Average age at onset of episode</td>
<td>3 years</td>
</tr>
<tr>
<td>Duration of episode</td>
<td>0.1 year</td>
</tr>
<tr>
<td>Average age at death</td>
<td>3 years</td>
</tr>
<tr>
<td>Life loss function</td>
<td>89.41</td>
</tr>
<tr>
<td>DW of episode</td>
<td>0.230</td>
</tr>
<tr>
<td>Proportion having Polyneuritis</td>
<td>2%</td>
</tr>
<tr>
<td>DW of Polyneuritis</td>
<td>0.078</td>
</tr>
<tr>
<td>Duration of Polyneuritis</td>
<td>Life-long</td>
</tr>
<tr>
<td>Proportion having Myocarditis</td>
<td>15%</td>
</tr>
<tr>
<td>DW of Myocarditis</td>
<td>0.323</td>
</tr>
<tr>
<td>Duration of Myocarditis</td>
<td>Life-long</td>
</tr>
<tr>
<td><strong>Pertussis:</strong></td>
<td></td>
</tr>
<tr>
<td>Average age at onset of episode</td>
<td>5.5 years</td>
</tr>
<tr>
<td>Duration of episode</td>
<td>0.1 year</td>
</tr>
<tr>
<td>Average age at death</td>
<td>5.5 years</td>
</tr>
<tr>
<td>Life loss function</td>
<td>84.52</td>
</tr>
<tr>
<td>DW of episode</td>
<td>0.178</td>
</tr>
<tr>
<td>Proportion having Mental retardation</td>
<td>0.5%</td>
</tr>
<tr>
<td>DW of Mental retardation</td>
<td>0.469</td>
</tr>
<tr>
<td>Duration of mental retardation</td>
<td>Life-long</td>
</tr>
<tr>
<td><strong>Tetanus:</strong></td>
<td></td>
</tr>
<tr>
<td>Average age at onset of episode</td>
<td>15</td>
</tr>
<tr>
<td>Duration of episode</td>
<td>2 months</td>
</tr>
<tr>
<td>Average age at death</td>
<td>15</td>
</tr>
<tr>
<td>Life loss function</td>
<td>74.54</td>
</tr>
<tr>
<td>DW of episode</td>
<td>0.640</td>
</tr>
<tr>
<td>Proportion having Motor deficit</td>
<td>15%</td>
</tr>
<tr>
<td>DW of Motor deficit</td>
<td>0.388</td>
</tr>
<tr>
<td>Duration of Motor deficit</td>
<td>Life-long</td>
</tr>
<tr>
<td>Proportion having Mental retardation</td>
<td>15%</td>
</tr>
<tr>
<td>DW of Mental retardation</td>
<td>0.469</td>
</tr>
<tr>
<td>Duration of Mental retardation</td>
<td>Life-long</td>
</tr>
<tr>
<td><strong>Measles:</strong></td>
<td></td>
</tr>
<tr>
<td>Average age at onset of episode</td>
<td>2 years</td>
</tr>
<tr>
<td>Duration of episode</td>
<td>2 weeks</td>
</tr>
<tr>
<td>Average age at death</td>
<td>2 years</td>
</tr>
<tr>
<td>Life loss function</td>
<td>89.41</td>
</tr>
<tr>
<td>DW of episode</td>
<td>0.152</td>
</tr>
</tbody>
</table>

Source: (Birmingham & Stein, 2003; Brenzel et al., 2006; Griffiths et al., 2004; WHO, 2013a)

**Attributing DALYs averted by vaccination to IPL:**

Since IPL was an additional effort on top of the National vaccination programme where vaccines were procured by a different source of funding, marginal gains attributable to IPL activities and funding were calculated by comparing the rates of changes of vaccination coverage for DPT3 and Measles from the baseline (coverage for the year 2010) to the end of the project (coverage for 2013 till September) for the districts covered by IPL and the results were compared with the results for the rest of the districts which were not covered by IPL. The results for the districts not covered by IPL intervention gave a “null” (without IPL intervention) scenario whereas the results for
the districts covered by IPL intervention gave the “intervention” (with IPL intervention) scenario to be able to calculate the incremental gains by IPL.

The expected gains by IPL if the targets of the project had been achieved fully were also calculated by using the same methods assuming 81.5% (i.e. the target coverage) of the eligible children had been vaccinated for DPT3 and Measles in the districts covered by IPL intervention. This gave the calculation of expected DALYs averted in the optimum scenario. A proportion of expected DALYs saved was then attributed to IPL activities by comparing the outcomes in optimum scenario in districts covered by IPL with the outcomes in districts not covered by IPL.

**Calculating the cost effectiveness ratio and technical efficiency of IPL:**

The cost per DALY saved by the IPL intervention and the cost per DALY saved expected in the best case (optimum) scenario were calculated by dividing the total programme cost of IPL by attributable DALYs averted by IPL in actual and optimum scenario. The technical efficiency of IPL was then calculating by dividing the cost per DALY saved in the optimum scenario (optimum gains attributable to IPL if 100% project targets were achieved) by the cost per DALY saved in intervention scenario (actual gains attributable to IPL).
8.5.2 Results for IPL

Table 8.24 presents a summary of design, outputs and outcomes of IPL:

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Overview</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evaluation period</td>
<td>April 2011 to September 2013</td>
</tr>
<tr>
<td>Goal</td>
<td>Support the MOH in Timor-Leste to improve immunization coverage to the level of 81.5% for DPT3 and Measles and to strengthen routine immunization services</td>
</tr>
<tr>
<td>Objectives</td>
<td>1) Strengthen service delivery to identify and reach unimmunized children at least five times a year; 2) Strengthen district and CHC-level programme management capacity and technical skills among government health personnel; 3) Strengthen SISCa as an effectively functioning community-based outreach mechanism for providing immunization and other health services; 4) Strengthen programme monitoring and reporting through better collection of routine data and the routine analysis and use of data for decision-making and targeted action</td>
</tr>
<tr>
<td>Inputs</td>
<td>Total funding: $2.63 million; Total estimated expenditure: $2.63 million</td>
</tr>
<tr>
<td>Reported outputs</td>
<td>Percentage of CHCs with current micro-plans, maps, full service strategies increased from 31% to 92%; Improved vaccinator ranking on quality measures increased from 0% to 54%; 192 staff in focus districts trained; percentage of health facilities with good vaccine management increased from 17% to 74%; 46 teachers and religious leaders received EPI orientation; Percentage of CHCs with updated list of missed children by suco increased from 0% to 31%; Percentage of CHCs holding quarterly micro-plan reviews with wide civil society participation increased from 24% to 92%; percentage of CHCs with active system for identifying and following up left-outs and drop-outs increased from 4% to 33%.</td>
</tr>
<tr>
<td>Reported outcomes</td>
<td>Average national coverage of DPT3 and Measles increased from 61.6% (June 2011) to 78% (June 2013)</td>
</tr>
</tbody>
</table>

Source: Author's review of programme planning, budget, performance framework, results and expenditure documents of IPL

The IPL targeted 7 under-performing districts in Timor-Leste to raise the national coverage of vaccination for three doses of Diphtheria, Pertussis and Tetanus (DPT3) vaccines and for a single dose of Measles vaccine to the level of 81.5%. Table 8.25 summarizes the total number of children eligible for vaccination, distribution of number of eligible children in the IPL-focus and IPL-non-focus areas, and the absolute number of children vaccinated for DPT3 and Measles in both areas with calculation of coverage rates for the years from 2010 (baseline prior to intervention) and 2013 (end of intervention):
Table 8.25: Summary of DPT3 and Measles vaccination coverage in IPL-focus and IPL-non-focus districts in Timor-Leste, 2010-2013

<table>
<thead>
<tr>
<th></th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of children eligible for vaccination</td>
<td>42,188</td>
<td>42,188</td>
<td>38,915</td>
<td>39,497</td>
</tr>
<tr>
<td>Number of children eligible for vaccination in IPL-focus districts</td>
<td>28,574</td>
<td>28,574</td>
<td>26,357</td>
<td>26,751</td>
</tr>
<tr>
<td>Number of children eligible for vaccination in IPL-non-focus districts</td>
<td>13,614</td>
<td>13,614</td>
<td>12,558</td>
<td>12,746</td>
</tr>
<tr>
<td>Number of children vaccinated for DPT3 in IPL-focus districts</td>
<td>19,408</td>
<td>18,479</td>
<td>22,101</td>
<td>17,236</td>
</tr>
<tr>
<td>Coverage of DPT3 vaccination in IPL-focus districts</td>
<td>67.92%</td>
<td>64.67%</td>
<td>83.85%</td>
<td>64.43%</td>
</tr>
<tr>
<td>Number of children vaccinated for DPT3 in IPL-non-focus districts</td>
<td>11,116</td>
<td>9,800</td>
<td>10,043</td>
<td>6,741</td>
</tr>
<tr>
<td>Coverage of DPT3 vaccination in IPL-non-focus districts</td>
<td>81.65%</td>
<td>71.98%</td>
<td>79.97%</td>
<td>52.89%</td>
</tr>
<tr>
<td>Number of children vaccinated for measles in IPL-focus districts</td>
<td>16,890</td>
<td>18,083</td>
<td>19,362</td>
<td>18,454</td>
</tr>
<tr>
<td>Coverage of measles vaccination in IPL-focus districts</td>
<td>59.11%</td>
<td>63.29%</td>
<td>73.46%</td>
<td>68.98%</td>
</tr>
<tr>
<td>Number of children vaccinated for measles in IPL-non-focus districts</td>
<td>11,032</td>
<td>9,207</td>
<td>9,039</td>
<td>5,776</td>
</tr>
<tr>
<td>Coverage of measles vaccination in IPL-non-focus districts</td>
<td>81.03%</td>
<td>67.63%</td>
<td>71.98%</td>
<td>45.32%</td>
</tr>
</tbody>
</table>

Source: IPL Final Programme Report 2013 (MCHIP, 2013a)

The following chart shows the combined targets and achievements of vaccination coverage for both DPT3 and Measles in the IPL-focus and IPL-non-focus districts:

![Figure 8.5: Vaccination coverage targets and achievements in IPL-focus and IPL-non-focus districts in Timor-Leste, 2010-2013](image)

Before starting the IPL intervention, the vaccination coverage in 2010 for both DPT3 and measles in IPL-non-focus districts were close to the National targets. It seems that the IPL programme design assumed that non-IPL focus districts would maintain the already achieved vaccination coverage level at 81.5% from 2010 to 2013, whereas with IPL activities, IPL-focus districts would catch up with the national target of vaccination
coverage of 81.5% for DPT3 and Measles. However, in practice, while IPL-focus districts saw improvements in vaccination coverage from 63.52% in 2010 to 66.71% in 2013, vaccination coverage in IPL-non-focus districts dropped from 81.34% in 2010 to 49.10% in 2013.

**New infections and deaths averted by vaccination:**

Table 8.26 presents the estimated number of infections and deaths averted by the vaccination intervention in the IPL-focus districts from 2011 to 2013:

**Table 8.26: Infections and deaths averted by vaccination in IPL-focus districts from 2011-2013**

<table>
<thead>
<tr>
<th>Formula component</th>
<th>Diphtheria</th>
<th>Pertussis</th>
<th>Tetanus</th>
<th>Measles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of children eligible for vaccination</td>
<td>81682</td>
<td>81682</td>
<td>81682</td>
<td>81682</td>
</tr>
<tr>
<td>Proportion of vaccination coverage</td>
<td>70.78%</td>
<td>70.78%</td>
<td>70.78%</td>
<td>68.44%</td>
</tr>
<tr>
<td>Total number of children vaccinated (V)</td>
<td>57,816</td>
<td>57,816</td>
<td>57,816</td>
<td>55,899</td>
</tr>
<tr>
<td>Susceptibility: Proportion of cumulative background incidence among unvaccinated during the period of immunity provided by vaccination (S)</td>
<td>0.08%</td>
<td>80%</td>
<td>65%</td>
<td>100%</td>
</tr>
<tr>
<td>Vaccine effectiveness (VE)</td>
<td>80%</td>
<td>80%</td>
<td>90%</td>
<td>90%</td>
</tr>
<tr>
<td>Number of prevented cases, I = V<em>S</em>VE</td>
<td>35</td>
<td>37,002</td>
<td>33,822</td>
<td>50,309</td>
</tr>
<tr>
<td>Case fatality ratio (CFR)</td>
<td>10%</td>
<td>2%</td>
<td>35%</td>
<td>3%</td>
</tr>
<tr>
<td>Number of prevented deaths, N = I*CFR</td>
<td>3.47</td>
<td>740.05</td>
<td>11,837.86</td>
<td>1,509.28</td>
</tr>
</tbody>
</table>

Source: Author’s calculations

Table 8.27 presents the estimated number of infections and deaths expected to be averted in IPL-focus districts from 2011 to 2013 in the optimum scenario, that is, if IPL could achieve all its targets fully.

**Table 8.27: Target infections and deaths expected to be averted in IPL-focus districts from 2011-2013**

<table>
<thead>
<tr>
<th>Formula component</th>
<th>Diphtheria</th>
<th>Pertussis</th>
<th>Tetanus</th>
<th>Measles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of children eligible for vaccination</td>
<td>81682</td>
<td>81682</td>
<td>81682</td>
<td>81682</td>
</tr>
<tr>
<td>Proportion of vaccination coverage</td>
<td>0.815</td>
<td>0.815</td>
<td>0.815</td>
<td>0.815</td>
</tr>
<tr>
<td>Total number of children vaccinated (V)</td>
<td>66,571</td>
<td>66,571</td>
<td>66,571</td>
<td>66,571</td>
</tr>
<tr>
<td>Susceptibility: Proportion of cumulative background incidence among unvaccinated during the period of immunity provided by vaccination (S)</td>
<td>0.08%</td>
<td>80%</td>
<td>65%</td>
<td>100%</td>
</tr>
<tr>
<td>Vaccine effectiveness (VE)</td>
<td>80%</td>
<td>80%</td>
<td>90%</td>
<td>90%</td>
</tr>
<tr>
<td>Number of prevented cases, I = V<em>S</em>VE</td>
<td>40</td>
<td>42,605</td>
<td>38,944</td>
<td>59,914</td>
</tr>
<tr>
<td>Case fatality ratio (CFR)</td>
<td>10%</td>
<td>2%</td>
<td>35%</td>
<td>3%</td>
</tr>
<tr>
<td>Number of prevented deaths, N = I*CFR</td>
<td>3.99</td>
<td>852.11</td>
<td>13,630.38</td>
<td>1,797.41</td>
</tr>
</tbody>
</table>

Source: Author’s calculations
It can be seen that with 65% background susceptibility (based on pre-vaccination era estimations for the full duration of immunity provided by vaccination) and 35% case fatality ratio, the number of tetanus related deaths prevented by vaccination was quite substantial (11,838). This was followed by measles vaccination with 1,797 deaths prevented based on 100% background susceptibility (based on pre-vaccination era estimations for the full duration of immunity provided by vaccination) and 3% case fatality ratio.

**DALYs saved by vaccination:**

Table 8.28 presents the calculations of DALYs saved by the vaccination intervention in IPL-focus districts from 2011 to 2013 based on the number of infections and deaths averted, natural history of diseases, and life expectancy in the GHE methods:

<table>
<thead>
<tr>
<th>Calculation parameters</th>
<th>Diphtheria</th>
<th>Pertussis</th>
<th>Tetanus</th>
<th>Measles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average age of disease onset (years)</td>
<td>3</td>
<td>5.5</td>
<td>15</td>
<td>2</td>
</tr>
<tr>
<td>Average life expectancy at disease onset</td>
<td>89.41</td>
<td>84.52</td>
<td>74.54</td>
<td>89.41</td>
</tr>
<tr>
<td>Average age at death due to disease</td>
<td>3</td>
<td>5.5</td>
<td>15</td>
<td>2</td>
</tr>
<tr>
<td>Life expectancy at average age at death (LE)</td>
<td>89.41</td>
<td>84.52</td>
<td>74.54</td>
<td>89.41</td>
</tr>
<tr>
<td>Year of Life Lost (YLL) averted = N*LE</td>
<td>310.16</td>
<td>62,548.75</td>
<td>882,393.89</td>
<td>134,944.28</td>
</tr>
<tr>
<td>Average duration of disease episode (in years) (L0)</td>
<td>0.1</td>
<td>0.1</td>
<td>0.16</td>
<td>0.038</td>
</tr>
<tr>
<td>Disability weight (DW) for the disease episode (DW0)</td>
<td>0.23</td>
<td>0.178</td>
<td>0.64</td>
<td>0.152</td>
</tr>
<tr>
<td>Averted Years Lived in Disability (YLD) during disease episode, YLD0 = I<em>L0</em>DW0</td>
<td>0.80</td>
<td>658.64</td>
<td>3,463.42</td>
<td>290.59</td>
</tr>
<tr>
<td>Proportion having disease sequelae 1 from the disease outcomes (Polyneuritis for Diphtheria, Mental retardation for Pertussis, Motor deficit for Tetanus, None for Measles) (P1)</td>
<td>2%</td>
<td>0.50%</td>
<td>15%</td>
<td></td>
</tr>
<tr>
<td>Average duration of disease sequelae 1 (in years) (L1)</td>
<td>89.41</td>
<td>84.52</td>
<td>74.54</td>
<td></td>
</tr>
<tr>
<td>Disability weight (DW) for the disease sequelae 1 (DW1)</td>
<td>0.078</td>
<td>0.469</td>
<td>0.388</td>
<td></td>
</tr>
<tr>
<td>Averted Years Lived in Disability (YLD) during disease sequelae 1, YLD1 = I<em>P1</em>L1*DW1</td>
<td>4.84</td>
<td>7,333.84</td>
<td>146,729.50</td>
<td></td>
</tr>
<tr>
<td>Proportion having disease sequelae 2 from the disease outcomes (Myocarditis for Diphtheria, None for Pertussis, Mental retardation for Tetanus, None for Measles) (P2)</td>
<td>15%</td>
<td>15%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average duration of disease sequelae 2 (in years) (L2)</td>
<td>89.41</td>
<td>74.54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disability weight (DW) for the disease sequelae 2 (DW2)</td>
<td>0.323</td>
<td>0.469</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Averted Years Lived in Disability (YLD) during disease sequelae 2, YLD2 = I<em>P2</em>L2*DW2</td>
<td>150.27</td>
<td>177,361.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Years Lived in Disability (YLD) averted, YLD = YLD0+YLD1+YLD2</td>
<td>155.91</td>
<td>7,992.48</td>
<td>327,554.09</td>
<td>290.59</td>
</tr>
<tr>
<td>Total DALYs averted = YLL + YLD</td>
<td>466.1</td>
<td>70,541.2</td>
<td>1,209,948</td>
<td>135,234.9</td>
</tr>
</tbody>
</table>

Source: Author’s calculations
Table 8.29 presents the DALYs expected to be saved in the IPL-focus districts from 2011 to 2013 in the optimum scenario, that is, if IPL could achieve all its targets fully.

Table 8.29: DALYs expected to be saved in optimum scenario IPL-focus districts, 2011-2013

<table>
<thead>
<tr>
<th>Calculation parameters</th>
<th>Diphtheria</th>
<th>Pertussis</th>
<th>Tetanus</th>
<th>Measles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average age of disease onset (years)</td>
<td>3</td>
<td>5.5</td>
<td>15</td>
<td>2</td>
</tr>
<tr>
<td>Average life expectancy at disease onset</td>
<td>89.41</td>
<td>84.52</td>
<td>74.54</td>
<td>89.41</td>
</tr>
<tr>
<td>Average age at death due to disease</td>
<td>3</td>
<td>5.5</td>
<td>15</td>
<td>2</td>
</tr>
<tr>
<td>Life expectancy at average age at death (LE)</td>
<td>89.41</td>
<td>84.52</td>
<td>74.54</td>
<td>89.41</td>
</tr>
<tr>
<td>Year of Life Lost (YLL) averted = N*LE</td>
<td>357.1</td>
<td>72,020.1</td>
<td>1,016,008.3</td>
<td>160,706.6</td>
</tr>
<tr>
<td>Average duration of disease episode (in years) (L0)</td>
<td>0.1</td>
<td>0.1</td>
<td>0.16</td>
<td>0.038</td>
</tr>
<tr>
<td>Disability weight (DW) for the disease episode (DW0)</td>
<td>0.23</td>
<td>0.178</td>
<td>0.64</td>
<td>0.152</td>
</tr>
<tr>
<td>Averted Years Lived in Disability (YLD) during disease episode, YLD0 = L0*DW0</td>
<td>0.92</td>
<td>758.37</td>
<td>3,987.86</td>
<td>346.06</td>
</tr>
<tr>
<td>Proportion having disease sequelae 1 from the disease outcomes (Polyneuritis for Diphtheria, Mental retardation for Pertussis, Motor deficit for Tetanus, None for Measles) (P1)</td>
<td>2%</td>
<td>0.50%</td>
<td>15%</td>
<td></td>
</tr>
<tr>
<td>Average duration of disease sequelae 1 (in years) (L1)</td>
<td>89.41</td>
<td>84.52</td>
<td>74.54</td>
<td></td>
</tr>
<tr>
<td>Disability weight (DW) for the disease sequelae 1 (DW1)</td>
<td>0.078</td>
<td>0.469</td>
<td>0.388</td>
<td></td>
</tr>
<tr>
<td>Averted Years Lived in Disability (YLD) during disease sequelae 1, YLD1 = L1<em>P1</em>DW1</td>
<td>5.57</td>
<td>8,444.35</td>
<td>168,947.67</td>
<td></td>
</tr>
<tr>
<td>Proportion having disease sequelae 2 from the disease outcomes (Myocarditis for Diphtheria, None for Pertussis, Mental retardation for Tetanus, None for Measles) (P2)</td>
<td>15%</td>
<td>15%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average duration of disease sequelae 2 (in years) (L2)</td>
<td>89.4</td>
<td>74.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disability weight (DW) for the disease sequelae 2 (DW2)</td>
<td>0.323</td>
<td>0.469</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Averted Years Lived in Disability (YLD) during disease sequelae 2, YLD2 = L2<em>P2</em>DW2</td>
<td>173</td>
<td>204,217.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Years Lived in Disability (YLD) averted, YLD = YLD0+YLD1+YLD2</td>
<td>179.5</td>
<td>9,202.7</td>
<td>377,153.2</td>
<td>346.1</td>
</tr>
<tr>
<td>Total DALYs averted = YLL + YLD</td>
<td>536.6</td>
<td>81,222.8</td>
<td>1,393,161.5</td>
<td>161,052.7</td>
</tr>
</tbody>
</table>

Source: Author’s calculations

As seen in earlier calculations, with a substantial number of infections and deaths prevented, higher disability rates, and a higher proportion of cases having long term disease outcomes, the DALYs calculated for tetanus were much higher than the other diseases.
Attributing DALYs saved and calculating cost effectiveness ratio:

It can be seen from the above calculations that a total of 466.1 DALYs saved for Diphtheria, 70,541.2 for Pertussis, 1,209,948 for Tetanus, and 135,234.9 for Measles by the vaccination intervention in IPL-focus districts from 2011 to 2013. Since IPL was an add-on intervention on top of the Government’s programme and vaccines were procured from other sources of funding and administered through the Government’s existing mechanism, the vaccination coverage performance in the IPL-focus districts was compared with the vaccination coverage performance in the IPL-non-focus districts for the same period. In order to compare these performances the average percentage of children vaccinated from 2011 to 2013 was calculated by adding the total number of children vaccinated for each disease in those three years and dividing the total by the total number of children eligible for vaccination in those three years. The percentages change in coverage in both IPL-focus and IPL-non-focus districts from the baseline (2010) to end of the project average achievements (from 2011 to September 2013) were then worked out to derive the percentage of change attributable to IPL intervention.

Table 8.30 summarizes this calculation of attribution:

<table>
<thead>
<tr>
<th></th>
<th>Baseline % of coverage of DPT3 (2010)</th>
<th>Average % of coverage of DPT3 end of Sept 2013</th>
<th>% change in DPT3 coverage</th>
<th>Baseline % of coverage of Measles (2010)</th>
<th>Average % of coverage of Measles end of Sept 2013</th>
<th>% change in measles coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPL focus areas</td>
<td>67.92%</td>
<td>70.78%</td>
<td>2.86%</td>
<td>59.11%</td>
<td>68.43%</td>
<td>9.32%</td>
</tr>
<tr>
<td>Non-IPL focus areas</td>
<td>81.65%</td>
<td>68.31%</td>
<td>-13.34%</td>
<td>81.03%</td>
<td>61.72%</td>
<td>-19.31%</td>
</tr>
<tr>
<td>% of change attributable to IPL</td>
<td>-13.73%</td>
<td>2.47%</td>
<td>16.20%</td>
<td>-21.92%</td>
<td>6.71%</td>
<td>28.63%</td>
</tr>
</tbody>
</table>

Source: Calculated from (MCHIP, 2013g)

It was found that a 16.20% coverage increase for DPT3 (Diphtheria, Pertussis, and Tetanus) and 28.63% coverage increase for Measles vaccination in the IPL-focus districts were attributable to the IPL intervention. These attribution rates were used to calculate the total DALYs saved attributable to the IPL intervention. The total cost of IPL from 2011 to 2013 (assuming each disease component would have 25% of the total cost) was then divided by the attributed DALYs saved by the IPL intervention for each disease to calculate the costs per DALY saved by IPL. Table 8.31 provides a summary for the calculation of cost per DALY saved by IPL:
Table 8.31: Cost per DALY saved by IPL

<table>
<thead>
<tr>
<th></th>
<th>Diphtheria</th>
<th>Pertussis</th>
<th>Tetanus</th>
<th>Measles</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total DALYs saved by</td>
<td>466.07</td>
<td>70,541.24</td>
<td>1,209,947.98</td>
<td>135,234.87</td>
<td>1,416,190</td>
</tr>
<tr>
<td>vaccination</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proportion of</td>
<td>16.2%</td>
<td>16.2%</td>
<td>16.2%</td>
<td>28.63%</td>
<td></td>
</tr>
<tr>
<td>achievement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>attributable to IPL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DALYs saved</td>
<td>75.50</td>
<td>11,427.68</td>
<td>196,011.57</td>
<td>38,717.74</td>
<td>246,232.50</td>
</tr>
<tr>
<td>attributable to IPL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost of IPL</td>
<td>$659,812.50</td>
<td>$659,812.50</td>
<td>$659,812.50</td>
<td>$659,812.50</td>
<td>$2,639,250</td>
</tr>
<tr>
<td>Cost per DALYs saved</td>
<td>$8,738.86</td>
<td>$57.74</td>
<td>$3.37</td>
<td>$17.04</td>
<td>$10.72</td>
</tr>
<tr>
<td>by IPL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Author’s calculations

It can be seen that while the cost per DALY saved varied widely for different diseases based on their susceptibility, duration, disease outcomes, and case fatality ratio, overall IPL achieved a cost effectiveness ratio of $10.72 for the cost per DALY saved.

**Optimum cost effectiveness scenario and technical efficiency of IPL:**

In order to calculate the technical efficiency of IPL, first the proportion of programmatic outcomes attributable to IPL in the optimum scenario was calculated. This calculation followed the same method as in Table 8.30 but assumed 81.5% vaccination coverage achieved in the IPL focus districts. Table 8.32 presents the calculation of expected programmatic outcomes that would be attributable to IPL, if IPL could fully achieve all its targets.

Table 8.32: Expected programmatic outcomes attributable to IPL in the optimum scenario

<table>
<thead>
<tr>
<th></th>
<th>Baseline % of coverage of DPT3 (2010)</th>
<th>Average % of coverage of DPT3 end of Sept 2013</th>
<th>% change in DPT3 coverage</th>
<th>Baseline % of coverage of Measles (2010)</th>
<th>Average % of coverage of Measles end of Sept 2013</th>
<th>% change in measles coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPL focus areas</td>
<td>67.92%</td>
<td>81.50%</td>
<td>13.58%</td>
<td>59.11%</td>
<td>81.50%</td>
<td>22.39%</td>
</tr>
<tr>
<td>Non-IPL focus areas</td>
<td>81.65%</td>
<td>68.31%</td>
<td>-13.34%</td>
<td>81.03%</td>
<td>61.72%</td>
<td>-19.31%</td>
</tr>
<tr>
<td>% of change</td>
<td>-13.73%</td>
<td>13.19%</td>
<td>26.92%</td>
<td>-21.92%</td>
<td>19.78%</td>
<td>41.70%</td>
</tr>
<tr>
<td>attributable to IPL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Author’s calculations

The calculated attribution rates were then applied to the estimated DALYs expected to be saved in the optimum case scenario to work out the total DALYs that would be attributable to IPL intervention, if IPL could fully achieve all its targets.
The costs per DALY saved by IPL in the optimum scenario were then calculated by dividing the total allocated costs for IPL by the total DALYs saved attributable to IPL in the optimum scenario. Finally, the technical efficiency of IPL was calculated by comparing the cost per DALY saved in the optimum scenario by the actual cost per DALY saved as calculated in Table 8.31. Table 8.33 summarizes the calculation of cost per DALY saved by IPL in the optimum scenario and the technical efficiency of IPL intervention.

### Table 8.33: Cost per DALY saved by IPL in optimum scenario

<table>
<thead>
<tr>
<th></th>
<th>Diphtheria</th>
<th>Pertussis</th>
<th>Tetanus</th>
<th>Measles</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total DALYs expected to be saved in optimum scenario (if targets fully achieved)</td>
<td>536.64</td>
<td>81,222.78</td>
<td>1,393,161.54</td>
<td>161,052.71</td>
<td>1,635,973.67</td>
</tr>
<tr>
<td>Proportion of achievement attributable to IPL in optimum scenario</td>
<td>26.92%</td>
<td>26.92%</td>
<td>26.92%</td>
<td>41.7%</td>
<td></td>
</tr>
<tr>
<td>Expected DALYs averted attributable to IPL</td>
<td>144.46</td>
<td>21,865.17</td>
<td>375,039.09</td>
<td>67,158.98</td>
<td>464,207.70</td>
</tr>
<tr>
<td>Total cost allocated to IPL</td>
<td>$659,812.50</td>
<td>$659,812.50</td>
<td>$659,812.50</td>
<td>$659,812.50</td>
<td>$2,639,250</td>
</tr>
<tr>
<td>Expected cost per DALYs saved by IPL in optimum scenario</td>
<td>$4,567.30</td>
<td>$30.18</td>
<td>$1.76</td>
<td>$9.82</td>
<td>$5.69</td>
</tr>
<tr>
<td>Technical efficiency of IPL (actual results against optimum scenario)</td>
<td>52.26%</td>
<td>52.26%</td>
<td>52.26%</td>
<td>57.65%</td>
<td>53.04%</td>
</tr>
</tbody>
</table>

Source: Author’s calculations

It can be seen that on average the cost per DALY saved for IPL would be $5.69 if IPL could achieve all its targets fully. However, it appeared that IPL used 100% of its allocated budget to achieve a $10.72 the cost per DALY saved. This gives IPL an overall technical efficiency of 53.04% for its intervention.
8.6 Comparison of the Cost-effectiveness and technical efficiency of the three projects

Table 8.34 compares the duration, total costs, total DALYs saved, cost per DALY saved, and technical efficiency of three projects:

Table 8.34: Comparison of cost effectiveness and technical efficiency of three projects

<table>
<thead>
<tr>
<th>Intervention duration (years)</th>
<th>HSSP-SP</th>
<th>NAP</th>
<th>IPL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total costs</td>
<td>$16,240,000</td>
<td>$11,942,703</td>
<td>$2,639,250</td>
</tr>
<tr>
<td>Average cost per year</td>
<td>$3,608,889</td>
<td>$1,821,954</td>
<td>$1,055,700</td>
</tr>
<tr>
<td>Total DALYs saved</td>
<td>44,628</td>
<td>35,325.19</td>
<td>246,232.50</td>
</tr>
<tr>
<td>DALY saved per year</td>
<td>9,917</td>
<td>5,435</td>
<td>98,493</td>
</tr>
<tr>
<td>Cost per DALY saved</td>
<td>$363.90</td>
<td>$335.25</td>
<td>$10.72</td>
</tr>
<tr>
<td>Technical efficiency</td>
<td>53.62% (against pessimistic targets) 29.27% (against optimistic targets)</td>
<td>67.77%</td>
<td>53.04%</td>
</tr>
</tbody>
</table>

Source: Author’s calculations

The comparison shows that overall the NAP had a higher technical efficiency against the pessimistic targets (67.77%) than that of IPL (53.04%) and HSSP-SP (53.62%).

8.7 Sensitivity analysis

In order to derive the possible range of variations in the uncertainty of results, a one-way sensitivity analysis was conducted for all projections and DALY calculations by varying key assumptions to the lowest and highest assumed points. The variations were then compared with the values and results used in the calculations of cost effectiveness analysis to assess validity and uncertainty of results.

8.7.1 Uncertainty range for HSSP-SP

For HSSP-SP, the uncertainty was assumed in estimating the total DALYs in the counterfactual (no-intervention) scenario, which would affect the calculation of DALYs averted by the HSSP-SP intervention. Table 8.35 presents the lower range and upper range of estimated annual reduction rates of DALYs in the absence of HSSP-SP intervention.
Table 8.35: Uncertainty ranges for HSSP-SP

<table>
<thead>
<tr>
<th></th>
<th>Median range (value used)</th>
<th>Lower range</th>
<th>Upper range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reduction of DALY from 2008 to 2009</td>
<td>5%</td>
<td>5.50%</td>
<td>4.50%</td>
</tr>
<tr>
<td>Reduction of DALY from 2009 to 2010</td>
<td>4%</td>
<td>4.50%</td>
<td>3.50%</td>
</tr>
<tr>
<td>Reduction of DALY from 2010 to 2011</td>
<td>3%</td>
<td>3.50%</td>
<td>2.50%</td>
</tr>
<tr>
<td>Reduction of DALY from 2011 to 2012</td>
<td>2%</td>
<td>2.50%</td>
<td>1.50%</td>
</tr>
<tr>
<td>Total reduction of DALY from 2008 to 2012</td>
<td>13.30%</td>
<td>15.10%</td>
<td>11.50%</td>
</tr>
</tbody>
</table>

Source: Author’s estimations based on (World Bank, 2007a)

8.7.2 Uncertainty range for the NAP

The uncertainty analysis for number of deaths and new infections averted by the NAP was conducted by the built-in uncertainty analysis tool in the Spectrum software. By using the ranges provided in Table 8.36, the uncertainty analysis engine in the Spectrum software was run for 1000 iterations for 95% plausibility.

Table 8.36: Uncertainty analysis ranges for the NAP

<table>
<thead>
<tr>
<th>Epidemiology</th>
<th>Value used</th>
<th>Lower range</th>
<th>Upper range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transmission of HIV per act (female to male)</td>
<td>0.0011</td>
<td>0.00099</td>
<td>0.00121</td>
</tr>
<tr>
<td>Transmission multiplier for male to female</td>
<td>1</td>
<td>0.9</td>
<td>1.1</td>
</tr>
<tr>
<td>Transmission multiplier for STI</td>
<td>4</td>
<td>3.6</td>
<td>4.4</td>
</tr>
<tr>
<td>Transmission multiplier for MSM contacts</td>
<td>2.6</td>
<td>2.34</td>
<td>2.86</td>
</tr>
<tr>
<td>Relative infectiousness during Primary Infection</td>
<td>8</td>
<td>7.2</td>
<td>8.8</td>
</tr>
<tr>
<td>Months in Primary Stage</td>
<td>3</td>
<td>2.7</td>
<td>3.3</td>
</tr>
<tr>
<td>Relative infectiousness during Asymptomatic stage</td>
<td>1</td>
<td>0.9</td>
<td>1.1</td>
</tr>
<tr>
<td>Relative infectiousness during Symptomatic stage</td>
<td>4</td>
<td>3.6</td>
<td>4.4</td>
</tr>
<tr>
<td>(no ART)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relative infectiousness during Symptomatic stage</td>
<td>0.6</td>
<td>0.54</td>
<td>0.66</td>
</tr>
<tr>
<td>(with ART)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Condom efficacy for transmission reduction (%)</td>
<td>80</td>
<td>72</td>
<td>88</td>
</tr>
<tr>
<td>Reduction in male susceptibility when circumcised (%)</td>
<td>60</td>
<td>54</td>
<td>66</td>
</tr>
<tr>
<td>Reduction in male infectiousness when circumcised</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Size of initial pulse of infection (0 - 0.01)</td>
<td>0.00012</td>
<td>0.000108</td>
<td>0.000132</td>
</tr>
</tbody>
</table>

Source: (Futures Institute, 2014f)
8.7.3 Uncertainty range for IPL

The uncertainty analysis for IPL in the calculation of the number of deaths and new infections averted by vaccination was conducted by using the ranges provided in Table 8.37.

Table 8.37: Uncertainty ranges for IPL

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value used</th>
<th>Lower range</th>
<th>Upper range</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Diphtheria:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vaccine effectiveness</td>
<td>80%</td>
<td>70%</td>
<td>90%</td>
</tr>
<tr>
<td>Proportion of cumulative background incidence among unvaccinated during the period of immunity provided by vaccination</td>
<td>0.08%</td>
<td>0.07%</td>
<td>0.09%</td>
</tr>
<tr>
<td>Case fatality ratio</td>
<td>10%</td>
<td>5%</td>
<td>20%</td>
</tr>
<tr>
<td><strong>Pertussis:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vaccine effectiveness</td>
<td>80%</td>
<td>70%</td>
<td>90%</td>
</tr>
<tr>
<td>Proportion of cumulative background incidence among unvaccinated during the period of immunity provided by vaccination</td>
<td>80%</td>
<td>70%</td>
<td>90%</td>
</tr>
<tr>
<td>Case fatality ratio</td>
<td>2%</td>
<td>0.50%</td>
<td>10%</td>
</tr>
<tr>
<td><strong>Tetanus:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vaccine effectiveness</td>
<td>90%</td>
<td>85%</td>
<td>95%</td>
</tr>
<tr>
<td>Proportion of cumulative background incidence among unvaccinated during the period of immunity provided by vaccination</td>
<td>65%</td>
<td>55%</td>
<td>75%</td>
</tr>
<tr>
<td>Case fatality ratio</td>
<td>35%</td>
<td>30%</td>
<td>40%</td>
</tr>
<tr>
<td><strong>Measles:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vaccine effectiveness</td>
<td>90%</td>
<td>85%</td>
<td>95%</td>
</tr>
<tr>
<td>Proportion of cumulative background incidence among unvaccinated during the period of immunity provided by vaccination</td>
<td>100%</td>
<td>80%</td>
<td>100%</td>
</tr>
<tr>
<td>Case fatality ratio</td>
<td>3%</td>
<td>0.05%</td>
<td>10%</td>
</tr>
</tbody>
</table>

Source: (Birmingham & Stein, 2003; Brenzel et al., 2006; Murray et al., 2004; WHO, 2002, 2004c)
8.7.4 Results of sensitivity analysis

Table 8.38 presents the lowest and highest values of uncertainties for cost per DALY saved and technical efficiencies of the three projects.

Table 8.38: Uncertainty ranges of cost per DALY saved and technical efficiency of three projects

<table>
<thead>
<tr>
<th></th>
<th>Value used</th>
<th>Lower range</th>
<th>Upper range</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HSSP-SP:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost per DALY saved</td>
<td>$363.90</td>
<td>$932.74</td>
<td>$225.16</td>
</tr>
<tr>
<td>with intervention</td>
<td>$195.14</td>
<td>$202.17</td>
<td>$188.72</td>
</tr>
<tr>
<td>Technical efficiency</td>
<td>53.62%</td>
<td>21.68%</td>
<td>83.82%</td>
</tr>
<tr>
<td>Technical efficiency</td>
<td>$106.50</td>
<td>$109.61</td>
<td>$103.60</td>
</tr>
<tr>
<td>against pessimistic</td>
<td>29.27%</td>
<td>11.75%</td>
<td>46.01%</td>
</tr>
<tr>
<td>targets</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>NAP:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost per DALY saved</td>
<td>$335.25</td>
<td>$764.33</td>
<td>$260.91</td>
</tr>
<tr>
<td>with intervention</td>
<td>$227.19</td>
<td>$612.72</td>
<td>$163.14</td>
</tr>
<tr>
<td>Technical efficiency</td>
<td>67.77%</td>
<td>62.53%</td>
<td>80.16%</td>
</tr>
<tr>
<td><strong>IPL:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost per DALY saved</td>
<td>$10.72</td>
<td>$18.35</td>
<td>$5.67</td>
</tr>
<tr>
<td>with intervention</td>
<td>$5.69</td>
<td>$9.59</td>
<td>$3.05</td>
</tr>
<tr>
<td>Technical efficiency</td>
<td>53.04%</td>
<td>28.42%</td>
<td>89.50%</td>
</tr>
</tbody>
</table>

Source: Author’s calculations

8.8 Discussion

DALYs calculated in this cost effectiveness analysis may appear as different from the DALYs calculated in other burden of disease studies. This can be explained by different methodological approaches followed in different studies. Many health economists calculate DALYs based on the local life expectancy of a country with male-female differentiation. These analysts also consider an age weighting factor assuming that people under the age of 15 and over 65 have much less economic productivity than people in the 15 to 65 years age group. Many health economists also consider a time discounting (usually 3%) factor to calculate the net present value (NPV) of the DALYs in the future (Murray & Lopez, 1996).

Although calculating DALYs in the local context is useful for more pragmatic cost-benefit analysis, DALYs calculated in the local context cannot be readily compared with DALYs calculated in different contexts and cannot be aggregated to estimate the global burden of diseases. The 2010 GBD and 2012 GHE studies, therefore, used a single standard life expectancy...
table without any male-female differentiation, age weighting or time discounting to calculate DALYs in any context. While the 2010 GBD method considered 86 years standard life expectancy at birth based on the frontier life expectancy for Japanese women, the GHE method (in GHE 2012 study) increased this standard life expectancy to 92 years based on the projected frontier life expectancy in 2050 by Japanese and Korean women as projected by the UN population projection 2012. The assumption behind this standard life table is, in an ideal situation, and without any disease/condition/causes, people in any country would be able to live up to 92 years (WHO, 2013a).

Since the estimated DALYs saved by HSSP-SP was based on the GHE 2012 estimates for Timor-Leste, in order to be comparable and consistent in estimating DALYs for NAP and IPL, this chapter used the GHE method of calculating DALYs (i.e. with 92 years standard life expectancy at birth, no age-weightage, no time discounting) for all interventions. This approach, in fact, gave much higher DALYs estimates than the WHO DALYs estimates for Timor-Leste for 2002 (WHO, 2004a) and 2004 (WHO, 2009)*.

These high estimates of DALYs averted is very much apparent in the calculation of DALYs averted by the vaccination programme. As explained above, without age-weightage and time discounting, DALYs saved calculated for children in the GHE method are significantly higher than DALYs calculated in GBD 1990 method with age weightage. As it can be seen in the GHE standard life table (WHO, 2013a), a death averted in 1-4 age group will be considered as 89.41 DALYs saved in the GHE method compared to a calculation of 35.17 DALYs saved in the GBD 1990 method.

In addition to the removal of age weighting and time discounting, the estimated background susceptibility of the childhood disease cluster also contributed to the high DALYs estimates for the vaccination programme. Being consistent with the DALY calculation methods used in the 2010 GBD and 2012 GHE studies, this chapter

* For example, DALY for all causes per 100,000 population in Timor-Leste was calculated by WHO as 20,705.5 for 2002 and 28,090.4 for 2004 using the old methods. However, DALY estimates per 100,000 population for all causes in Timor-Leste recalculated by using GHE methods came out as 77,095.8 for 2000 and 48,752.1 for 2010.
considered the susceptibility of the diseases based on the likely infection during the “pre-vaccination” era. It was assumed that without any vaccination 0.075% of the vaccinated children would eventually be infected with Diphtheria, 80% with Pertussis, 65% with Tetanus, and 100% with Measles during the same period of immunity that would be provided by vaccination. Out of those infected, 10% of Diphtheria cases, 2% of Pertussis cases, 35% of Tetanus cases, and 3% of Measles cases would die. Considering 92 years life expectancy at birth and no age-weightage, this pre-vaccination era susceptibility and case fatality ratio gave very high estimates of DALYs averted. However, these counterfactual susceptibility estimations represent a hypothetical scenario and due to the prevailing vaccination programme, the actual incidences of these diseases in Timor-Leste had been much lower.

It also needs to be noted that as this chapter calculated the DALYs averted by the interventions compared to their hypothetical counterfactual (no-intervention) scenarios, the calculated number of DALYs averted by each intervention are not generally comparable to the current DALY estimates for a cause as they appear in GBD and GHE studies. The DALYs estimated in GBD and GHE studies indicate the current burden of disease (estimated deaths and disabilities) for a cause and they are based either on the incidence or the current prevalence of a cause. Therefore, DALYs estimated for a country in GBD and GHE studies do not reflect DALYs saved by a specific intervention.

Calculating DALYs for a SWAp such as HSSP-SP is extremely rare. In fact, in a recent study commissioned by AusAID to assess the economic inputs to the Timor-Leste Health Design, Fabricant (2013) commented that:

“While technical interventions have evidence-based cost-effectiveness and can be compared in terms of DALYs (or the equivalent), this is not possible for health system interventions due to lack of robust available data” (p.13).

However, the World Bank attempted an indicative cost-benefit analysis as part of the ex-ante project appraisal of HSSP-SP. The appraisal document (World Bank, 2007a) suggested that the cost per DALY saved by HSSP-SP would be in the range of US $501 to $822 for the optimistic and pessimistic targets of the project. But this calculation was based on the WHO Global Burden of Disease for 2002 (WHO, 2004a), which used
significantly different methods (such as use of different standard life table with age
discounting, time discounting, and less number of disease sequelae) in calculating
DALYs than the methods used in GHE 2012. Converting the 2002 GBD DALYs to GHE
methods would give a higher number of DALYs averted and a lower cost per DALY saved
comparable to the optimistic and pessimistic targets scenarios for HSSP-SP.

Unlike other interventions, due to indirect results and uncertainty, HSSP-SP considered
two target scenarios such as pessimistic and optimistic scenarios about its expected
outcomes in the economic analysis. As calculated in the GHE method, DALYs saved by
HSSP-SP in fact appeared as lower than the calculated total DALYs expected to be saved
if the intervention had achieved its pessimistic targets, and no wonder World Bank rated
the performance of HSSP-SP as "marginally unsatisfactory". These pessimistic and
optimistic target scenarios, however, do not represent the uncertainty range for the
calculation of DALYs averted by HSSP-SP.

As mentioned earlier, a modeling exercise using Asian Epidemic Model (AEM) was
undertaken in 2005 for the projection of HIV/AIDS in Indonesia, Papua New Guinea and
Timor-Leste (HEMI Study Team, 2006). This projection gave much lower estimates for
HIV/AIDS in Timor-Leste and the currency of those estimates in programme planning
and policy making has been questioned. A recent review of the NAP in Timor-Leste
(Chan et al., 2013) commented that:

"The HEMI [2006] estimated HIV prevalence in the general population aged 15-49 to be 0.07 in
2010. The prevalence among female sex workers and MSM would be 6.2 per cent and 1.36 per
cent respectively. With the exception of MSM, these estimates have proven to be very different
from data empirically obtained in 2010. A figure of 0.68 % prevalence is reported by MOH
(UNGASS 2010) from sentinel surveillance based on antenatal population. This calls into question
the utility of using the HEMI model for future policy and programming.” (pp.16-17)

It should be noted that HEMI (HEMI Study Team, 2006) was an ex-ante projection in the
absence of wide range of evidence, whereas the exercise with Spectrum software used in
this chapter is an ex-post study with the advantage of having more available evidence.
Therefore, there are significant differences in the assumptions used in these two
projections. For example, the Spectrum exercise in this chapter used more updated
projections of populations based on the UN Population Projection 2012, whereas HEMI
used population projections based on Timor-Leste’s census data from the National
Directorate of Statistics for 2004. The size estimates for female sex workers, clients of
female sex workers, and men who have sex with men used in the Spectrum exercise are based on the programmatic coverage and more recent behavioral surveillance survey (2008, 2010, 2011) results. These estimates varied significantly from the estimates used in the HEMI exercise. Being consistent with the Modes of Transmission tool of UNAIDS (2012), infection risk multiplying factor for STI was considered as 4 in this Spectrum exercise, whereas HEMI considered this factor as 3.5. Finally, the risk behavior and programmatic coverage data used in this Spectrum exercise are based on more updated behavioral surveillance, demographic and health survey, and programme reports. In the absence of such evidence, HEMI had to use wider assumptions.

Unlike the health SWAp and the NAP, the cost-effectiveness of vaccination programme is well documented around the world (Atherly et al., 2009; Griffiths & Miners, 2009; Lydon et al., 2008; Sinha et al., 2007; Tu et al., 2012). Lydon et al. (2008) estimated that the cost per DALY averted with the traditional expanded programme of immunization (EPI) would range from US$ 7 to US$ 438. Comparing US$10.72 cost per DALY averted achieved by IPL with this range would place IPL among the highly cost-effective interventions. However, one should again consider the DALY calculation methods used in the Lydon et al. (2008) study and in this chapter before coming to any such conclusion.

IPL conducted a baseline assessment and a final programme review of IPL intervention. But the programme review of IPL (MCHIP, 2013a) did not attempt a cost effectiveness analysis. There also remains the question why IPL's inherent programme design assumption that non-IPL focus districts would maintain their coverage level achieved in 2010 did not work. In other words, why did the vaccination coverage in the initial high performing non-IPL focus districts decrease from 2010 to 2013? The IPL Programme Review (MCHIP, 2013a) indicated that there was a national stock out of vaccines in the first half of 2013 affecting the national coverage of vaccination. However, the stock out of vaccines in 2013 does not explain the decreasing trend in the vaccination coverage in 2011 and 2012 in IPL-non-focus districts.

The final programme report of IPL (MCHIP, 2013a) acknowledged, “Prior to IPL, the USAID-funded child health project, namely TAIS (Timor-Leste Asisténsia Integradu
Saúde), supported the MOH from 2005 to 2011 to improve the quality of immunization (and other child health) services” (p.7). While building on TAIS’s work enabled IPL to demonstrate good performance of vaccination coverage in IPL-focus districts, discontinuation of TAIS and lack of IPL support in the IPL-non-focus districts probably left a gap in those districts resulting in decreasing vaccination coverage.

The performance gap in the first year of intervention is, in fact, evident in all three projects. Looking at the yearly performance of each project, one can see that programmatic performance for the first year of each project was very poor or even negative against their targets. This is probably because of the delayed start of the implementation, transition from a previous project, and initial confusion and coordination issues. This also reflects the fragility situation and the presence of a difficult national environment for programme implementation.

The fragility situation, decreasing vaccination coverage in non-IPL focus districts, and the national stock out of vaccines were largely beyond the project design and resource allocations for IPL. Thus, the issues remained beyond the control of IPL although they affected the achievement of overall programmatic goal of raising the national vaccination coverage to 81.5%.

From an overall look at the comparison table (Table 8.34) it appears that IPL saved maximum DALYs with minimum costs and in minimum time. However, this seems largely due to the nature of the programme and cannot form a sole basis for decision-making. DALYs for different diseases and causes vary widely and decision related to the prevention and treatment of different diseases largely depend on social values and other decision-making considerations rather than their respective disease burden alone. As pointed out in WHO Guide for Cost-Effectiveness Analysis (Edejer et al., 2003):

"resource allocation decisions affecting the entire health sector must also take into account social concerns such as prioritizing the sick, reducing inequalities in health, or addressing the well-being of future generations” (p.4).

However, along with other social concerns and decision making considerations, cost-effectiveness analysis based on the DALY calculations may provide useful information for the decision makers.
It can also be noted that the programme duration for the NAP was the highest (6.5 years) for this evaluation giving this project the advantage of achieving higher economy and efficiency with the initial set up costs diffused over longer period of time than the IPL or HSSP-SP.

The calculation of cost effectiveness ratios such as cost per DALY saved connects the effectiveness of interventions with the effectiveness of budget allocations. Due to different objectives and disease components of the comparing projects, it would be unfair to compare the cost per DALY saved by each project with each other without additional considerations. However, we can compare the cost per DALY saved by each project with similar interventions. Table 8.39 suggests some reference cost per DALY saved by best performing interventions in South Asia:
Table 8.39: "Best buy" health interventions in South Asia

<table>
<thead>
<tr>
<th>Health Intervention</th>
<th>Cost (in US$) per DALY Averted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Childhood immunization</td>
<td>$8.00</td>
</tr>
<tr>
<td>HIV/AIDS prevention</td>
<td>$9 to $126</td>
</tr>
<tr>
<td>Maternal and neo-natal care</td>
<td>$127 to $394</td>
</tr>
</tbody>
</table>

Source: Adapted from Disease Control Priorities Project (Laxminarayan & Ashford, 2008)

Compared to these examples, the cost per DALY saved achieved by HSSP-SP ($363.90) and IPL ($10.72) seem well within the comparable range. However, the cost per DALY saved achieved by the NAP ($335.25) seems more than the double of the highest range of the "best buy" examples above. This indicates that the NAP was highly resource-intensive, although its technical efficiency (67.77%) was the highest among three projects.

However, according to the WHO’s Commission on Macro-economics and Health (Walker et al., 2010), an intervention that gains a year of healthy life (i.e., a DALY averted) at a cost that is less than the GDP per capita of the country where the intervention takes place is “very cost-effective”. An intervention is “cost-effective” if the cost per DALY averted by the intervention is within the range of GDP per capita to three times of GDP per capita. Interventions whose cost per DALY averted are more than three times of the GDP per capita are not considered as cost-effective.

Using this broad classification, one can compare the costs per DALY saved by the HSSP-SP, NAP, and IPL against Timor-Leste’s GDP per capita of $876 in 2010*. It seems that all three interventions, even with their uncertainty ranges, fall within the classification of “very cost-effective” interventions.

The idea of using a cost effectiveness threshold is also recommended by the UK National Institute of Health and Care Excellence (NICE) in evaluating new health technologies. According to NICE, if a health technology has an incremental cost effectiveness ratio (ICER) below £20,000, the decision to recommend it can be normally based on cost-

effectiveness. Above this threshold, NICE suggests to consider other decision making criteria including uncertainty, innovation, non-health outcomes, end-of-life considerations, stakeholder perspectives on quality of life gains, and special considerations for technologies aimed at children, disadvantaged populations, and severe diseases (National Institute for Health and Care Excellence, 2013).

Cost effectiveness analysis using DALYs as health programme outcome measures and calculating technical efficiency against the programmatic targets can be useful tools for comparing and evaluating a health programme’s performance against the performance of other programmes. However, several limitations need to be considered in interpreting the results.

First, calculating DALYs averted by a SWAp such as HSSP-SP is highly dependent on the underlying assumptions and, therefore, should only be taken as indicative. Other than the effect of HSSP-SP, there could be many other confounding factors for the changes in DALYs. For example, the construction of roads and increase of number of vehicles and their possible implications for increasing DALYs due to road accidents and injuries were not considered in the cost effectiveness analysis. This chapter consistently used the simpler method of calculating DALYs as used in the GHE (i.e. without time and age discounting and considering 92 years standard life expectancy at birth) in all calculations. However, converting DALYs for 2005 and 2010 from the GBD (Institute for Health Metrics and Evaluation, 2013) studies and estimating DALYs for the interim years from 2005 to 2012 were based on a gross conversion rate calculated from the total DALYs in 2000 as per GHE and GBD methods. Therefore, all these estimates should be considered as indicative. A more sophisticated model that could address the methodological difference issues and confounding factors would be more useful in getting more plausible estimates. However, developing and testing such model was beyond the scope of this research, as this research focused more on lessons learned in piloting and testing the current evaluation approach rather than establishing the empirical results.

Second, the use of the Spectrum software for estimating the number of new infections and deaths averted by the NAP is also dependent on many assumptions. Results from
any modeling tool are only as good as the inputs used for the modeling. Although this chapter attempted to triangulate and synthesize all available information to approximate a best-informed scenario, the estimations should not be considered as absolute.

Third, the number of cases and deaths prevented by the vaccination programme facilitated by IPL were calculated from the generalized assumptions used in GBD studies. These calculations used a hypothetical counterfactual scenario of the susceptibility of infection in the event of no vaccination during the period of immunity that would be provided by the vaccination. This estimated susceptibility was based on the infectiousness and disease prevalence of diphtheria, pertussis, tetanus, and measles during the pre-vaccination era for the full period of immunity that would be provided by vaccination. Due to the uncertainty of counterfactual, these results also need to be considered as indicative.

Fourth, the cost effectiveness analysis in this chapter only attempted to measure the direct health benefits in terms of the total DALYs saved by an intervention and calculate the cost-effectiveness ratio in terms of cost per DALY saved. As this is not a full cost-benefit analysis, the social perspectives of all costs and benefits are missing from this analysis. For example, the cost components only considered the programme planner’s point of view of implementing the programme and did not include the cost for seeking treatment such as out of pocket expenses by the treatment seekers. The cost of treatment saved by preventing the number of cases was also not considered. Such additional costs and benefits data were not available and as such it was difficult to measure directly the number of cases of any specific cause prevented by a sector-wide health system strengthening effort such as HSSP-SP.

Fifth, due to the difficulties in measurement and attribution, the benefits from strengthening health systems by the projects were not considered in the cost effectiveness analysis. For the similar reasons of lack of data, any indirect benefits including social benefits were not included in the analysis. Therefore, the ratio for cost per DALY saved should be interpreted accordingly.
8.9 Conclusion

This chapter used cost effectiveness analysis techniques for calculating and comparing the cost per health benefit (i.e. DALY averted) for the intervention outcomes and best-case scenario for three different projects in Timor-Leste’s health sector. Although each project was very different in their nature, had different objectives, and employed a different approach to implementation, considering the fact that the projects operated within the same health system’s context and around the same time, the outcomes and technical efficiencies of the projects were relatively comparable to assess their effectiveness.

The cost effectiveness ratio and technical efficiency based on the cost per DALY saved in actual and optimum scenario attempted to connect the inputs of the projects with the outcomes to evaluate their performance. While comparing the figures related to the cost effectiveness ratio and technical efficiency can provide useful information about value of an intervention, it does not necessarily indicate whether adherence to the aid effectiveness principles and management modalities of these projects had any effect on their effectiveness and efficiency. The next chapter will attempt to compare the degree of adherence to the Paris principles of aid effectiveness by an intervention with its economy, efficiency, and effectiveness to see if there are any correlations.
9 The Paris principles and comparative economy, efficiency, and effectiveness

9.1 Introduction

This chapter serves two purposes. First, it attempts further comparison and analysis of performance of the three projects through the technique of data envelopment analysis (DEA). The findings of the DEA are compared with the findings from the stakeholder perspectives and cost effectiveness analysis for greater validation and triangulation of the results. Second, this chapter attempts a correlations analysis among the variables of the research findings to see if there is any correlation between the Paris principles of aid effectiveness and any other evaluation variables.

The methods and results sections of this chapter are organized in two parts: the first part introduces the concept and technique of DEA, describes the methods used for such analysis in this chapter, and presents the results of such analysis along with their comparison with findings in other chapters. The second part presents the methods and results from the correlations analysis. This is followed by discussion and conclusion.

9.2 Data Envelopment Analysis: concept and techniques

As defined in the Steering Committee for the Review of Commonwealth/State Service Provision (1997), effectiveness is the “degree to which the outputs of a service provider achieve the stated objectives of that service”; and efficiency is the “degree to which the observed use of resources to produce outputs of a given quality matches the optimal use of resources to produce outputs of a given quality” (p.14).

Simply put, the basic measure of efficiency is the ratio between the output and the input:

\[ \text{Efficiency} = \frac{\text{Output}}{\text{Input}} \]

The current Chapter uses these concepts of efficiency and effectiveness to compare the economy, performance efficiency, and outcome efficiency (effectiveness) of the three externally funded projects through a DEA technique.
Typically in DEA there are a number of producers, which are called ‘Decision Making Units’ or DMU in DEA literature (Banker et al., 1984; Banker et al., 1986; Charnes et al., 1978; Farrell, 1957). Each DMU uses a varying set of inputs to produce varying levels of outputs. The fundamental assumption behind DEA is: in a given context, if a given DMU A is capable of producing Y(A) units of output with X(A) inputs, then other DMUs should also be able to do the same if they work at the same level of efficiency.

Similarly, if DMU B is capable of producing Y(B) units of outputs with X(B) units of inputs, then other DMUs should also be able to do the same. Therefore, by using a virtual combination of inputs and outputs from the best performing DMUs, it is possible to sketch a best practice data envelop or efficiency frontier.

DMUs that fall on the best practice frontier are assigned an efficiency score of 1 (or 100%) and are considered technically efficient compared to their peers. The efficiency of the DMUs that fall below the efficiency frontier is measured in terms of their distance from the frontier. The inefficient DMUs are assigned a score between 1 (one) and 0 (zero). The larger the score the more efficient a DMU is (Coelli, Rao, & Battese, 2005).

To illustrate further how DEA works in evaluating health projects, let us imagine three projects, A, B, and C, from a fragile state’s health sector. Let us suppose each project used 1 million dollars equivalent of inputs and produced different levels of outputs related to the achievement of project objectives (project outcomes) and their net effect on the health systems (health system outcomes). Project-A produced 20 units of health system outcomes and 1000 units of project outcomes; Project-B produced 40 units of health system outcomes and 400 units of project outcomes; and Project-C produced 140 units of health system outcomes and 200 units of project outcomes. Assuming a constant return to scale (i.e. outputs are equally proportionate to any increase or decrease of the inputs), we can sketch the performance of three projects as follows:
The figure above shows that among the three projects, Project C and Project A are on the efficiency frontier line (CA), created by the performance of Project C and Project A. But the efficiency of Project B falls below the frontier line. Therefore Project B’s outcomes are inefficient compared to the outcomes of the other two projects.

Now, the efficiency of Project B can be compared to a virtual Project V falling on the same efficiency line CA. If it is assumed that a combination of inputs and outputs of Project A and Project C is possible, then this virtual Project V would be able to produce the same level of outputs by using AV/AC units of Project C and CV/AC units of Project A inputs.

The efficiency of Project B then can be calculated by finding the fraction of inputs that Project V would need to produce the same level of outputs as Project B. This can be calculated by looking at the line from the origin O to V. The efficiency of Project B in this case is OB/OV. On the other hand, any virtual Project for analyzing Project C (say V2) or Project A (say V1) would fall on the same point C and A. Therefore, their efficiency OC/OV2 or OA/OV1 would be equal to 1.

Using linear programming can solve the same example to calculate Project B’s efficiency as a ratio of total weighted outputs and inputs. Let us say we will form a virtual project by maximizing weighted outputs for Project B for L1 weight of project outcomes, L2
weight of health system outcomes, and L3 weight of inputs combinations as illustrated in Table 9.1:

Table 9.1: Illustrating DEA: efficiency measurement using linear programming

<table>
<thead>
<tr>
<th>Project</th>
<th>Project outcomes</th>
<th>Health system outcomes</th>
<th>Inputs (Million dollars)</th>
<th>Weighted outputs</th>
<th>Weighted inputs</th>
<th>Efficiency</th>
<th>Constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1000</td>
<td>20</td>
<td>1</td>
<td>1000<em>L1+20</em>L2</td>
<td>1*L3</td>
<td>(1000<em>L1+20</em>L2)/1*L3</td>
<td>(1000<em>L1+20</em>L2)–1*L3</td>
</tr>
<tr>
<td>B</td>
<td>400</td>
<td>40</td>
<td>1</td>
<td>400<em>L1+40</em>L2</td>
<td>1*L3</td>
<td>(400<em>L1+40</em>L2)/1*L3</td>
<td>(400<em>L1+40</em>L2)–1*L3</td>
</tr>
<tr>
<td>C</td>
<td>200</td>
<td>140</td>
<td>1</td>
<td>200<em>L1+140</em>L2</td>
<td>1*L3</td>
<td>(200<em>L1+140</em>L2)/1*L3</td>
<td>(200<em>L1+140</em>L2)–1*L3</td>
</tr>
<tr>
<td>Weight</td>
<td>L1</td>
<td>L2</td>
<td>L3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In this example, the virtual project consisting of fractions of Project A, B and C should use less than or equal to Project B’s level of inputs and produce more than or equal to Project B’s level of outputs. If 100% of the weighted inputs are converted into 100% weighted outputs, then the difference between weighted outputs and inputs must be equal to or less than zero. Also the value of the multiplying factors (weights) L1, L2, and L3 cannot be negative, that is L1, L2, and L3 must be greater than or equal to 0 (Zero).

So, mathematically we can write the problem as:

Maximize:
400*L1+40*L2 (i.e. maximize the total weighted outputs of Project B)

By changing values of:
L1, L2, L3 (i.e. calculate the multiplying factors for L1, L2 and L3 that would maximize the weighted outputs of Project B)

Subject to:
1*L3 = 1 (i.e. the weighted inputs of Project B is set to 1 so that the weighted outputs of Project B denotes the relative efficiency of Project B compared to the inputs)
(1000*L1+20*L2) – 1*L3 <= 0 (i.e. the difference between weighted outputs and weighted inputs of Project A must be less than or equal to Zero)
(400*L1+40*L2) – 1*L3 <= 0 (i.e. the difference between weighted outputs and weighted inputs of Project B must be less than or equal to Zero)
(200*L1+140*L2) – 1*L3 <= 0 (i.e. the difference between weighted outputs and weighted inputs of Project C must be less than or equal to Zero)
L1 >= 0 (i.e. value of L1 cannot be negative)
L2 >= 0 (i.e. value of L2 cannot be negative)
L3 >= 0 (i.e. value of L3 cannot be negative)

Solving this problem using a Microsoft Excel Solver with the above parameters and constraints gives the following result:
The table above shows that compared to the performance of Project A and Project C, the performance of Project B is 0.59 or 59% efficient in terms of producing project outcomes and health system outcomes with a similar level of inputs used by their comparator projects.

As explained in the methods section below, this chapter used a similar DEA technique with Microsoft Excel Solver for comparing the relative economy, efficiency, and effectiveness of the three projects with multiple outputs.

### 9.3 Methods

#### 9.3.1 Data sources

The DEA technique in this chapter made use of primary sources of data such as project documents as well as findings from the stakeholder interviews in Chapter 7 and cost effectiveness analysis in Chapter 8. The project documents, as listed in Annex 5, included the project asset register, staff list, project results framework, project progress reports, project financial reports, and available audit reports of the three projects for the evaluation period. This chapter also used the health system outputs and health system outcome ranking for three projects as assessed by the key informants during the stakeholder interviews conducted for analyzing stakeholders’ perspectives in Chapter 7. The findings from the Chapter 8 on the DALYs saved by each project (reference to Table 8.34) were also used for calculating relative effectiveness of each project.
9.3.2 Calculating economy

The procurement items considered in the comparison for economy of procurement included the project goods such as equipment and vehicles as well as human resources such as staff and technical assistance. Accordingly, the economy of procurement of common items was measured at two levels: the economy of procurement of common goods, and the economy of procurement of staff time and technical assistance inputs. For the DEA exercise, sample costs for the comparable procurement items in three projects were collected from the asset lists of the three projects as listed in Annex 5.

In order to avoid the risk of all projects being highly specialized and non-comparable to other projects, the DEA exercise only considered the common items from the asset lists of the projects. These common items included printer, laptop computers, desktop computers, motorbikes, and projectors. The number of units and total costs of procuring these items were collected from the asset register of each project. This calculation, however, provides only some indicative economy of procurement of common goods.

The local staff time used by each project was measured in terms of total number of person-months of input from the local staff and consultants for the year 2012. Similarly, technical assistance input for each project was measured in terms of total number of person-months of inputs from the international consultants in the year 2012. Total cost of local and international staff and consultants for each project was collected from the staff list, and corresponding cost categories of project financial reports including any available audit reports.

9.3.3 Calculating performance efficiency

The performance efficiency of each project was measured by considering the extent to which each project achieved the performance targets and the degree of health system strengthening related outputs achieved by each project. These outputs were then compared with the percentage of project budget expended by each project.

Measuring the extent of achievement of project targets by each project considered the performance indicators of each project as appeared in their respective results
framework and assessed each project’s progress against those indicators from periodic progress reports. For measuring the performance by each project over a period of time, the calculation of the performance ratio considered the baseline value of each indicator and compared the actual change from the baseline value based on the project’s results with the expected change from the baseline value based on the project’s target:

\[
\text{Performance ratio for an indicator} = \frac{\text{Actual result for the indicator} - \text{baseline value for the indicator}}{\text{Target set for the indicator} - \text{baseline value for the indicator}}
\]

The result of the calculation of the performance ratio was adjusted for the indicators where the actual performance was worse than the baseline value, and where the actual performance exceeded the target performance. Being consistent with the efficiency measurement in a scale of ‘0’ (zero) to ‘1’ (one), any negative ratio (i.e., when the actual result is worse than the baseline value) was considered as ‘0’ (zero) and any ratio that exceeded 1 (i.e., when the actual result exceeded the target) was considered as 1 or 100%. Performance ratios against each of the indicators of a project were then aggregated and averaged to calculate the overall project performance ratio of a project.

The health system strengthening related outputs of each project was assessed based on the stakeholders’ interviews. As described in Chapter 7, a total of 22 interviews were conducted with project stakeholders and key informants using a structured questionnaire. For assessing health system strengthening related outputs, each interviewee was asked to assess the percentage of achievement of health system outputs achieved by each project (related to service delivery, workforce, information, supplies and commodities, health financing, and governance) compared to their expected targets. The interviewees were asked to rank the performance of each project in a scale of 1-5, where 1 represented the lowest and 5 the highest score (reference to Question 13 of Table 7.3).

The DEA model considered both the extent of project outputs (project performance ratio) and the extent of health system outputs (based on stakeholder interviews) of each project and compared them with the extent of expenditure by each project. The extent of
expenditure was measured by the ratio of total project expenditure and total budget of a project.

9.3.4 Calculating effectiveness

Like the performance efficiency measurement, the total outcomes of a project considered both the value of project outcomes and the value of health system related outcomes. The value of project outcomes was measured by the DALYs saved by each project as calculated in Chapter 8 (reference to Table 8.34). However, for measuring health system related outcomes, the calculation used the health system outcome scores obtained from the stakeholders’ interviews as described in Chapter 7 (reference to Question 18 of Table 7.3).

The measurement of possible health system outcome values by each project first calculated the total investment related to health system strengthening by each project by considering their investments on technical assistance, training, durable goods, equipment and infrastructure. Regardless of the objectives of the project, it was assumed that these investments (i.e. on technical assistance, training, durable goods, equipment and infrastructure) would have longer lasting effects on the health system’s performance improvement than the investment on other project activities. The health system outcome performance ratio of each project as obtained from the stakeholder interviews referenced above was then applied to the total investment on health system strengthening by each project to derive a health system strengthening outcome value.

Both the project outcomes (i.e., DALYs saved) and health system outcome values were then compared with the total expenditure by each project to calculate and compare their relative effectiveness.

9.3.5 DEA model used

Following the instructions in Sherman and Zhu (2006) this chapter developed a multiplier DEA model by using Microsoft Excel 2011 and its Solver add-in component. The DEA model was customized for calculating relative efficiencies of each project for the aspects of 1) economy of procurement of goods and equipment, 2) economy of procurement of human resources and technical assistance, 3) performance efficiency,
and 4) effectiveness by using the outputs and inputs values discussed above. Each instance of the DEA model was run for each project to measure their relative efficiencies. Except for the case of economy of procurement of goods and equipment, the DEA model used the linear programming method (LP Simplex) of the Excel Solver to calculate the efficiencies.

It appeared that the DEA model for calculating economy of procurement of goods and equipment had a higher number of outputs (total 5) than the number of projects (total 3). Using LP Simplex method of Solver in this case returned a result showing that all three projects were 100% efficient on this aspect. In this case, a non-linear method (GRG Non-linear) was used in the DEA model to identify possible inefficiencies in the procurement of goods and equipment.

9.4 Results

9.4.1 Economy

The economy of procurement:
The asset list of three projects showed that HSSP-SP spent a total of $127,981 USD to procure 3 printers, 11 laptop computers, 13 desktop computers, 60 motorbikes, and 3 projectors during the period from 2008 to 2012. The NAP, on the other hand, procured 4 printers, 51 laptop computers, 7 desktop computers, 15 motorbikes and 4 projectors with a total $101,777 USD during the period from 2007 to 2013. IPL spent a total of $54,063 USD and procured 5 printers, 8 laptop computers, 2 desktop computers, 21 motorbikes, and 2 projectors.

Solving these outputs (items procured) and input (total USD spent) for three projects in the DEA model gave the following results:

Table 9.3: DEA result: Economy of procurement of goods and equipment

<table>
<thead>
<tr>
<th>Project</th>
<th>Outputs (number of items procured)</th>
<th>Input used</th>
<th>Relative efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Printer</td>
<td>Laptop computer</td>
<td>Desktop computer</td>
</tr>
<tr>
<td>HSSP-SP</td>
<td>3</td>
<td>11</td>
<td>13</td>
</tr>
<tr>
<td>NAP</td>
<td>4</td>
<td>51</td>
<td>7</td>
</tr>
<tr>
<td>IPL</td>
<td>5</td>
<td>8</td>
<td>2</td>
</tr>
</tbody>
</table>
As noted earlier, HSSP-SP followed the government systems and procurement methods for procuring the items through the MoH’s procurement unit. While the NAP also followed the government systems and procurement methods, procurement of the items for this project was facilitated by the Division of Global Fund (a specialized project facilitation unit set up by the MoH for the Global Fund funded projects) and payment was made through the Division of Global Fund. Procurement of the items for IPL, on the other hand, was done by IPL itself outside the mechanism of the MoH and by following the procurement policies of USAID.

It can be seen from the DEA model results that the performance of HSSP-SP in terms of achieving economy of procurement (i.e. procuring the same number of items with least possible costs) was dominated by the performance of the NAP and IPL. The 87.2% relative efficiency for HSSP-SP indicates that HSSP-SP would need to reduce its procurement costs by 12.8% (i.e. 100% – 87.2%) to achieve the same level of efficiency with HSSP-SP and IPL for economy of procurement.

For further validation, this result can be compared with the similar item in the stakeholder interviews conducted in Chapter 7. Question 11 of the stakeholders’ interview questionnaire asked the respondents to rank these three projects for their performance to achieve economy of procurement on a scale of 1 to 5. Scoring 1 represented the lowest level of performance with “No discernable use of procurement to manage or reduce costs” and 5 represented the highest level of performance with “Costs are significantly reduced and managed to very good effect”. Table 9.4 provides the summary of responses from the stakeholders:

<table>
<thead>
<tr>
<th>Project</th>
<th>Total score</th>
<th>Number of respondents</th>
<th>Mean score</th>
<th>Standard deviation</th>
<th>Economy of procurement performance ratio (mean score : highest point in scale)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSSP-SP</td>
<td>41.5</td>
<td>13</td>
<td>3.19</td>
<td>0.663</td>
<td>63.85%</td>
</tr>
<tr>
<td>NAP</td>
<td>50</td>
<td>15</td>
<td>3.33</td>
<td>0.816</td>
<td>66.67%</td>
</tr>
<tr>
<td>IPL</td>
<td>29</td>
<td>8</td>
<td>3.63</td>
<td>0.744</td>
<td>72.5%</td>
</tr>
</tbody>
</table>

Source: Chapter 7 Table 7.3

It can be seen that the stakeholder views are quite consistent with the DEA results on economy of procurement as provided in Table 3. Like the DEA results showing 87.20% relative efficiency of HSSP-SP compared to 100% relative efficiency of the NAP and IPL.
for procurement of goods and equipment, the stakeholder views also assigned the lowest efficiency rate (63.85%) to HSSP-SP.

**The economy of procuring local staff and technical assistance inputs:**

Comparing the staff time and international technical assistance (TA) inputs used by the three projects in a single year in 2012, it was found that HSSP-SP spent a total of $646,438 USD ($104,255 for local staff and $544,183 for TA) to procure 152.6 person-months of local staff/consultant inputs and 61.6 person months of international technical assistance input (excluding the technical assistance procured through the World Bank but charged to the project). The NAP spent a total of $744,44 USD ($309,961 for local staff and $434,483 for TA) to procure 576 person-months of local staff inputs and 55 person-months of international technical assistance inputs. IPL, on the other hand, spent a total of $581,076 USD ($236,775 for local staff and $344,301 for TA) and procured 468 person-months of local staff inputs and 19.32 person months of international technical assistance inputs.

Calculating the cost per person-month of inputs by local staff and international technical assistance, gives the following unit costs for three projects:

**Table 9.5: Unit costs of local staff and international technical assistance inputs**

<table>
<thead>
<tr>
<th>Project</th>
<th>Average monthly cost per local staff input (USD)</th>
<th>Average monthly cost per international technical assistance input (USD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSSP-SP</td>
<td>$670.08</td>
<td>$8,834.15</td>
</tr>
<tr>
<td>NAP</td>
<td>$538.13</td>
<td>$7,899.70</td>
</tr>
<tr>
<td>IPL</td>
<td>$505.93</td>
<td>$17,820.98</td>
</tr>
</tbody>
</table>

Source: Calculated from staff list and financial reports of HSSP-SP, NAP, and IPL

It can be seen that while IPL’s local staff cost was the lowest among three projects, IPL’s international technical assistance cost appeared as very high compared to those costs of HSSP-SP and the NAP. This was probably due to the fact that the US based organisation JSI, which was implementing IPL in Timor-Leste, recovered some US based TA costs from IPL and included their international travel costs in the TA costs.
Solving all these outputs (staff and TA time procured) and inputs (money used for procuring) of three projects in the DEA model gave the following results:

**Table 9.6: DEA results for economy of procuring local staff and international technical assistance inputs**

<table>
<thead>
<tr>
<th>Project</th>
<th>Outputs (Person-months procured)</th>
<th>Input used</th>
<th>Relative efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Staff</td>
<td>Technical assistance</td>
<td>USD</td>
</tr>
<tr>
<td>HSSP-SP</td>
<td>152.6</td>
<td>61.6</td>
<td>$646,438</td>
</tr>
<tr>
<td>NAP</td>
<td>576</td>
<td>55</td>
<td>$774,444</td>
</tr>
<tr>
<td>IPL</td>
<td>468</td>
<td>19.32</td>
<td>$581,076</td>
</tr>
</tbody>
</table>

The 100% relative efficiency for all three projects implies that there was no significant difference in the performance of three projects in terms of achieving the economy of procuring local staff time and international technical assistance inputs. In other words, none of these projects had any opportunity for efficiency saving compared to the performance of each other.

**9.4.2 Performance efficiency**

As explained in the methods section, calculation of performance efficiency considered project results and stakeholders’ score on health system outputs of three projects.

**Project results for HSSP-SP:**

Table 9.7 presents the project results and performance of HSSP-SP against its performance indicators, targets and baseline:
Table 9.7: Performance indicators and achievement of targets by HSSP-SP

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Baseline</th>
<th>Results</th>
<th>Target</th>
<th>Percentage of change against baseline &amp; target</th>
<th>Adjusted achievement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Percentage of children under 1 year of age vaccinated against (i) DPT3;</td>
<td>63%</td>
<td>83%</td>
<td>90%</td>
<td>74.07%</td>
<td>74.07%</td>
</tr>
<tr>
<td>2. Percentage of children under 1 year of age vaccinated against (ii) measles</td>
<td>61%</td>
<td>73%</td>
<td>90%</td>
<td>41.38%</td>
<td>41.38%</td>
</tr>
<tr>
<td>3. Percentage of births attended by skilled health personnel</td>
<td>27%</td>
<td>59%</td>
<td>45%</td>
<td>177.78%</td>
<td>100%</td>
</tr>
<tr>
<td>4. Percentage of pregnant women receiving four or more prenatal visits</td>
<td>36%</td>
<td>49%</td>
<td>55%</td>
<td>68.42%</td>
<td>68.42%</td>
</tr>
<tr>
<td>5. Percentage of children (6-59 months) receiving vitamin A supplements</td>
<td>36%</td>
<td>57%</td>
<td>80%</td>
<td>47.73%</td>
<td>47.73%</td>
</tr>
<tr>
<td>6. Contraceptive Prevalence Rate</td>
<td>10</td>
<td>26</td>
<td>25</td>
<td>106.67%</td>
<td>100%</td>
</tr>
<tr>
<td>7. Number of formal MOH-private sector/NGO contracts signed</td>
<td>0</td>
<td>6</td>
<td>10</td>
<td>60.00%</td>
<td>60%</td>
</tr>
<tr>
<td>8. Availability of tracer essential drugs at (i) SAMES</td>
<td>88</td>
<td>81</td>
<td>98</td>
<td>-70.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>9. Availability of tracer essential drugs at (ii) health facilities</td>
<td>85 (estimated)</td>
<td>60</td>
<td>95</td>
<td>-250.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>Overall performance against targets and baseline</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>54.62%</td>
</tr>
</tbody>
</table>

Source: Calculated from (Lee, 2013)

It can be seen that achievements of HSSP-SP against indicators like births attended by skilled health personnel and contraceptive prevalence rate exceeded the targets and thus, the project received 100% adjusted performance scores against these indicators. However, there was shortage of tracer essential drugs at the central level (i.e. at the central medical warehouse called SAMES) and at the health facility levels. Performance against these two indicators, in fact, deteriorated from their baseline and HSSP-SP received 0% adjusted performance scores against these indicators. Overall, HSSP-SP achieved 54.62% of its targets by 31 December 2012. This rate of achievement is, in fact, comparable to the “Moderately Unsatisfactory” rating assigned by the World Bank for the overall progress of HSSP-SP towards its objectives (Lee, 2013).

**Project results for the NAP:**

Table 9.8 provides the performance indicators and results achieved by the NAP by 31 December 2013:
Table 9.8: Performance indicators and achievement of targets by NAP

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Baseline</th>
<th>Result</th>
<th>Target</th>
<th>Percentage of change</th>
<th>Adjusted achievement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Percentage of sex workers who are HIV infected</td>
<td>2.76%</td>
<td>2%</td>
<td>&lt;5%</td>
<td>294.74%</td>
<td>100%</td>
</tr>
<tr>
<td>2. Percentage of MSM who are HIV infected</td>
<td>1.3%</td>
<td>1%</td>
<td>&lt;3%</td>
<td>566.67%</td>
<td>100%</td>
</tr>
<tr>
<td>3. Percentage of adult and children known to be on treatment 12 months after initiation of antiretroviral therapy</td>
<td>57.14%</td>
<td>87%</td>
<td>80%</td>
<td>130.62%</td>
<td>100%</td>
</tr>
<tr>
<td>4. Percentage of female sex workers reporting the use of a condom during penetrative sex with their most recent client</td>
<td>68.5%</td>
<td>63%</td>
<td>75%</td>
<td>-84.62%</td>
<td>0</td>
</tr>
<tr>
<td>5. Percentage of men reporting the use of a condom the last time they had anal sex with a male partner</td>
<td>51.6%</td>
<td>67%</td>
<td>55%</td>
<td>452.94%</td>
<td>100%</td>
</tr>
<tr>
<td>6. Percentage of uniformed personnel reporting the use of condom last time they had sex with a non-regular partner</td>
<td>21.8%</td>
<td>53%</td>
<td>28%</td>
<td>503.23%</td>
<td>100%</td>
</tr>
<tr>
<td>7. Percentage of men who have sex with men (MSM) surveyed who both correctly identify ways of preventing the sexual transmission of HIV and who reject major misconceptions about HIV transmission</td>
<td>48.7%</td>
<td>11%</td>
<td>58%</td>
<td>-405.38%</td>
<td>0</td>
</tr>
<tr>
<td>8. Percentage of female sex workers surveyed who both correctly identify ways of preventing the sexual transmission of HIV and who reject major misconceptions about HIV transmission</td>
<td>56.5%</td>
<td>38%</td>
<td>66%</td>
<td>-194.74%</td>
<td>0</td>
</tr>
<tr>
<td>9. Number of men who have sex with men (MSM) reached through a basic package of HIV/AIDS prevention services</td>
<td>1250</td>
<td>1095</td>
<td>1050</td>
<td>129.03%</td>
<td>100%</td>
</tr>
<tr>
<td>10. Number of Female Sex Workers (FSW) reached through a basic package of HIV/AIDS prevention services</td>
<td>840</td>
<td>815</td>
<td>800</td>
<td>160%</td>
<td>100%</td>
</tr>
<tr>
<td>11. Number of people who receive HIV testing and counseling including provision of test results</td>
<td>1480</td>
<td>13310</td>
<td>16500</td>
<td>78.76%</td>
<td>78.76%</td>
</tr>
<tr>
<td>12. Number and percentage of HIV-positive pregnant women who received anti-retroviral to reduce the risk of mother-to-child transmission during the last 12 months</td>
<td>75%</td>
<td>78%</td>
<td>100%</td>
<td>12%</td>
<td>12%</td>
</tr>
<tr>
<td>13. Number and percentage of health facilities providing STI case management services as per national guidelines</td>
<td>30</td>
<td>61</td>
<td>80</td>
<td>62%</td>
<td>62%</td>
</tr>
<tr>
<td>14. Number of adults and children with advanced HIV infection currently receiving antiretroviral treatment</td>
<td>110</td>
<td>140</td>
<td>298</td>
<td>15.96%</td>
<td>15.96%</td>
</tr>
<tr>
<td>15. Number and percentage of PLHIVs (who know their HIV status) reached with a basic package of care and support services</td>
<td>30%</td>
<td>47%</td>
<td>74%</td>
<td>38.64%</td>
<td>38.64%</td>
</tr>
<tr>
<td>16. Number and percentage of blood bank and laboratory facilities conducting HIV screening/ testing as per national guidelines</td>
<td>12</td>
<td>13</td>
<td>13</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>17. Number of civil society/community based organizations received capacity development support</td>
<td>3</td>
<td>10</td>
<td>10</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Overall performance against targets and baseline 65.14%

Source: Calculated from Grant Performance Report, 22 October 2014 (The Global Fund, 2014)
It can be seen that out of a total 17 indicators, the NAP exceeded the targets for 7 indicators and achieved targets by 100% for 2 indicators. However, the Programme’s performance against 3 indicators was worse than their baseline. For indicator 9 and 10 (number of MSM and female sex workers reached), the targets were set lower than the baseline, as the Programme restarted its calculation of coverage at the beginning of 2012 with the start of a new grant (Round 10 HIV/AIDS Grant) from the Global Fund, whereas the baseline was set based on the achievements during the previous grant period. With the adjustments of all achievements to be within 0% to 100%, the overall performance of the NAP came out as 65.14%, which is comparable to the ‘B1’ (Adequate) performance rating given by the Global Fund for the period ended on 31 December 2013 (The Global Fund, 2014).

**Project results for IPL:**
Unlike HSSP-SP and the NAP, the performance framework for IPL had only targets set for the coverage indicators for DPT3 and Measles vaccinations and did not have any targets set for other process indicators. To measure the performance of IPL, the average vaccination coverage change from the baseline (i.e. for the year 2010) to the end of the project (i.e. end of September 2013) period in both groups of districts which were covered by IPL activities (IPL focus districts) and which were not covered by IPL activities (Non-IPL focus districts) were compared. The differences of these changes were then compared with the likely difference of changes in case the full coverage targets were achieved by IPL. Table 9.9 presents these calculations and results:
Table 9.9: Attributable vaccination coverage and achievement of targets by IPL

<table>
<thead>
<tr>
<th></th>
<th>Baseline % of coverage of DPT3 (2010)</th>
<th>Average % of coverage of DPT3 end of Sept 2013</th>
<th>% change in DPT3 coverage</th>
<th>Baseline % of coverage of Measles (2010)</th>
<th>Average % of coverage of Measles end of Sept 2013</th>
<th>% change in measles coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPL focus districts</td>
<td>67.92%</td>
<td>70.78%</td>
<td>2.86%</td>
<td>59.11%</td>
<td>68.43%</td>
<td>9.32%</td>
</tr>
<tr>
<td>Non-IPL focus districts</td>
<td>81.65%</td>
<td>68.31%</td>
<td>-13.34%</td>
<td>81.03%</td>
<td>61.72%</td>
<td>-19.31%</td>
</tr>
<tr>
<td>% of change attributable to IPL for results</td>
<td>-13.73%</td>
<td>2.47%</td>
<td>16.20%</td>
<td>-21.92%</td>
<td>6.71%</td>
<td>28.63%</td>
</tr>
<tr>
<td>Targets for IPL focus districts</td>
<td>67.92%</td>
<td>81.50%</td>
<td>13.58%</td>
<td>59.11%</td>
<td>81.50%</td>
<td>22.39%</td>
</tr>
<tr>
<td>Non-IPL focus districts</td>
<td>81.65%</td>
<td>68.31%</td>
<td>-13.34%</td>
<td>81.03%</td>
<td>61.72%</td>
<td>-19.31%</td>
</tr>
<tr>
<td>% of change to be attributable to IPL for targets</td>
<td>-13.73%</td>
<td>13.19%</td>
<td>26.92%</td>
<td>-21.92%</td>
<td>19.78%</td>
<td>41.70%</td>
</tr>
<tr>
<td>Achievement of targets for DPT3 (i.e. 16.20/26.92)</td>
<td>60.18%</td>
<td>Achievement of targets for Measles (i.e. 28.63/41.70)</td>
<td>68.67%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall performance for DPT3 and Measles</td>
<td>64.42%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Calculated from IPL Final Report, 2013

As discussed in Chapter 8, vaccination coverage in IPL focus districts increased from their baseline, whereas the coverage in non-IPL focus districts decreased giving a percentage of increased vaccination coverage of 16.20% for DPT3 and 28.63% for Measles attributable to IPL performance. However, IPL still fell short of its targets and a percentage change of 26.92% for DPT3 and 41.70% for Measles vaccination would be attributable to IPL performance if IPL had achieved its targets fully. This gave IPL an overall 64.42% achievement of its targets for DPT3 and Measles vaccination coverage.

Health system outputs:

As explained in the methods section the extent of health system related outputs by each project was calculated from the stakeholder interviews in Chapter 7. Table 9.10 provides the summary of responses from the stakeholder interviews in this regard:

Table 9.10: Extent of health system outputs achieved by each project

<table>
<thead>
<tr>
<th>Project</th>
<th>Total score</th>
<th>Number of respondents</th>
<th>Mean score</th>
<th>Standard deviation</th>
<th>Health system output performance ratio (mean score : highest point in scale)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSSP-SP</td>
<td>38</td>
<td>13</td>
<td>2.92</td>
<td>0.732</td>
<td>58.46%</td>
</tr>
<tr>
<td>NAP</td>
<td>43</td>
<td>12</td>
<td>3.58</td>
<td>0.821</td>
<td>71.67%</td>
</tr>
<tr>
<td>IPL</td>
<td>35</td>
<td>11</td>
<td>3.18</td>
<td>0.982</td>
<td>63.64%</td>
</tr>
</tbody>
</table>

Source: Chapter 7, Table 7.3
Question 13 of the stakeholder interview asked the respondents their views on the percentage of achievement of health system outputs by each project against their targets. The health systems outputs were related to one or more areas of health service delivery, health workforce, health information, health supplies and commodities, health financing, and governance. It appeared that the stakeholders viewed the health system related outputs by the NAP (71.7%) more favourably than those of IPL (63.6%) and HSSP-SP (58.5%) in terms of the average scores each project obtained.

**Use of allocated resources by three projects:**

Table 9.11 provides a summary of use of allocated resources by three projects:

<table>
<thead>
<tr>
<th>Project</th>
<th>Total budget (USD)</th>
<th>Total expenditure (USD)</th>
<th>Use of allocated resources</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSSP-SP</td>
<td>$21,000,000</td>
<td>$16,240,000</td>
<td>77.33%</td>
</tr>
<tr>
<td>NAP</td>
<td>$14,480,049</td>
<td>$11,842,703</td>
<td>81.97%</td>
</tr>
<tr>
<td>IPL</td>
<td>$2,639,250</td>
<td>$2,639,250</td>
<td>100%</td>
</tr>
</tbody>
</table>

While it was assumed that IPL utilized 100% of its allocated budget by the end of its project period, use of allocated by HSSP-SP was 77.33% by 31 December 2012 and 81.97% by NAP by 31 December 2013 against their budgeted amounts.

Dividing the 54.62%, 65.14% and 64.42% achievements of project results of HSSP-SP, NAP and IPL by their respective use of allocated resources a cost efficiency score can be calculated for achieving results by each project. These cost efficiency scores are 70.63% for HSSP-SP with their 77.33% use of resources; 79.47% for the NAP with 81.97% use of resources; and 64.42% for IPL with their 100% use of allocated resources.

**Calculating relative efficiencies with DEA model:**

The DEA model considered the data on percentage of achievement of project targets, percentage of achievement of health system related outputs, and percentage of use of allocated resources by three projects to calculate the relative efficiencies of each project. Table 9.12 provides the results of DEA for relative performance efficiency of three projects:
### Table 9.12: DEA result for performance efficiency of three projects

<table>
<thead>
<tr>
<th>Project</th>
<th>Outputs (achievement of targets)</th>
<th>Input</th>
<th>Relative efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Project results</td>
<td>Health system strengthening</td>
<td>Budget used</td>
</tr>
<tr>
<td>HSSP-SP</td>
<td>54.62%</td>
<td>58.46%</td>
<td>77.33%</td>
</tr>
<tr>
<td>NAP</td>
<td>65.14%</td>
<td>71.66%</td>
<td>81.97%</td>
</tr>
<tr>
<td>IPL</td>
<td>64.42%</td>
<td>63.63%</td>
<td>100%</td>
</tr>
</tbody>
</table>

It can be seen that with 65.14% achievement of project targets, 71.66% ranking of health system strengthening outputs and 81.97% use of allocated budget, the NAP appeared to be 100% efficient in relation to the performance of HSSP-SP and IPL.

Compared to the NAP and IPL, the relative efficiency of HSSP-SP was 88.89%. Although IPL’s project results outputs (64.42%) and health system strengthening output ranking (63.63%) were higher than those of HSSP-SP, due to the 100% use of allocated budget, IPL’s relative efficiency (81.06%) came out as lower than that of HSSP-SP.

These DEA results can be compared with the findings from the stakeholder interviews. Question 14 of the stakeholders’ interview questionnaire asked the respondents to rank these three projects for their performance on efficiency of productivity measure in a scale of 1 to 5. Scoring 1 represented the lowest level of performance with “no efficiencies and very poor input-output ratios” and 5 represented the highest level of performance with “Very efficient with high productivity ratio and performance expected”. Table 9.13 provides the summary of responses from the stakeholders on this item:

### Table 9.13: Stakeholders’ views on efficiency of productivity measure of three projects

<table>
<thead>
<tr>
<th>Project</th>
<th>Total score</th>
<th>Number of respondents</th>
<th>Mean score</th>
<th>Standard deviation</th>
<th>Efficiency of productivity ratio (mean score : highest point in scale)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSSP-SP</td>
<td>38</td>
<td>13</td>
<td>2.92</td>
<td>0.641</td>
<td>58.46%</td>
</tr>
<tr>
<td>NAP</td>
<td>57.5</td>
<td>15</td>
<td>3.83</td>
<td>0.523</td>
<td>76.67%</td>
</tr>
<tr>
<td>IPL</td>
<td>39.5</td>
<td>11</td>
<td>3.59</td>
<td>1.158</td>
<td>71.82%</td>
</tr>
</tbody>
</table>

Source: Chapter 7, Table 7.3

It can be seen that the relatively highest productivity efficiency assigned by the stakeholders to the NAP (76.67%) is consistent with the DEA results in Table 9.12 that showed 100% performance efficiency for the NAP. However, it seems the stakeholders viewed the performance of IPL more favourably than that of HSSP-SP, whereas the DEA
results showed 90.92% efficiency for HSSP-SP and 80.74% for IPL. This was probably because of the fact that the outputs of a vaccination project such as IPL were more directly visible to the stakeholders than the outputs of a sector wide health systems strengthening project such as HSSP-SP.

9.4.3 Effectiveness

As explained in the Methods section, the evaluation of the effectiveness of the three projects considered both the project outcomes and health system outcomes. The project outcomes were measured by the DALYs saved by each project as calculated in Chapter 8 (reference Table 8.34). But the health system outcomes were calculated from the total health system strengthening related investments by each project with the application of health system strengthening outcome scores obtained from the stakeholders’ interviews. Table 9.14 provides the summary of stakeholder interviews on health system outcomes:

<table>
<thead>
<tr>
<th>Project</th>
<th>Total score</th>
<th>Number of respondents</th>
<th>Mean score</th>
<th>Standard deviation</th>
<th>Health system outcome performance ratio (mean score : highest point in scale)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSSP-SP</td>
<td>40</td>
<td>13</td>
<td>3.08</td>
<td>0.760</td>
<td>61.54%</td>
</tr>
<tr>
<td>NAP</td>
<td>54</td>
<td>15</td>
<td>3.60</td>
<td>0.660</td>
<td>72.00%</td>
</tr>
<tr>
<td>IPL</td>
<td>38</td>
<td>11</td>
<td>3.45</td>
<td>0.907</td>
<td>69.09%</td>
</tr>
</tbody>
</table>

The calculation of total health system strengthening related investments considered the cost items such as technical assistance, civil work or infrastructure, training, and durable goods and equipment whose benefits would last beyond the project period with likely benefits to more than one disease component. The health system outcome scores in Table 9.14 were then applied to the total health system strengthening (HSS) investments by each project to covert the amounts to health system outcome points.

Table 9.15 summarizes the estimated investments and health system outcome points for each project:
Table 9.15: Health system strengthening investments and health system outcome points

<table>
<thead>
<tr>
<th></th>
<th>HSSP-SP</th>
<th>NAP</th>
<th>IPL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technical assistance</td>
<td>$4,883,423</td>
<td>$4,109,267</td>
<td>$887,052</td>
</tr>
<tr>
<td>Infrastructure, durable goods and equipment</td>
<td>$3,151,907</td>
<td>$815,357</td>
<td>$37,776</td>
</tr>
<tr>
<td>Training</td>
<td>$3,443,745</td>
<td>$501,947</td>
<td>$196,096</td>
</tr>
<tr>
<td>Total HSS Investments</td>
<td>$11,479,075</td>
<td>$5,426,571</td>
<td>$1,120,924</td>
</tr>
<tr>
<td>HSS investment as % of total expenditure</td>
<td>70.68%</td>
<td>45.82%</td>
<td>42.47%</td>
</tr>
<tr>
<td>Health system outcome conversion factor</td>
<td>61.54%</td>
<td>72%</td>
<td>69.09%</td>
</tr>
<tr>
<td>Health system outcome points (HSS investments x conversion factor)</td>
<td>7,064,223</td>
<td>3,907,131</td>
<td>774,446</td>
</tr>
</tbody>
</table>

Source: Author’s analysis of expenditure reports of HSSP-SP, NAP, & IPL

It was found in Chapter 8 (reference Table 8.34) that HSSP-SP spent a total of $16,240,000 and saved 44,628 DALYs. The NAP spent $11,842,703 and saved 39,250 DALYs. IPL, on the other hand, spent $2,639,250 and saved 246,233 DALYs. These data on project outcomes and inputs were inserted in the DEA model along with the data on health system outcome points as obtained in Table 9.15 to calculate the relative effectiveness of each project in terms of DALYs saved and health system outcome points. Table 9.16 presents the results of DEA for effectiveness:

Table 9.16: DEA results for effectiveness of three projects

<table>
<thead>
<tr>
<th></th>
<th>Outcomes</th>
<th>Input used</th>
<th>Relative Effectiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DALYs saved</td>
<td>HSS outcome points</td>
<td>USD</td>
</tr>
<tr>
<td>Project</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HSSP-SP</td>
<td>44,628</td>
<td>7,064,223</td>
<td>$16,240,000</td>
</tr>
<tr>
<td>NAP</td>
<td>39,250.21</td>
<td>3,907,131</td>
<td>$11,884,703</td>
</tr>
<tr>
<td>IPL</td>
<td>246,232.5</td>
<td>774,446</td>
<td>$2,639,250</td>
</tr>
</tbody>
</table>

The DEA results in this case show that considering the DALYs saved and health system strengthening outcome points, the performance of the NAP was much lower than the performance of HSSP-SP and IPL. The DEA results show a relative effectiveness of 76.28% of the NAP compared to the 100% relative effectiveness of HSSP-SP and IPL.

The DEA results on effectiveness can be further compared with the stakeholders’ views on the effectiveness of the three projects as obtained in Chapter 7 (reference Table 7.3). In order to gather stakeholders’ views on the effectiveness of three projects, in Chapter 7 respondents were asked two questions: Question 19 of the interview questionnaire asked the respondents to rank the three projects for their effectiveness in terms of relevance and robustness of evidence in a scale of 1 to 5. Scoring 1 represented the lowest level of effectiveness with “Indicators and achievements are neither relevant nor
robust” and 5 represented the highest level of effectiveness with “Indicators and achievements are relevant and robust”.

In a similar manner Question 20 of the stakeholders’ interview questionnaire asked the respondents to rank the three projects for their effectiveness of theory of change behind the project design in a scale of 1 to 5. Scoring 1 represented the lowest level of effectiveness with “Little or no likelihood that outputs will deliver purpose” and 5 represented the highest level of effectiveness with “Outputs are necessary and sufficient to deliver purpose”. Table 9.17 provides the summary of responses from the stakeholders on these two questions:

**Table 9.17: Stakeholders’ views on project effectiveness**

<table>
<thead>
<tr>
<th>Project</th>
<th>Question</th>
<th>Total score</th>
<th>Number of respondents</th>
<th>Mean score</th>
<th>Standard deviation</th>
<th>Effectiveness ratio (mean score : highest point in scale)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSSP-SP</td>
<td>19. Relevance &amp; robustness</td>
<td>45.5</td>
<td>13</td>
<td>3.50</td>
<td>0.764</td>
<td>70%</td>
</tr>
<tr>
<td></td>
<td>20. Theory of change</td>
<td>41</td>
<td>13</td>
<td>3.15</td>
<td>0.875</td>
<td>63.08%</td>
</tr>
<tr>
<td>NAP</td>
<td>19. Relevance &amp; robustness</td>
<td>65</td>
<td>15</td>
<td>4.33</td>
<td>0.523</td>
<td>86.67%</td>
</tr>
<tr>
<td></td>
<td>20. Theory of change</td>
<td>60</td>
<td>15</td>
<td>4.00</td>
<td>0.655</td>
<td>80%</td>
</tr>
<tr>
<td>IPL</td>
<td>19. Relevance &amp; robustness</td>
<td>47</td>
<td>11</td>
<td>4.27</td>
<td>0.786</td>
<td>85.45%</td>
</tr>
<tr>
<td></td>
<td>20. Theory of change</td>
<td>42.5</td>
<td>11</td>
<td>3.86</td>
<td>0.636</td>
<td>77.27%</td>
</tr>
</tbody>
</table>

Source: Chapter 7, Table 7.3

These views provided by the stakeholders seem to be quite different from the DEA results in Table 9.16 that assigned lowest (74.44%) relative effectiveness to the NAP compared to 100% effectiveness of HSSP-SP and IPL. This difference is probably because of different considerations behind assessing the effectiveness. The DEA model considered the value of project outcomes such as DALYs saved and weighted health system strengthening investments by each project, whereas the stakeholder interviews considered the extent of project outputs with respect to delivering the project purpose.
9.4.4 Summary of DEA results

Table 9.18 below summarizes the DEA results for economy of procurement, economy of human resources and technical assistance, performance efficiency, and effectiveness of the three projects compared:

<table>
<thead>
<tr>
<th>Project</th>
<th>Economy of procuring goods &amp; equipment</th>
<th>Economy of procuring human resources &amp; technical assistance</th>
<th>Performance efficiency for project outputs &amp; health system outputs</th>
<th>Effectiveness for project outcomes &amp; health system outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSSP-SP</td>
<td>87.2%</td>
<td>100%</td>
<td>88.89%</td>
<td>100%</td>
</tr>
<tr>
<td>NAP</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>76.28%</td>
</tr>
<tr>
<td>IPL</td>
<td>100%</td>
<td>100%</td>
<td>81.06%</td>
<td>100%</td>
</tr>
</tbody>
</table>

It can be observed that even with relatively lower efficiency in procurement of goods and equipment than the NAP and IPL and relatively lower performance efficiency than the NAP, HSSP-SP came out as more effective than the NAP in saving DALYs and achieving health system strengthening outcomes. IPL was also more effective than the NAP in saving DALYs and achieving health system outcomes.

Although the NAP came out as comparatively more efficient than HSSP-SP and IPL in procurement and performance achievements, its effectiveness in saving DALYs and achieving health system outcomes was dominated by the performance of HSSP-SP and IPL on these aspects.

9.5 Looking for correlations

With all these findings from the stakeholder interviews, cost effectiveness analysis, and DEA, correlations were explored between the Paris principles of aid effectiveness and any other aspects of the projects. The other aspects of the projects included a range of dependent variables such as economy, efficiency, effectiveness, technical efficiency of DALYs saved, achievements of project targets, health system outputs, health system outcomes, cost per year, total expenditure, total health system investment, percentage of use of resources, and cost efficiency of achieving results. The purpose was to investigate if any of these project aspects were influence by the degree of adherence to the Paris principles by the projects.
9.5.1 Data sources

For the purpose of defining the extent of compliance with the Paris principles of aid effectiveness, the correlation analysis used the results from the stakeholder interviews in Chapter 7 for different aspects of Paris principles such as ownership, alignment, harmonization, managing for results, mutual accountability and overall score for Paris principles. The correlations analysis also used the stakeholder interview results for scores on economy, efficiency, effectiveness, sustainability, health system output, and health system outcome.

The stakeholder interviews results in Chapter 7, however, were converted from the weighted scores (i.e. average score of each question multiplied by their importance weight) to an equal scale of 100 for all domains such as the Paris principles, economy, efficiency, effectiveness and sustainability. In order to remove the effect of this conversion to scale, the correlation analysis used the original weighted scores (i.e. mean score multiplied by the importance weight) from the stakeholder interviews.

For the rest of the variables such as technical efficiency of DALYs saved, achievements of project targets, cost per year, total expenditure, total health system investment, percentage of use of resources, and cost efficiency of achieving results, the results from the cost-effectiveness analysis in Chapter 8 and calculations based on project documents in this current chapter were used.

Table 9.19 provides the list of variables, their respective values, and data sources used for running the correlations analysis:
### Table 9.19: Variables considered for correlations analysis

<table>
<thead>
<tr>
<th>Variable</th>
<th>Total score</th>
<th>HSSP-SP</th>
<th>NAP</th>
<th>IPL</th>
<th>Data source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ownership</td>
<td>50</td>
<td>28.85</td>
<td>38.83</td>
<td>30.00</td>
<td>Stakeholder interviews (Chapter 7)</td>
</tr>
<tr>
<td>Alignment</td>
<td>135</td>
<td>84.04</td>
<td>94.13</td>
<td>74.05</td>
<td>Stakeholder interviews (Chapter 7)</td>
</tr>
<tr>
<td>Harmonization</td>
<td>45</td>
<td>31.15</td>
<td>31.20</td>
<td>36.20</td>
<td>Stakeholder interviews (Chapter 7)</td>
</tr>
<tr>
<td>Managing Results</td>
<td>50</td>
<td>26.15</td>
<td>39.00</td>
<td>36.36</td>
<td>Stakeholder interviews (Chapter 7)</td>
</tr>
<tr>
<td>Accountability</td>
<td>45</td>
<td>33.58</td>
<td>37.50</td>
<td>25.98</td>
<td>Stakeholder interviews (Chapter 7)</td>
</tr>
<tr>
<td>Paris Principles</td>
<td>325</td>
<td>203.77</td>
<td>240.67</td>
<td>202.59</td>
<td>Stakeholder interviews (Chapter 7)</td>
</tr>
<tr>
<td>Economy</td>
<td>130</td>
<td>73.46</td>
<td>87.70</td>
<td>87.75</td>
<td>Stakeholder interviews (Chapter 7)</td>
</tr>
<tr>
<td>Efficiency</td>
<td>225</td>
<td>124.62</td>
<td>156.33</td>
<td>151.77</td>
<td>Stakeholder interviews (Chapter 7)</td>
</tr>
<tr>
<td>Effectiveness</td>
<td>165</td>
<td>101.04</td>
<td>128.84</td>
<td>126.27</td>
<td>Stakeholder interviews (Chapter 7)</td>
</tr>
<tr>
<td>Sustainability</td>
<td>90</td>
<td>51.35</td>
<td>46.13</td>
<td>59.82</td>
<td>Stakeholder interviews (Chapter 7)</td>
</tr>
<tr>
<td>Cost effectiveness efficiency</td>
<td>100</td>
<td>53.62%</td>
<td>67.77%</td>
<td>53.04%</td>
<td>Cost effectiveness analysis (Chapter 8)</td>
</tr>
<tr>
<td>Achievement of results</td>
<td>100</td>
<td>54.62%</td>
<td>65.14%</td>
<td>64.42%</td>
<td>Project results (Chapter 9)</td>
</tr>
<tr>
<td>Health System outputs</td>
<td>100</td>
<td>58.46%</td>
<td>71.67%</td>
<td>63.64%</td>
<td>Stakeholder interviews (Chapter 7)</td>
</tr>
<tr>
<td>Health System outcomes</td>
<td>100</td>
<td>61.54%</td>
<td>72.00%</td>
<td>69.09%</td>
<td>Stakeholder interviews (Chapter 7)</td>
</tr>
<tr>
<td>Cost per year</td>
<td>100</td>
<td>70.68%</td>
<td>45.82%</td>
<td>42.47%</td>
<td>Project financial reports (Chapter 8)</td>
</tr>
<tr>
<td>% of investments for Health System</td>
<td>100</td>
<td>77.33%</td>
<td>81.97%</td>
<td>100.00%</td>
<td>Project financial reports (Chapter 9)</td>
</tr>
<tr>
<td>Total expenditure</td>
<td>Thousand USD</td>
<td>$3,608.89</td>
<td>$1,821.95</td>
<td>$1,055.70</td>
<td>Project financial reports (Chapter 8)</td>
</tr>
<tr>
<td>Use of resources</td>
<td>100</td>
<td>70.68%</td>
<td>45.82%</td>
<td>42.47%</td>
<td>Project financial reports (Chapter 9)</td>
</tr>
<tr>
<td>% of investments for Health System</td>
<td>100</td>
<td>77.33%</td>
<td>81.97%</td>
<td>100.00%</td>
<td>Project financial reports (Chapter 8)</td>
</tr>
<tr>
<td>Total expenditure</td>
<td>Thousand USD</td>
<td>$16,240</td>
<td>$11,842.70</td>
<td>$2,639.25</td>
<td>Project financial reports (Chapter 8)</td>
</tr>
<tr>
<td>Health System expenditure</td>
<td>Thousand USD</td>
<td>$11,479.08</td>
<td>$5,426.57</td>
<td>$1,120.92</td>
<td>Project financial reports (Chapter 9)</td>
</tr>
<tr>
<td>Cost efficiency of Results achieved</td>
<td>100</td>
<td>70.63%</td>
<td>79.47%</td>
<td>64.42%</td>
<td>Project results and financial reports (Chapter 9)</td>
</tr>
</tbody>
</table>

### 9.5.2 Method and results

Running Spearman rank-order correlations with the above variables and values gave the following results:
<table>
<thead>
<tr>
<th>Sample size</th>
<th>Ownership</th>
<th>Alignment</th>
<th>Harmonization</th>
<th>Managing</th>
<th>Accountability</th>
<th>Paris Principles</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>Spearman Correlation Coefficient</td>
<td>0.5</td>
<td>0.75</td>
<td>0.57735</td>
<td>0.66667</td>
<td>accepted</td>
</tr>
<tr>
<td>0.0</td>
<td>R Standard Error</td>
<td>0.75</td>
<td>0.57735</td>
<td>0.66667</td>
<td>0.57735</td>
<td>accepted</td>
</tr>
<tr>
<td>0.5</td>
<td>p-value</td>
<td>0.66667</td>
<td>0.66667</td>
<td>0.66667</td>
<td>0.66667</td>
<td>accepted</td>
</tr>
<tr>
<td>0.0</td>
<td>(2%)</td>
<td>accepted</td>
<td>accepted</td>
<td>accepted</td>
<td>accepted</td>
<td>accepted</td>
</tr>
<tr>
<td>0.0</td>
<td>Harmonization</td>
<td>Spearman Correlation Coefficient</td>
<td>0.5</td>
<td>0.57735</td>
<td>0.66667</td>
<td>accepted</td>
</tr>
<tr>
<td>0.5</td>
<td>R Standard Error</td>
<td>0.57735</td>
<td>0.66667</td>
<td>0.66667</td>
<td>0.66667</td>
<td>accepted</td>
</tr>
<tr>
<td>0.0</td>
<td>p-value</td>
<td>0.66667</td>
<td>0.66667</td>
<td>0.66667</td>
<td>0.66667</td>
<td>accepted</td>
</tr>
<tr>
<td>0.0</td>
<td>(2%)</td>
<td>accepted</td>
<td>accepted</td>
<td>accepted</td>
<td>accepted</td>
<td>accepted</td>
</tr>
<tr>
<td>Accountability</td>
<td>Spearman Correlation Coefficient</td>
<td>0.5</td>
<td>0.75</td>
<td>0.57735</td>
<td>0.66667</td>
<td>accepted</td>
</tr>
<tr>
<td>0.5</td>
<td>R Standard Error</td>
<td>0.75</td>
<td>0.57735</td>
<td>0.66667</td>
<td>0.66667</td>
<td>accepted</td>
</tr>
<tr>
<td>0.0</td>
<td>p-value</td>
<td>0.66667</td>
<td>0.66667</td>
<td>0.66667</td>
<td>0.66667</td>
<td>accepted</td>
</tr>
<tr>
<td>0.0</td>
<td>(2%)</td>
<td>accepted</td>
<td>accepted</td>
<td>accepted</td>
<td>accepted</td>
<td>accepted</td>
</tr>
<tr>
<td>Paris Principles</td>
<td>Spearman Correlation Coefficient</td>
<td>0.5</td>
<td>0.75</td>
<td>0.57735</td>
<td>0.66667</td>
<td>accepted</td>
</tr>
<tr>
<td>0.5</td>
<td>R Standard Error</td>
<td>0.75</td>
<td>0.57735</td>
<td>0.66667</td>
<td>0.66667</td>
<td>accepted</td>
</tr>
<tr>
<td>0.0</td>
<td>p-value</td>
<td>0.66667</td>
<td>0.66667</td>
<td>0.66667</td>
<td>0.66667</td>
<td>accepted</td>
</tr>
<tr>
<td>0.0</td>
<td>(2%)</td>
<td>accepted</td>
<td>accepted</td>
<td>accepted</td>
<td>accepted</td>
<td>accepted</td>
</tr>
<tr>
<td>Sustainability</td>
<td>Spearman Correlation Coefficient</td>
<td>0.5</td>
<td>0.75</td>
<td>0.57735</td>
<td>0.66667</td>
<td>accepted</td>
</tr>
<tr>
<td>0.5</td>
<td>R Standard Error</td>
<td>0.75</td>
<td>0.57735</td>
<td>0.66667</td>
<td>0.66667</td>
<td>accepted</td>
</tr>
<tr>
<td>0.0</td>
<td>p-value</td>
<td>0.66667</td>
<td>0.66667</td>
<td>0.66667</td>
<td>0.66667</td>
<td>accepted</td>
</tr>
<tr>
<td>0.0</td>
<td>(2%)</td>
<td>accepted</td>
<td>accepted</td>
<td>accepted</td>
<td>accepted</td>
<td>accepted</td>
</tr>
<tr>
<td>Arrival</td>
<td>Spearman Correlation Coefficient</td>
<td>0.5</td>
<td>0.75</td>
<td>0.57735</td>
<td>0.66667</td>
<td>accepted</td>
</tr>
<tr>
<td>0.5</td>
<td>R Standard Error</td>
<td>0.75</td>
<td>0.57735</td>
<td>0.66667</td>
<td>0.66667</td>
<td>accepted</td>
</tr>
<tr>
<td>0.0</td>
<td>p-value</td>
<td>0.66667</td>
<td>0.66667</td>
<td>0.66667</td>
<td>0.66667</td>
<td>accepted</td>
</tr>
<tr>
<td>0.0</td>
<td>(2%)</td>
<td>accepted</td>
<td>accepted</td>
<td>accepted</td>
<td>accepted</td>
<td>accepted</td>
</tr>
<tr>
<td>Health system</td>
<td>Spearman Correlation Coefficient</td>
<td>0.5</td>
<td>0.75</td>
<td>0.57735</td>
<td>0.66667</td>
<td>accepted</td>
</tr>
<tr>
<td>0.5</td>
<td>R Standard Error</td>
<td>0.75</td>
<td>0.57735</td>
<td>0.66667</td>
<td>0.66667</td>
<td>accepted</td>
</tr>
<tr>
<td>0.0</td>
<td>p-value</td>
<td>0.66667</td>
<td>0.66667</td>
<td>0.66667</td>
<td>0.66667</td>
<td>accepted</td>
</tr>
<tr>
<td>0.0</td>
<td>(2%)</td>
<td>accepted</td>
<td>accepted</td>
<td>accepted</td>
<td>accepted</td>
<td>accepted</td>
</tr>
<tr>
<td>Cost per year</td>
<td>Spearman Correlation Coefficient</td>
<td>0.5</td>
<td>0.75</td>
<td>0.57735</td>
<td>0.66667</td>
<td>accepted</td>
</tr>
<tr>
<td>0.5</td>
<td>R Standard Error</td>
<td>0.75</td>
<td>0.57735</td>
<td>0.66667</td>
<td>0.66667</td>
<td>accepted</td>
</tr>
<tr>
<td>0.0</td>
<td>p-value</td>
<td>0.66667</td>
<td>0.66667</td>
<td>0.66667</td>
<td>0.66667</td>
<td>accepted</td>
</tr>
<tr>
<td>0.0</td>
<td>(2%)</td>
<td>accepted</td>
<td>accepted</td>
<td>accepted</td>
<td>accepted</td>
<td>accepted</td>
</tr>
<tr>
<td>% of investments</td>
<td>Spearman Correlation Coefficient</td>
<td>0.5</td>
<td>0.75</td>
<td>0.57735</td>
<td>0.66667</td>
<td>accepted</td>
</tr>
<tr>
<td>0.5</td>
<td>R Standard Error</td>
<td>0.75</td>
<td>0.57735</td>
<td>0.66667</td>
<td>0.66667</td>
<td>accepted</td>
</tr>
<tr>
<td>0.0</td>
<td>p-value</td>
<td>0.66667</td>
<td>0.66667</td>
<td>0.66667</td>
<td>0.66667</td>
<td>accepted</td>
</tr>
<tr>
<td>0.0</td>
<td>(2%)</td>
<td>accepted</td>
<td>accepted</td>
<td>accepted</td>
<td>accepted</td>
<td>accepted</td>
</tr>
<tr>
<td>Use of resources</td>
<td>Spearman Correlation Coefficient</td>
<td>0.5</td>
<td>0.75</td>
<td>0.57735</td>
<td>0.66667</td>
<td>accepted</td>
</tr>
<tr>
<td>0.5</td>
<td>R Standard Error</td>
<td>0.75</td>
<td>0.57735</td>
<td>0.66667</td>
<td>0.66667</td>
<td>accepted</td>
</tr>
<tr>
<td>0.0</td>
<td>p-value</td>
<td>0.66667</td>
<td>0.66667</td>
<td>0.66667</td>
<td>0.66667</td>
<td>accepted</td>
</tr>
<tr>
<td>0.0</td>
<td>(2%)</td>
<td>accepted</td>
<td>accepted</td>
<td>accepted</td>
<td>accepted</td>
<td>accepted</td>
</tr>
<tr>
<td>Health system</td>
<td>Spearman Correlation Coefficient</td>
<td>0.5</td>
<td>0.75</td>
<td>0.57735</td>
<td>0.66667</td>
<td>accepted</td>
</tr>
<tr>
<td>0.5</td>
<td>R Standard Error</td>
<td>0.75</td>
<td>0.57735</td>
<td>0.66667</td>
<td>0.66667</td>
<td>accepted</td>
</tr>
<tr>
<td>0.0</td>
<td>p-value</td>
<td>0.66667</td>
<td>0.66667</td>
<td>0.66667</td>
<td>0.66667</td>
<td>accepted</td>
</tr>
<tr>
<td>0.0</td>
<td>(2%)</td>
<td>accepted</td>
<td>accepted</td>
<td>accepted</td>
<td>accepted</td>
<td>accepted</td>
</tr>
</tbody>
</table>
It can be observed from the above table that ownership seems to be correlated with a number of aspects such as managing for results, efficiency, effectiveness, achievement of results, health system outputs, and health system outcomes. Alignment seems to be correlated with the overall Paris principle scores as well as with the harmonization, mutual accountability, cost effectiveness efficiency (cost efficiency in saving DALYs), and cost efficiency in achieving results. Harmonization also seems to be correlated with the overall Paris principle scores as well as with mutual accountability, cost efficiency in saving DALYs, and cost efficiency in achieving results. The principle of managing for results seems to be correlated with efficiency, effectiveness, achievement of results, health system outputs, and health system outcomes.

Like the principle of harmonization, the principle of mutual accountability seems to be correlated with the overall Paris principle scores as well as with the mutual accountability, cost efficiency in saving DALYs, and cost efficiency in achieving results. The overall score on compliance with the Paris principles seems to be correlated with cost effectiveness efficiency in saving DALYs and cost efficiency in achieving results. However, the scores on sustainability appear to be negatively correlated with the principles of alignment, harmonization, mutual accountability, and overall compliance with the Paris principles.

Apart from the results for sustainability, which was influenced by the poor score obtained by the NAP on this aspect, other correlations seem to be consistent with the expectations behind the Paris principles that better compliance with the Paris principles would ultimately produce better results at lower costs.

One caveat of these correlations, however, is the very small sample size of only 3 projects and the non-availability of the t-test and p-values in perfect correlations (i.e. with +1 and -1 correlation coefficients where the pair of variables are not the same) cases. This makes it impossible to ascertain whether these correlations were robust enough and were statistically significant to ascertain that they did not happen by chance.

Assuming close proximity of the values of each variable to the parametric nature and possible linear correlations among the variables, the correlations were further
investigated by running a Pearson Linear Correlations with the calculation of standard error, t-test, p-value, and testing of null hypothesis at 2 degree of freedom level. This attempt to find possible linear correlations returned the following results:
Table 9.21: Pearson correlation coefficients for the Paris principles and project performance
Sample size

Alignment

Pearson Correlation
R Standard Error
t
p-value
H0 (2%)
Harmonization Pearson Correlation
R Standard Error
t
p-value
H0 (2%)
Managing
Pearson Correlation
Results
R Standard Error
t
p-value
H0 (2%)
Accountability Pearson Correlation
R Standard Error
t
p-value
H0 (2%)
Paris
Pearson Correlation
Principles
R Standard Error
t
p-value
H0 (2%)
Economy
Pearson Correlation
R Standard Error
t
p-value
H0 (2%)
Efficiency
Pearson Correlation
R Standard Error
t
p-value
H0 (2%)
Effectiveness Pearson Correlation
R Standard Error
t
p-value
H0 (2%)
Sustainability Pearson Correlation
R Standard Error
t
p-value
H0 (2%)
Cost
Pearson Correlation
effectiveness R Standard Error
efficiency
t
p-value
H0 (2%)
Achievement Pearson Correlation
of results
R Standard Error
t
p-value
H0 (2%)
Health System Pearson Correlation
outputs
R Standard Error
t
p-value
H0 (2%)
Health System Pearson Correlation
outcomes
R Standard Error
t
p-value
H0 (2%)
Cost per year Pearson Correlation
R Standard Error
t
p-value
H0 (2%)
% of
Pearson Correlation
investments
R Standard Error
for Health
t
System
p-value
H0 (2%)
Pearson Correlation
R Standard Error
t
p-value
H0 (2%)
Total
Pearson Correlation
expenditure
R Standard Error
t
p-value
H0 (2%)
Health System Pearson Correlation
expenditure
R Standard Error
t
p-value
H0 (2%)
Cost efficiency Pearson Correlation
of Results
R Standard Error
achieved
t
p-value
H0 (2%)
Use of
resources

3 Critical value
(2%)
Ownership
Coefficient

Coefficient

Coefficient

Coefficient

Coefficient

Coefficient

Coefficient

Coefficient

Coefficient

Coefficient

Coefficient

Coefficient

Coefficient

Coefficient

Coefficient

Coefficient

Coefficient

Coefficient

Coefficient

31.82052
Alignment

Harmonization

Managing
Results

Accountability

Paris
Principles

0.81025
0.3435
1.38246
0.39867
accepted
-0.39813
0.84149
-0.43401
0.73932
accepted
0.73445
0.46059
1.08219
0.47488
accepted
0.68842
0.52607
0.94915
0.5166
accepted
0.9912

1.

-0.86022
0.26002
-1.68698
0.34065
accepted
0.19732
0.96106
0.20128
0.87355
accepted
0.98289
0.03393
5.33601
0.11794
accepted
0.88069

0.33015
0.891
0.34976
0.7858
accepted
-0.93943
0.11747
-2.74093
0.22271
accepted
-0.51605

0.01337
0.99982
0.01337
0.99149
accepted
0.63816

0.77837

1.

0.01752
7.48855
0.08451
accepted
0.58596
0.65665
0.7231
0.60144
accepted
0.69072
0.5229
0.9552
0.51459
accepted
0.65399
0.5723
0.86449
0.54619
accepted
-0.72117
0.47992
-1.041
0.48721
accepted
0.99018
0.01955
7.08222
0.0893
accepted
0.63684
0.59443
0.826
0.56048
accepted
0.95704
0.08407
3.3007
0.18728
accepted
0.78455
0.38447
1.26529
0.42578
accepted
-0.32625
0.89356
-0.34514
0.78843
accepted
-0.49705

0.22438
1.85923
0.31415
accepted
-0.00016
1.
-0.00016
0.9999
accepted
0.13584
0.98155
0.13711
0.91325
accepted
0.08651
0.99252
0.08684
0.94486
accepted
-0.99034
0.01922
-7.14321
0.08855
accepted
0.88423
0.21814
1.89322
0.30937
accepted
0.06413
0.99589
0.06426
0.95915
accepted
0.6055
0.63337
0.76083
0.58594
accepted
0.27227
0.92587
0.28296
0.82445
accepted
0.28968
0.91609
0.30265
0.8129
accepted
0.10583

0.73369
-0.60247
0.6548
accepted
0.51006
0.73984
0.59299
0.65925
accepted
0.38834
0.84919
0.42141
0.7461
accepted
0.43359
0.812
0.48117
0.71449
accepted
0.92261
0.14879
2.39184
0.2521
accepted
-0.52248
0.72702
-0.61277
0.65002
accepted
0.45371
0.79415
0.50912
0.7002
accepted
-0.11505
0.98676
-0.11582
0.92659
accepted
0.25644
0.93424
0.26531
0.8349
accepted
-0.73724
0.45647
-1.09119
0.47226
accepted
-0.59809

0.59276
0.82887
0.5594
accepted
0.98031
0.039
4.96404
0.12655
accepted
0.99806
0.00388
16.01406
0.0397
accepted
0.99373
0.01249
8.89053
0.07131
accepted
-0.0595
0.99646
-0.05961
0.9621
accepted
0.63235
0.60014
0.81627
0.56418
accepted
0.99097
0.01797
7.39241
0.0856
accepted
0.89968
0.19058
2.06084
0.2876
accepted
0.99703
0.00594
12.94015
0.0491
accepted
-0.88115
0.22358
-1.8635
0.31355
accepted
-0.95395

0.39414
1.23983
0.43209
accepted
-0.18436
0.96601
-0.18757
0.88196
accepted
-0.04897
0.9976
-0.04903
0.96881
accepted
-0.09848
0.9903
-0.09896
0.93721
accepted
-0.99893
0.00213
-21.64467
0.02939
accepted
0.78307
0.3868
1.25909
0.42731
accepted
-0.12079
0.98541
-0.12168
0.92291
accepted
0.44855
0.79881
0.50187
0.70388
accepted
0.09037
0.99183
0.09074
0.94239
accepted
0.46102
0.78746
0.51953
0.69497
accepted
0.28718

0.47355
0.77575
0.53765
0.68595
accepted
0.58893
0.65316
0.72871
0.5991
accepted
0.5481
0.69959
0.6553
0.6307
accepted
-0.80652
0.34953
-1.36417
0.4027
accepted
0.99997
0.00006
132.9854
0.00479
rejected
0.52919
0.71996
0.62367
0.64499
accepted
0.91024
0.17146
2.19825
0.27179
accepted
0.69558
0.51617
0.96817
0.5103
accepted
-0.19826
0.96069
-0.20228
0.87294
accepted
-0.37782

0.75294
-0.57282
0.66883
accepted
-0.2213
0.95103
-0.22692
0.85794
accepted
0.09561
0.99086
0.09605
0.93904
accepted
-0.20118
0.95953
-0.20538
0.87104
accepted
0.86354
0.2543
1.71243
0.33648
accepted

0.9888
0.10643
0.9325
accepted
-0.75086
0.4362
-1.13688
0.45928
accepted
0.66087
0.56325
0.88058
0.54037
accepted
0.4111
0.831
0.45097
0.73029
accepted
0.99523
0.00951
10.20525
0.06218
accepted

0.64229
-0.74628
0.59185
accepted
0.98269
0.03432
5.30435
0.11863
accepted
-0.95119
0.09524
-3.08225
0.19972
accepted
-0.81848
0.3301
-1.42457
0.38964
accepted
-0.80639
0.34973
-1.36358
0.40283
accepted

0.08998
-3.18022
0.19395
accepted
0.49931
0.75069
0.57629
0.66717
accepted
-0.60533
0.63357
-0.7605
0.58608
accepted
-0.81255
0.33977
-1.39398
0.39616
accepted
0.29199
0.91474
0.30529
0.81137
accepted

0.91753
0.29981
0.81456
accepted
-0.85967
0.26097
-1.68283
0.34134
accepted
0.78781
0.37936
1.27907
0.42243
accepted
0.57198
0.67284
0.69731
0.61235
accepted
0.96024
0.07794
3.43957
0.18012
accepted

0.85725
-0.40807
0.75334
accepted
-0.34843
0.8786
-0.37172
0.77343
accepted
0.22653
0.94869
0.23257
0.85453
accepted
-0.06975
0.99513
-0.06992
0.95556
accepted
0.92269
0.14864
2.39323
0.25197
accepted

1.

293

1.

1.


It can be seen from the above table that compliance with the Paris principles may have some possible correlation with a number of variables: cost-effectiveness efficiency ($r=0.999$, $p=0.0048$), health systems outputs ($r=0.910$, $p=0.272$), and cost efficiency of results ($r=0.922$, $p=0.252$) showed positive correlations with $r > 0.70$; efficiency ($r=0.589$, $p=0.599$), effectiveness ($r=0.548$, $p=0.630$), achievement of results ($r=0.529$, $p=0.645$), and health systems outputs ($r=0.696$, $p=0.510$) showed correlations ranging from $r=0.50$ to $0.70$; and sustainability showed a negative correlation ($r=-0.863$, $p=0.403$) for compliance with the Paris principles.

However, it appears that the $p$ values for these likely correlations are very high in most cases suggesting that the size of the sample is far too small in these cases to run the correlation analysis reliably. Based on a t-test with a 2 degrees of freedom and a critical value of 31.82 (Alpha value for confidence interval= 2%), the null hypothesis of correlations for all but one pair were accepted. This means except the correlation for one pair (compliance with the Paris principles and cost effectiveness efficiency), correlations in other pairs could happen by chance. The Pearson correlation coefficient, in this case, found a plausible correlation ($r=0.99997$, standard error= 0.00006 and $p=0.0048$) between the Paris principles and cost effectiveness efficiency in terms of technical efficiency of cost per DALY saved by an intervention. However, the sample size ($n=3$) was too small to establish its significance.

This pilot investigation indicates that significant correlations may exist between the compliance with the Paris principles and technical efficiency of a project, but this cannot be concluded because of the small sample size. Another investigation using a larger sample may provide a feasible approach to find such correlations.

In fact, the cost effectiveness efficiency (i.e., technical efficiency of cost per DALY saved) score combined the absolute value of outcomes by a single common measure such as DALYs saved by a project with the actual amount of money utilized to calculate the cost per DALY saved. This cost per DALY saved was then compared with the likely cost per DALY saved in the optimum scenario if the project had achieved its full targets by utilizing 100% of its allocated resources. Therefore, the cost effectiveness efficiency score of a project represents both the overall effectiveness of a project and the technical
efficiency of its performance against its targets. In this regard, the seeming correlation between the Paris principles and the cost effectiveness efficiency provides a good indication that it may be possible to find a linear causal relationship between the compliance with the Paris principles of aid effectiveness and development outcomes even at the project level.

9.6 Discussion

The measurement of efficiencies in this chapter involved more than one output and different forms of measurement. Therefore, the DEA technique was useful to identify the apparent best performing projects by their ability to produce the highest level of outputs or outcomes for a given set of inputs in comparison to the performance of other projects. However, the results of DEA exercise depend on the set of projects and variables considered in the DEA model. Compared to the DEA results for economy of procurement of common items and performance efficiency, where the NAP came out as 100% efficient compared to the other two projects, the results for relative effectiveness of the NAP may appear surprising. This was probably due to the consideration of health systems related investments by the projects. In fact, 70.68% of the total expenditure of HSSP-SP was invested on health system strengthening giving the project a higher score on health system outcome points, whereas the NAP invested 45.82% and IPL invested 42.47% of their expenditure for health system strengthening activities.

Although IPL had lower investment on health system strengthening than that of the NAP, due to the nature of children’s immunization programmes, DALYs saved by IPL was much higher than DALYs saved by the NAP. Moreover, IPL used much lower levels of inputs (i.e. total expenditure) than that of the NAP. Therefore, both HSSP-SP and IPL surpassed the performance of the NAP for the effectiveness in terms of DALYs saved and health system outcome points.

It can also be noted that the DEA results for effectiveness of the NAP differ from the stakeholder views, which perceived the NAP as most effective among the three projects. However, the stakeholders rated the projects in terms of their achievement of objectives. The primary purpose of the NAP was to reduce HIV/AIDS related mortality and morbidity in Timor-Leste by expanding HIV/AIDS prevention, treatment and care
services. The health system strengthening only came as a secondary purpose of the NAP as a strategy to achieve the primary goal. Therefore, stakeholders' views of achieving purpose by the NAP were quite consistent with the assessments of project results. Nevertheless, the relative effectiveness score obtained from the DEA model draws attention to the room for improvement for the NAP especially related to the health systems strengthening performance.

There are examples of using DEA techniques to support project evaluations. For example, Nunamaker (1985) analyzed the potential of the use of DEA technique in measuring the performance efficiency of non-profit organisations. Huang and McLaughlin (1989) applied DEA techniques to the evaluation of rural primary healthcare programmes and argued,

“DEA can provide the basis for carrying out detailed field studies to find out what the efficient program[me]s are doing differently than the others. It also can point toward issues of organizational change and its implementation. DEA was suitable for a portion of the evaluation even for program[me]s as heterogeneous as rural primary health care clinics” (p.155).

More recently, Satyanarayana et al. (2012) applied DEA technique for evaluation of Indian rural health centre programmes. Zhang and Pan (2013) used DEA techniques for evaluating the efficiency of basic public health service project in Beijing rural areas.

All these studies differentiated the inefficient projects from the frontier performance of the efficient projects based on the DEA techniques and identified room for improvements for the inefficient projects. Although the DEA technique used in this chapter compared the performance of three different projects with different objectives, the technique was useful in running a primary assessment of efficiencies and triangulating findings from other techniques.

However, it is important to keep in mind several limitations of the DEA technique in interpreting the results. First, DEA assumes a constant return to scale. This means that the model assumed that increase or decrease of a single unit of input would result in proportionate increase or decrease in the output levels. However, in reality this assumption might not be true for all projects. Especially in a fragile situation such as Timor-Leste, there could be necessary programme set up costs such as reconstructing the infrastructure, supplying necessary equipment and recruiting and training human
resources before the programme could be fully functional in producing expected outputs. Therefore, comparing the total costs against total outputs might not be favourable to all projects.

Second, the optimum production frontier in DEA is defined by the best performing units among the total number of units compared in the model. This means that the efficiency score of a project obtained from a DEA model is only relative to the comparing units that fall on the optimum production frontier in that model. DEA assigns 100% relative efficiency to those units on the production frontier. Since this chapter only compared three DMUs (i.e. three projects), it did not yield detailed information about the performance of the units that received a 100% efficiency score from the DEA model. In cases of two or more units receiving a 100% efficiency score, it is not possible to ascertain which one among them was better than the other(s).

Third, while the DEA model in this chapter used real data from the project documents, the scores on health system related outputs and outcomes were estimated from the stakeholders’ interviews. These scores, in fact, represent stakeholders’ impressions about the outputs and outcomes of the projects, rather than a direct measurement.

Fourth, in absence of available data in some cases of cost categories, the DEA exercise had to use best estimations based on available evidence and triangulation. These estimations were related to the health system investments by IPL and the NAP in 2013.

Fifth, results from a DEA model depend on the selection of inputs and outputs for consideration. In order to avoid all DMUs being highly specialized with different procurement items, the calculation of economy in the DEA model only considered the procurement of selected goods and equipment, which were common to all three projects. This indicates a limitation of the DEA model used in this chapter, as inclusion of other items did not return useful results.

Sixth, DEA only considered limited criteria of decision-making such as inputs and outputs. However, decision-makers often need to consider additional criteria such as access, equity, social values, longer term investments related to decision-making on
health investments and health policies, and so on. These additional criteria may not be
directly related to an output, and therefore, were not considered in the DEA model. But
they could still form important additional basis for decision-making. As the Steering
out:

"DEA can be a very useful analytical technique by providing an important 'first step' tool in
comparative analysis. But users also need to recognise its limitations as an input to the
development of public policy. Its theoretical predictions of potential efficiency gains may not be
translatable into actual gains when factors such as service quality, fundamental differences
between individual services and the costs of implementing changes are fully accounted for. Non-
efficiency objectives such as access and equity are also important policy considerations for
government, against which efficiency benefits will inevitably be balanced" (p. ix).

Similarly, the DEA model did not consider any value judgment related to the DALYs
saved and achieving health system outcomes in measuring relative effectiveness. It
depends on the policy and decision makers whether they would value health system
strengthening over DALYs saved or otherwise and to what extent. Inserting such value
judgment constraints in the DEA model would have changed the results.

Apart from the above-mentioned limitations, statistical tests were also not possible with
the DEA results due to their non-parametric nature of data. Moreover, due to the relative
values of the DEA results, they were not considered in the correlations analysis.

With the caveat of a small sample size of only three projects, the attempt to evaluate
correlations between the compliance with the Paris principles and other variables
provided some indication of plausible association between the compliance with the aid
effectiveness principles and programme effectiveness. With increased sample size, this
may provide a feasible approach to test if there is a causal relationship between the aid
effectiveness principles and development effectiveness as attempted in a number of
studies discussed below.

Paul, et al. (2013) proposed that results from the aid effectiveness principles need to be
evaluated in three causally linked levels as shown in the following diagram:
Proposing the process level, health systems level, and health outcome level evaluations, Paul et al. (2013) argued,

“although it is impossible to prove that observed changes [in health outcomes] are directly attributable to aid management reforms, qualitative methods can help understand the role of particular processes, and how far positive changes can be attributed to improvements in aid delivery systems and HSS [health systems strengthening]” (p.145).

Although Paul et al. (2013) could not establish a linear causal link between the implementation of the Paris principles and health outcomes, the Working Party on Aid Effectiveness of OECD that undertook an evaluation of the Paris principles in the health sector found some indirect but clear pathway of better results from better implementation of the Paris principles of aid effectiveness in the health sector. Their report, however, acknowledged the challenges of measuring and attributing the impact of aid effectiveness principles on health outcomes:

“Efforts have focused more on aid effectiveness processes than on the impact of better aid on health service delivery and outcomes. The transaction costs of aid effectiveness processes are high, in particular for donor country staff, and there is a risk that the costs are disproportionate to the benefits. In addition, demonstrating and attributing the impact of aid effectiveness is challenging and there is no common understanding of what results can realistically be expected or how these will be measured” (OECD Working Party on Aid Effectiveness Task Team on Health as a Tracer Sector, 2011, p. 10).

An analysis by Dickinson (2011) also suggested that implementation of aid effectiveness principles is contributing to development through creating conditions for sustainable impact:

“There is evidence that aid effectiveness is improving sector planning, budgeting, strengthening national systems and increasing resource allocations. More efficient funding of the health sector, through programme based approaches including SWAPs is helping implement health sector reforms, which are contributing to better health results” (Dickinson, 2011, p. 9).
Considering the indications from these studies, underlying theory of change, and a plausible causal chain behind the Paris principles and health intervention outcomes, the indication of possible correlation between the Paris principles and technical efficiency of health interventions in achieving cost effectiveness in the current study shows merit in pursuing such an approach with increased sample size to be able to establish any significance of such finding.

9.7 Conclusion

By using the DEA technique, project data, and findings from other chapters, this chapter attempted further analysis of comparative economy of procurement of common items, performance efficiency, and effectiveness of the three projects. It then attempted to find any possible correlations between the compliance with the Paris principles of aid effectiveness and programmatic performance and effectiveness.

The findings from the DEA exercise confirmed and validated the findings from the stakeholder analysis and cost effectiveness. Moreover, the indication of correlation between the compliance with the Paris principles and programmatic effectiveness backed up by the programme theory indicates a plausible causal relationship between the implementation of the Paris principles and positive outcomes of aid funded interventions.

However, due to the very small sample size and limitations of the methods that generated the data, further research will be required to investigate and explain if there is any causal link between the Paris principles of aid effectiveness and the programmatic outcomes in the health sector. By addressing the limitation of sample size and measurement techniques, this approach might be modified as a tool to gather empirical evidence to support the programme logic behind the aid effectiveness principles.
10 Lessons learned

10.1 Introduction

This research pilot tested a theory-driven realist evaluation approach with a range of analytical tools to compare and critically evaluate three different aid management modalities in Timor-Leste’s health sector. The country context, state fragility, and health systems analyses provided the broad context in which the aid projects under evaluation were implemented; analysis of aid projects provided useful information to understand how the aid mechanisms interacted with the context to produce the project and health system outcomes.

In order to evaluate the aid projects’ performance, interactions with the context, and outcomes, this research pilot tested a range of analytical tools such as balanced scorecards, cost effectiveness analysis, data envelopment analysis, and correlation analysis. This chapter attempts to capture lessons learned from these pilot studies to assess their feasibility to evaluate large-scale national projects in the health sector. Especially, this chapter attempts to seek answer to the following questions:

- How the analytical techniques were implemented in this and other evaluation research;
- Data requirements, strengths, and weaknesses of the analytical techniques; and
- Limitations of the pilot and how the limitations can be overcome to produce meaningful results of the evaluation.

10.2 Balanced scorecards

As explained in Chapter 7, this research used a balanced scorecards approach to assess the perceptions of stakeholders about the performance of each project under evaluation. Originally, the balanced scorecard approach was used as a performance measurement approach focusing on the corporate strategy from four perspectives: financial, customer, internal process, and learning and growth (Kaplan and Norton 1992, 1997, 2001). As a strategic planning and management tool, a balanced scorecard approach seeks to align business activities to the vision and strategy of the organization, improve internal and external communications, and monitor an organization’s performance against the
strategic goals. This is done by identifying key performance indicators (KPI) in all four perspectives (or ‘domains’) and then measuring the status of the organization against those KPIs and their benchmarks. Results are then presented side by side for each domain to see the status of the organization from those perspectives (https://balancedscorecard.org/Resources/About-the-Balanced-Scorecard accessed 31 January 2016).

The balanced scorecard approach has been modified and used for performance measurement in the healthcare sector in various studies (Inamdar, Kaplan, and Bower, 2001; Chan et al., 2010; Chen et al., 2006; Commonwealth Fund Commission on a High Performance Health System, 2008; Gauld et al., 2011; Hansen et al., 2008; Peters et al., 2007; Pink et al., 2001; Schoen et al., 2006; Ten Asbroek et al., 2004; Voelker, Rakich, & French, 2001; Zelman, Pink, & Matthias, 2003). These studies tried to align the performance of a healthcare organization with its vision and objectives and then attempted to measure different domains of the organization against their KPIs and benchmarks. However, the balanced scorecards approach has not yet been adapted and tested in comparative evaluation of aid funded projects from the perspectives of the compliance with the Paris principles, economy, efficiency, effectiveness, and sustainability.

Following the principles of balanced scorecards, a 22 item questionnaire (Annex 1) was developed in this research to assess each aid funded project’s compliance with the Paris principles, economy, efficiency, effectiveness, and sustainability. This questionnaire was guided by the OECD (2011) Survey on Monitoring the Paris Declaration for the monitoring indicators of compliance with the Paris principles, and the Grant Performance Assessment Methodology of the Global Fund (http://www.theglobalfund.org/en/performancebasedfunding/decisionmaking/methodology/ accessed 22 September 2012) for the analysis methods, and OECD-DAC (2010) guidelines for development evaluation for selecting the domains of evaluation.

In absence of established benchmarks against each domains of compliance with the Paris principles, economy, efficiency, effectiveness, and sustainability, this questionnaire
attempts to rate the status of a particular indicator against the best practice (or the highest attainable desired status) in a scale of 1 to 5.

While corporate balanced scorecards are often scored by management consultants or experts, this questionnaire was used in this research to collect opinions from stakeholders and their rating on each item of the questionnaire for an aid funded project under evaluation. Scores were then used for further analysis along with triangulation of the rating with other available evidences.

Strengths of such approach were the flexibility of selecting and designing key performance indicators according to the purpose of evaluation, involvement of stakeholders including the implementers and decision makers in self-evaluation of the projects, the ability to analyze and present multiple aspects of three projects at the same time, and the ability to capture and analyze a wide range of qualitative information related to each project. However, several challenges and limitations of the approach were also identified through the pilot. For example, since filling in the questionnaire was based on stakeholders’ opinions, the scores on the questionnaire, in fact, provided the perception of the stakeholders on the performance and status of each project under evaluation. Such perceptions can be influenced by the visibility and communication of each project, rather than their actual performance. It is also difficult to isolate and perceive health systems’ performance and effect of a project on the health system.

Another limitation of this approach was the limited number of stakeholders who were knowledgeable on all three projects. A total of 22 respondents took part in this study. However, not all of them were equally knowledgeable on all the projects. Beyond these 22 respondents, there were no other respondents knowledgeable on the projects available to fill in the questionnaire. Thus the total number of respondents was limited and posed a difficulty for further statistical analysis.

There have been issues around the validity of score analysis. In a sense, the 1 to 5 ranks in the questionnaire represent an ordinal scale, as there is no guarantee that the distances between the ranks or steps are equal. However, for the sake of analysis, the scores were analyzed as 'likert items' with importance weighting, summing up, and
calculating the average and standard deviations. Thus the scores do not represent an absolute performance measurement, rather they represent an indication of average perceptions of the stakeholders about the projects.

The pilot also involved three top level decision-makers in rating the importance of each item of the questionnaire. Their average ratings were used in weighting and converting average scores from the questionnaire survey. As these importance ratings totally depend on management decisions, they do not represent absolute values for empirical analysis. Similarly, converting score on a particular domain to equal scales with other domains’ scores also has an effect on the overall scores. It is, again, a management decision for the decision makers of a health sector to decide which domain they would value over others. For example, some decision makers may value compliance with the Paris principles over any other aspects of a project; some decision makers may value effectiveness of a project over other aspects; whereas some other decision makers may value sustainability of the project outcomes over any other aspects.

As the primary purpose of the questionnaire was to collect and analyze stakeholders’ perspectives, and not to measure the absolute performance through this method, the questionnaire was not further investigated for validity measures such as construct validity, or test-retest validity. Rather, validity of the findings were achieved through triangulation with other available evidence such as project documents, focus group discussion, and other analytical methods used in other chapters including cost-effectiveness analysis, and data envelopment analysis.

The balanced scorecards approach used in this pilot, therefore, provides a perception spectrum on different dimensions, not a comparison to decide which project is better than others. This approach can provide a good starting point to gather and analyze stakeholders’ views, which are useful for further analysis and triangulation. Suggestions for further refining and improving this approach may include:

- Conducting a validity test of the questionnaire by testing-retesting, and construct validity analysis;
- Redesigning the questionnaire in the form of a cardinal (i.e., equal distance between the steps) scale, for example, presenting the questionnaire items in the
form of statements with a ‘strongly agree to strongly disagree likert scale of 1 to 5’ (i.e., 1 = strongly disagree, 2 = disagree, 3 = neither agree nor disagree, 4 = agree, 5 = strongly agree);

- Independently filling in the questionnaire by two experts based on available evidence and then comparing their scores with the scores obtained from the stakeholders’ responses for further validity of scores;
- Using an expert panel or research reference group to validate the findings.

So, conditions for the balanced scorecard approach to be useful for a comparative evaluation are: having a clear set of domains and indicators to measure the status of a project; having a tested, validated, and statistically analyzable questionnaire to measure the indicators; and having a group of reliable scorers who can score the indicators.

### 10.3 Cost effectiveness analysis

Apart from the balanced scorecards filled in through the stakeholders’ opinion survey, this research employed cost-effectiveness analysis to measure and compare cost-effectiveness of the three projects and their technical efficiency against a theoretical optimum project performance. As explained in Chapter 8, cost-effectiveness analysis for health projects attempts to calculate and compare cost per a health outcome for different projects or strategies. In order to avoid the incomparability of different health outcomes, cost-effectiveness analysis often uses a common measurement unit such as DALY (disability adjusted life years) to convert the health outcomes into measurable values incorporating values for deaths and disease related disabilities (McCabe, 2009).

Cost effectiveness analysis is used by a number of writers for comparative effectiveness analysis of different health programmes or strategies for informed resource allocation decisions (Drummond et al., 2005; Gold et al., 1996; Jamison et al., 2006a, 2006b; Jamison et al., 1993; Levin & McEwan, 2001; Miguel & Kremer, 2004). These writers calculated cost per DALY saved or cost per QALY (quality adjusted life years) gained by different health programmes or strategies to compare their value returns against the resource allocations to help making more informed decisions.
This research also employed a cost-effectiveness technique to pilot a comparative effectiveness evaluation of three aid-funded projects. The strengths of such cost-effectiveness analysis approach are: the use of common measurement for different health outcomes, generating more comparable data about project performance and outcomes, and providing additional useful information for decision making.

However, the pilot also faced several challenges and limitations. One such challenge was related to the calculation of DALYs saved by a health sector-wide programme such as HSSP-SP. Since HSSP-SP did not target any specific disease but contributed to overall strengthening of health sector’s performance, the pilot attempted to calculate DALYs saved by HSSP-SP from the overall disease burden of the country derived from WHO Global Health Estimates (WHO, 2014). Here, the pilot relied on the economic analysis of the World Bank conducted as part of the project appraisal (World Bank, 2007) for the assumptions related to the counterfactual scenario (i.e., what would be the country’s disease burden without HSSP-SP), and pessimistic and optimistic targets for HSSP-SP as envisioned by the World Bank. Due to possible contribution of multiple interventions in the health sector, without availability of such assumptions, it would be difficult to draw a counterfactual scenario and calculate DALYs saved by HSSP-SP.

Calculating DALYs saved by the NAP also relied on statistical modeling through Spectrum software (Futures Institute, 2014) as well as on a number of assumptions as described in Chapter 8. Therefore, such calculations represent estimated values that cannot be taken as absolute values. Similarly, for IPL, although DALYs saved were calculated based on the actual number of children vaccinated in IPL focus districts, for a counterfactual scenario, disease prevalence and possible infected cases were calculated based on pre-vaccination era prevalence rates. These calculations were then compared with the vaccination coverage trends in non-IPL focus districts to calculate DALYs saved attributable to IPL. Again, these calculations represent estimates, not absolute values.

As described in Chapter 8, the cost-effectiveness analysis employed in this research also faced some methodological issues due to adopting Global Health Estimates (GHE) methods followed by WHO (2013). The GHE method differs from the traditional method of DALY calculation that often calculate DALY with local life table, age weighting, and
time discounting. As WHO (2013) summarized, along with using a revised disability table and updated information on death and diseases, in the GHE method:

- A simpler form of DALY, used by the GBD 2010 study (Murray et al., 2012b), has been adopted. This form is easier to explain and use. Age-weighting and time discounting are dropped, and the YLDs are calculated from prevalence estimates rather than incidence estimates. YLDs are also adjusted for independent comorbidity.
- The standard life table used for calculation of years of life lost for a death at a given age is based on the projected frontier life expectancy for 2050, with a life expectancy at birth of 92 years...(p.2)"

As age-weighting and time discounting would present the social values of DALYs saved, calculation of DALYs saved by following the GHE method, therefore, presented only the death and disease related estimates, not their social or economic values. However, the pilot study used a gross conversion rate to convert the baseline year’s GBD (i.e., global burden of disease calculated in traditional method) estimates to GHE estimates.

Since the three projects under evaluation were very different from each other, served different populations, and had very different outcomes, DALYs saved by each project were not readily comparable to each other. The cost-effectiveness analysis was, therefore, complemented by a technical efficiency analysis for each project by comparing DALYs saved by each project with the likely DALYs saved by each project in a theoretical optimum performance scenario (i.e., if the project had optimally achieved all its targets as per the project design). This technical efficiency of cost effectiveness gave an indication of a project’s achievements against its own standards, which appeared to be reasonably comparable to each other, as can be seen in Table 10.1:

**Table 10.1: Comparison of cost effectiveness and technical efficiency of three projects**

<table>
<thead>
<tr>
<th></th>
<th>HSSP-SP</th>
<th>NAP</th>
<th>IPL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intervention duration (years)</td>
<td>4.5</td>
<td>6.5</td>
<td>2.5</td>
</tr>
<tr>
<td>Total budget</td>
<td>$21,000,000</td>
<td>$14,446,916</td>
<td>$2,639,250</td>
</tr>
<tr>
<td>Total expenditure</td>
<td>$16,240,000</td>
<td>$11,842,703</td>
<td>$2,639,250</td>
</tr>
<tr>
<td>Utilization of budget</td>
<td>77.33%</td>
<td>81.97%</td>
<td>100%</td>
</tr>
<tr>
<td>Total DALYs saved (current scenario)</td>
<td>44,628</td>
<td>35,325.19</td>
<td>246,232.50</td>
</tr>
<tr>
<td>Cost per DALY saved (current scenario)</td>
<td>$363.90</td>
<td>$335.25</td>
<td>$10.72</td>
</tr>
<tr>
<td>Total DALYs saved (target scenario)</td>
<td>107,616</td>
<td>63,591</td>
<td>464,208</td>
</tr>
<tr>
<td>Cost per DALY saved (target scenario)</td>
<td>$195.14</td>
<td>$227.19</td>
<td>$5.69</td>
</tr>
<tr>
<td>Technical efficiency (current cost per DALY saved / target cost per DALY saved)</td>
<td>53.62% (against pessimistic targets)</td>
<td>67.77%</td>
<td>53.04%</td>
</tr>
</tbody>
</table>

Source: Author’s calculations (Chapter 8)
As can be seen in the above table, total DALYs saved or cost per DALY saved by HSSP-SP, NAP, and IPL are not at all comparable to each other. But when the cost per DALY saved are compared against the likely cost per DALY saved in the optimum performance (target) scenario, their technical efficiencies are fairly comparable to each other. These technical efficiencies, in fact, indicate to what extent a project was successful in implementing its design and achieving the desired outcomes.

In order to further improve the cost effectiveness and technical efficiency analysis approach, the following suggestions can be made:

- Validating all the assumptions used in the analyses by collecting credible evidence or by conducting uncertainty analysis using the full range of assumptions;
- Conducting another version of cost-effectiveness and technical efficiency analysis with social value considerations and assess the differences in findings from the current approach;
- Developing and using statistical models to calculate DALYs saved by a sector wide programme such as HSSP-SP and to isolate contributions by different interventions in the health sector;
- Identifying and addressing confounding factors in the measurement of achievements of different projects.

So, conditions for the cost effectiveness analysis technique to be useful for comparative evaluation are: availability of cost and project outcome information; availability of statistical modeling software and techniques for different scenario and uncertainty analysis; and the ability to identify and isolate the confounding factors from the analysis.

10.4 Data Envelopment Analysis (DEA)

As explained in Chapter 9, DEA is a linear programming technique that can be used for evaluating the relative efficiency of a group of production units with similar outputs in the context of multiple inputs and multiple outputs. DEA evaluates the relative efficiency of a production unit by comparing it against a virtual efficiency frontier created by the best performing units (Farrel, 1957; Charnes et al., 1978; Banker et al., 1984).
As described in Chapter 9, use of DEA techniques to evaluate efficiency in the healthcare sector is not new. For example, Huang and McLaughlin (1989) applied DEA techniques to the evaluation of rural primary healthcare programmes. Satyanarayana et al. (2012) applied a DEA technique for evaluation of Indian rural health centre programmes. Zhang and Pan (2013) used DEA techniques for evaluating the efficiency of basic public health service project in Beijing rural areas. By considering inputs such as the labor force, materials, and capital used by a hospital to produce outputs such as the quantity of emergency, inpatient, and outpatient services provided by the hospital, these studies differentiated the inefficient projects from the frontier performance of the efficient projects based on the DEA.

The DEA technique piloted in this research attempted to compare the performance of three different projects in terms of their economy of procurement of common items, efficiency in achieving project targets and health system outputs, and effectiveness of achieving outcome values (in DALYs) and health system outcomes. This pilot was conducted to further test and triangulate the findings from other analyses. The strengths of this technique were the ability to use non-parametric data and the ability to analyze multiple inputs and outputs at the same time. However, the pilot also revealed several challenges and limitations of the technique.

The DEA technique works best when there are a large number of units to study with a common output or outcome. However, in the current pilot, the total number of units to study was only three and they had different objectives. In DEA at least two of these units would form the comparatively best performing efficiency frontier. Therefore, DEA in this case, only revealed the relative inefficiency of one project in comparison with other two projects on the efficiency frontier.

Requirement for a common output or outcome also created some challenges for DEA. For example, running DEA for relative economy of three projects, in fact, only tested the procurement costs of the common items procured by three projects and did not represent the overall economy of a project. Running DEA for relative efficiency also faced the challenge of different outputs of different projects. However, these outputs
were converted into percentage of achieving project outputs against their performance targets and baseline set by each project and thus they became comparable to each other.

Similarly, for running DEA for relative effectiveness, the common health outcomes calculated in Chapter 8 in terms of DALYs saved were used. However, due to different programme objectives and different target populations, comparing DALYs saved by the projects, in this case, only provides a partial view, and does not really represent the effectiveness of an intervention.

Another challenge with the use of DEA was related to measuring the health system strengthening related outputs and outcomes by three projects. Since all three projects were implemented in the state fragility context, it was expected that they would strengthen the country’s health system and help the country to recover from fragility. However, there was no available method to isolate and measure the contribution of each project in health system strengthening outputs and outcomes. The DEA piloted in this research used the stakeholders’ perception obtained in Chapter 7 for the health systems outputs and outcomes index of each project to convert the health system strengthening related investments (expenditure) by each project to some indicative health system strengthening output and outcome values. However, the validity of such calculation was not investigated.

It appears that, with the above limitations, the DEA technique by itself cannot provide a complete evaluation approach for a small sample size with different programme objectives. However, along with other qualitative and analytical techniques, the DEA was still useful to triangulate and test the findings from other analyses. Running DEA with three project samples could still identify the sub-optimally performing project and potential areas of its improvement. For more useful results, use of the DEA technique could be further improved in the following way:

- Increasing the size of the sample (along with the three projects under study, at least two other benchmark projects can be included in the DEA analysis to derive more useful results);
- Developing and using tools for measuring health systems outputs and outcomes contribution of each project;
• Developing more sophisticated methods (e.g., a cost-benefit analysis) for the analysis of relative economy of a project;
• Including social values of project outcomes in relative effectiveness analysis of a project.

So, conditions for the DEA technique to be useful in comparative evaluation are: having a large sample of projects with similar outputs and outcomes for analysis; having reliable measurement of the outputs and inputs used in the DEA model (for example, measurement for health system related outputs and outcomes); and availability of required data for such measurements.

### 10.5 Correlation analysis

Along with a range of comparative and critical analyses, this research also attempted to explore if there are any significant correlations between the variables on compliance with the Paris principles and economy, efficiency, effectiveness, and sustainability of the three projects under study. As seen in Chapter 9, this attempt of investigating correlations was seriously constrained by the small sample size (n=3).

As seen in Chapter 9, most of the pairs of variables under investigation returned insignificant results except for one pair: overall compliance with the Paris principles and technical efficiency of cost per DALY saved. This pair returned a Pearson correlation coefficient of 0.99997, standard error of 0.00006, and p-value of 0.0048. Values used for this pair of variables for the investigation of correlations can be closely looked at in Table 10.2 below:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Total score</th>
<th>HSSP-SP</th>
<th>NAP</th>
<th>IPL</th>
<th>Data source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compliance with the Paris Principles</td>
<td>325</td>
<td>203.77</td>
<td>240.67</td>
<td>202.59</td>
<td>Stakeholder interviews (Chapter 7)</td>
</tr>
<tr>
<td>Technical efficiency of cost per DALY saved</td>
<td>100</td>
<td>53.62</td>
<td>67.77</td>
<td>53.04</td>
<td>Cost effectiveness analysis (Chapter 8)</td>
</tr>
</tbody>
</table>

As can be seen in the above table, compliance with the Paris principles were calculated from the stakeholders’ perception collected and analyzed in Chapter 7. These scores
indicate to what extent each project under study adhered to the Paris principles as assessed by the stakeholders. Technical efficiency of cost per DALY saved, on the other hand, were calculated from the cost-effectiveness analysis in Chapter 8. These technical efficiency scores were derived by comparing the cost per DALY saved by each project with their likely cost per DALY saved in optimum performance scenario (i.e., if the project achieved all its targets by utilizing 100% of its budget). Therefore, the technical efficiency scores for cost per DALY saved indicates a project’s performance of achieving health outcomes against its own standards.

Clearly, the variable of compliance with the Paris principles, and technical efficiency of cost per DALY saved indicate two different things and they are measured in two different ways. However, the almost perfect correlation found between them suggests that there could be some possible measurement issues, as the perfect correlation otherwise may indicate that they are measuring the same thing.

It is obvious that the sample size (n=3) in this pilot correlation analysis was far too small to establish any significance. But since the above pair of variables definitely included two different things, the approach shows some merit and indicates that with increased sample size it may be possible to run a correlation analysis to investigate a causal relation between the compliance with the Paris principles and performance efficiency of an aid funded project. The following recommendations can be made to further improve the correlation analysis:

• Clarifying the variables and unit of observations for correlation analysis;
• Increasing sample size for the observations;
• Conducting additional statistical analysis such as confidence intervals to infer significance of correlations.

So, conditions for a correlation analysis to be useful in a comparative evaluation are: identifying specific variables for correlation analysis and their units of observation; having a feasible sample size of projects for the observations; and availability of other tests to interpret the significance and strengths of associations.
10.6 Conclusion

This research employed a series of pilot studies using multiple techniques and analytical tools to analyze broader context and project instruments, elicit underlying programme theories of the projects under study, and collect supporting evidence from multiple sources and approaches about the performance of each project. It appeared that, in an isolated way, none of the analytical tools could provide a complete evaluation approach by themselves. However, findings from this series of analysis using different tools and approaches were compared with underlying programme logic of each project. It was also investigated to what extent the findings from different approaches agreed to or differed from each other with possible explanations for the differences. Together, this high-level analysis, synthesis, and triangulation, in fact, provided a feasible approach to a comparative evaluation of management modalities of aid funded projects in their interplay with the compliance with the Paris principles and interactions with the health systems in a state fragility context.

The pilots, in fact, represent the real life situation and challenges in evaluating large-scale national programmes, especially in a fragile state situation. Such challenges include: lack of a control group making an experimental research design not feasible, limited availability of required information and data, and absence of necessary tools and methods for assessing system-wide effects involving multiple intervention. With these challenges, the evaluation approach piloted in this research attempted to build best-informed and best-guessed scenarios to enable managers and policy makers make informed decisions.

As the projects under study were very different from each other, a mere finding that a project is outperforming another and the outperforming project is more compliant with the Paris principles does not provide enough evidence to infer a causal pathway. The pilots, therefore, attempted more sophisticated techniques and analysis from different angles to derive more comparable measurement values. The pilot showed that the evaluation framework, tools, and techniques employed in this research can provide a feasible approach for comparative evaluation of aid funded project. The approach can be further developed and tested with increased sample size and by following other
recommendations made in this chapter. With these lessons learnt, Chapter 11 provides a summative discussion and conclusion of this thesis.
11 Summative discussion and conclusion

11.1 Introduction

In order to improve the effectiveness of foreign aid for the recipient country, the Paris Declaration of Aid Effectiveness (2005) provided a set of process and coordination principles. As Manning and Trzeciak-Duval (2010) identified, “It is the underlying assumption of the Paris Declaration that improved aid effectiveness [i.e. adherence to the set of aid effectiveness principles] will in turn improve development outcomes as measured by progress towards the MDGs” (p.125).

However, the number of studies on the effect of the Paris principles on the development outcomes in the health sector is very limited. Grap-Pa Sante (2011) pointed out that there is a notable dearth of rigorous analysis and country studies that look at the impact on health results in relation to the use of the Paris principles.

Assessing effects of the Paris principles on the health intervention’s outcomes in fragile states and situations can be even more challenging than in other situations. This is because of the confounding factor of the state fragility affecting the performance of health systems and any intervention within the health systems. The health systems in a post-conflict fragile state often suffer from conflict, destruction of clinic and hospital infrastructure, the flight of health professionals, and the interruption of drugs and other medical supplies (Kruk et al., 2010). Although it is believed that the Paris principles of aid effectiveness are still applicable for achieving better development results in fragile situations (Leader & Colenso, 2005; OPM/IDL, 2008), to this author’s knowledge, no studies have demonstrated so far a linear causal relationship between the implementation of the Paris principles and development results achieved by the aid funded health programmes in fragile situations.

Against this backdrop, this thesis wanted pilot test an evaluation approach to see which management modality of aid funded health intervention is more compliant with the
Paris principles of aid effectiveness in a fragile state’s health sector and if this compliance has any effect on the achievement of the programmatic objectives and health systems strengthening by the aid funded intervention. For this purpose, this research selected three externally funded projects from Timor-Leste’s health sector: the HSSP-SP funded by AusAID and World Bank through a MDTF mechanism; the NAP that channelled the funding through the government but not fully using the government systems; and the IPL or ‘Immunizasaun Proteje Labarik’ funded by USAID with funding channelled through an NGO outside the government.

As discussed in Chapter 5, following a theory driven realist evaluation approach, this research used mixed methods with different qualitative and quantitative techniques. The qualitative techniques included analyzing the country context and fragility; analyzing health systems and project backgrounds; and collecting stakeholders’ views on adherence to the Paris principles, economy, efficiency, effectiveness, and likely sustainability of each project. The quantitative techniques included analyzing stakeholders’ views in the form of balanced scorecards; conducting cost-effectiveness analysis of three projects; using a non-parametric DEA method for comparing relative economy, efficiency and effectiveness of each project; and finally running statistical methods for finding possible correlations between the Paris principles and other variables.

By providing an overall summary and discussion, this chapter attempts to synthesize the findings from different methods, evaluates the validity and robustness of the research methods followed in this thesis, and draws conclusions with a set of recommendations for each aid funded project.
11.2 Summary of research findings

11.2.1 Contextual analysis

In an attempt to understand the project designs and their possible implications for project results this research conducted contextual analysis at four levels: the country context, state fragility context, health systems context, and project backgrounds context.

The contextual analysis of Timor-Leste's geography, people and culture, history, economy, politics, and human development indicated some strengths, challenges and risk factors for the health sector. It was noted that Timor-Leste's health and human development outcomes are still very low and many of the MDGs are unlikely to be met by 2015. Maternal mortality is very high. The country ranks number one in the world in terms of the percentage of under-weight children under five years of age. More than half of the country’s children under five are stunted indicating chronic mal-nutrition. 31% of the population does not have access to improved drinking water and 61% does not have access to improved sanitation facility.

The state fragility assessment of Timor-Leste showed that while the country made good progress in the areas of gaining state legitimacy, improving public services, and improving security, there are still weaknesses in the justice sector and the country may face increasing threats from the demographic pressure and uneven development. Some remaining factors of fragility such as poverty and lack of economic opportunities especially in the districts and rural areas, weak institutional services, horizontal inequalities between the people from the east and people from the west of the country, non-reconciled past, global economic shock and climate change may push Timor-Leste again into the cycles of conflicts and fragility.

An analysis of the health systems context of Timor-Leste also showed that the country saw increased life expectancy, reduced child mortality and reduced incidence of infectious diseases from those of 2002 and 2008. Vaccination coverage has also been increased.
However, a significant section of the country’s population living in the remote rural areas are still beyond the reach and coverage of most of the health services. Quality assurance of pharmaceutical procurement, shortage of medical supplies, and stock out of essential drugs have been frequent issues. The country still needs to develop an efficient and functioning health management information system at various levels. Finding longer-term sources of health financing without being dependent on the Petroleum Fund and donor support is another challenge for the country.

11.2.2 Project backgrounds and instruments

Timor-Leste’s health sector saw a number of different aid instruments and aid management modalities. The HSSP-SP used a MDTF mechanism managed by the World Bank with an aim to facilitate a SWAp in the health sector by channelling the funds through the government mechanism. The USAID funded IPL, on the other hand, used a contracting mechanism by channeling the funds outside the government mechanism to implement the project. The NAP was a hybrid approach in the continuum of state-building and state avoidance.

11.2.3 Stakeholders’ views, cost effectiveness analysis and DEA

This research employed three different methods-- balanced scorecards with stakeholder interviews, cost-effectiveness analysis, and DEA for collecting information, analyzing data, validating findings, and interpreting results.

Stakeholders’ views:

The following chart presents and compares the key findings from the stakeholder interviews:
As the above chart shows, apart from the sustainability domain, HSSP-SP scored lower than the NAP and IPL in terms of its compliance with the Paris principles, economy, efficiency, and effectiveness. While the NAP scored higher on its compliance with the Paris principles than other two projects, NAP’s score on sustainability was lower than the other two.

The balanced scorecards showed USAID-MCC funded IPL as comparatively more balanced in its scores in all five domains than the HSSP-SP and the NAP. IPL scored moderately on its compliance with the Paris principles but outperformed the other projects on economy and sustainability aspects. IPL’s scores on efficiency and effectiveness were also quite close to those of the NAP.

Cost effectiveness analysis:
Table 11.1 presents the summary of cost effectiveness analysis and technical efficiency of three projects in terms of cost per DALY saved:
### Table 11.1: Summary of cost effectiveness analysis

<table>
<thead>
<tr>
<th></th>
<th>HSSP-SP</th>
<th>NAP</th>
<th>IPL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Programme duration (years)</td>
<td>4.5</td>
<td>6.5</td>
<td>2.5</td>
</tr>
<tr>
<td>Total budget</td>
<td>$21,000,000</td>
<td>$14,446,916</td>
<td>$2,639,250</td>
</tr>
<tr>
<td>Total expenditure</td>
<td>$16,240,000</td>
<td>$11,842,703</td>
<td>$2,639,250</td>
</tr>
<tr>
<td>Utilization of budget</td>
<td>77.33%</td>
<td>81.97%</td>
<td>100%</td>
</tr>
<tr>
<td>Total DALYs saved</td>
<td>44,628</td>
<td>35,325.19</td>
<td>246,232.50</td>
</tr>
<tr>
<td>Cost per DALY saved</td>
<td>$363.90</td>
<td>$335.25</td>
<td>$10.72</td>
</tr>
<tr>
<td>Total DALYs saved</td>
<td>107,616</td>
<td>63,591</td>
<td>464,208</td>
</tr>
<tr>
<td>Cost per DALY saved</td>
<td>$195.14</td>
<td>$227.19</td>
<td>$5.69</td>
</tr>
<tr>
<td>Technical efficiency</td>
<td>53.62% (against pessimistic targets)</td>
<td>67.77%</td>
<td>53.04%</td>
</tr>
<tr>
<td></td>
<td>29.27% (against optimistic targets)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

It appears from this table that IPL saved maximum DALYs with minimum costs and in minimum time. However, calculating the technical efficiency of each project based on their own benchmarks and targets set by their programme design, shows that overall the NAP had a higher technical efficiency (67.77%) than those of IPL (53.04%) and HSSP-SP against their pessimistic targets (53.62%).

### Data Envelopment Analysis (DEA):

Table 11.2 summarizes the findings from the DEA:

### Table 11.2: Summary of results from DEA

<table>
<thead>
<tr>
<th>Project</th>
<th>Economy of procuring goods &amp; equipment</th>
<th>Economy of procuring human resources &amp; technical assistance</th>
<th>Performance efficiency for project outputs &amp; health system outputs</th>
<th>Effectiveness for project outcomes &amp; health system outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSSP-SP</td>
<td>87.2%</td>
<td>100%</td>
<td>88.89%</td>
<td>100%</td>
</tr>
<tr>
<td>NAP</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>76.28%</td>
</tr>
<tr>
<td>IPL</td>
<td>100%</td>
<td>100%</td>
<td>81.06%</td>
<td>100%</td>
</tr>
</tbody>
</table>

It appears that other than the domain of effectiveness, the findings of DEA on economy and performance efficiency of the three projects were quite consistent with the stakeholders’ views as shown in Figure 11.1. While the NAP scored high on economy, performance efficiency and rate of achieving health system outcomes as per stakeholders’ views, comparing its actual investment on health system strengthening activity had an effect on its relative effectiveness in the DEA model. HSSP-SP with much higher investment on health system strengthening, and IPL with much higher DALYs
saved clearly outperformed the NAP in terms of DALYs saved and health system outcomes achieved.

11.2.4 Correlation

The variables from the results from the stakeholder interviews, cost effectiveness analysis, and project results were considered to find if there were any significant correlations between the variables. The results showed a Pearson correlation coefficient correlation of 0.99997 between the compliance with the Paris principles and technical efficiency of cost per DALYs saved by each project with a standard error of 0.00006 and p-value of 0.0048. However, the sample size (n=3) was far too small to establish any significance of this finding.

11.3 Discussion

The strapline of the cover-page of Timor-Leste’s 2013 State Budget Book for Development Partners says, “Good bye conflict, welcome development” (Ministry of Finance of Timor-Leste, 2013f). The rapid economic growth, low human development outcomes, and challenging achievements against MDGs in Timor-Leste indicated a transition for the country from the state of fragility to a state of reconstruction and development. However, this transition to post-conflict status, as Collier and Hoeffler (2004) warned, may not be linear, as political settlements often take years, and about 40% of fragile states collapse back into conflict.

The findings from the contextual analysis of this research are, in fact, not much different from what Nathan Associates found in their economic recovery assessment conducted for Timor-Leste in 2008. For the health sector, Nathan Associates (2008) highlighted that:

“Despite progress in reestablishing some basic health infrastructure since independence, the disrupted or short supply of drugs and skilled health professionals has resulted in poor health outcomes. Maternal mortality and child malnutrition are high; communicable diseases such as TB and malaria are widespread; the supply of basic facilities such as improved water source and sanitation remain weak; and public education on health is needed” (p.vi).

A number of other writers also confirmed that conflict affected fragile states are often worse off than the non-fragile states in terms of key health indicators and social determinants of health due to the weak and incomplete health system building blocks
and the health system's inability to provide quality health services to a large proportion of population (Eldon et al., 2008; Newbrander et al., 2011; Ranson et al., 2007a).

The choice of aid instruments and aid management modality in fragile situations also reflect their transitional states and the need for different mechanisms. Typically, fragile countries move through a number of different stages including relief, rehabilitation, reconstruction, and development and there is always a dilemma between the short-term need for ensuring essential services and the long-term need for state capacity building. As Rocha-Menocal (2009) commented:

"In fragile and conflict-affected states, there may be tensions between the imperative to provide basic services to the population urgently, by any means, and the imperative to prioritise state-building...Second, at least in the short term, faced with weak state systems and low capacity, international agencies may find it necessary to further bypass the state by funding urgent services through international NGOs to generate quick and visible improvements in living conditions" (p.3).

A similar opinion is also found in case of the fragile situation of South Sudan, as Pantuliano (2009) commented:

"It is clear that the conventional aid architecture is ill equipped to cater for a situation that spans humanitarian and development needs. As in many 'post-conflict' countries, Southern Sudan still needs support for direct service delivery, alongside support for the building of government capacity" (p.1).

It was, therefore, not surprising that Timor-Leste's health sector saw multiple aid instruments and aid disbursement mechanisms with their own strengths, weaknesses and challenges.

Although it is generally expected that a SWAp or pooled funding mechanism offers better opportunities for sector coordination, harmonization and alignment to the country's systems than the vertical project approaches, the results for the HSSP-SP from this research indicates that such opportunities for implementing the Paris principles of aid effectiveness were not fully utilized through the process of implementing the HSSP-SP. Such difficulty of implementing a pooled funding mechanism in fragile situations is also documented in Chapman and Vaillant (2010):

"Pooled funds have been widely used in the countries studied, often as a cost-effective way to mobilise resources and coordinate response in the absence of sound host government systems. Despite being strongly endorsed at corporate level as demonstrating DFID’s compliance with harmonized aid practices, the experience at country level is that they are slow to set up, costly to manage, and (with some notable exceptions) have achieved modest development impact” (p. ix).
Chapman and Vaillant (2010) also identified some typical reasons of such weak performance of a pooled funding mechanism in a fragile situation saying that:

“Operating in a fragile state is typically more labour-intensive and expensive than elsewhere for a variety of reasons including the weakness of host governments, a risky operating environment, difficult communications and ill-adapted internal procedures and regulations” (p.xiii).

The contextual analysis, fragile situation, and capacity challenge of Timor-Leste’s health sector also confirm a challenging environment for the implementation of HSSP-SP. HSSP-SP, in fact, provides an example of good intent but weak programme theory and weak implementation of a health SWAp in a fragile situation.

It can be seen that apart from the questions of sustainability and amount of investments on health system strengthening, the NAP demonstrated better compliance with the Paris principles and better performance efficiency than HSSP-SP and IPL in achieving results against the targets as well as the ratio of cost per DALY saved against the target scenario. However, the question remains if the modality used by the Global Fund grant management was most effective in strengthening the country’s health system and if the results obtained from the Global Fund grant would sustain by the country’s ownership and capacity if the external funding discontinues. As Newbrander et al. (2011) identified:

“GHPs [Global Health Partnerships such as the Global Fund] can aid fragile states in filling gaps, such as restarting a national tuberculosis (TB) programme with a grant from the Global Fund to Fight AIDS, Tuberculosis and Malaria. Potential disadvantages are that such programmes may not be integrated into basic health services, may not be sustainable, or may not provide support for health system development” (p.652).

WHO also noted the potential and effect of the global health initiative such as the Global Fund on the recipient country’s health systems especially in a fragile situation:

“Since 2000, the emergence of several large disease-specific global health initiatives (GHIs) has changed the way in which international donors provide assistance for public health. Some critics have claimed that these initiatives burden health systems that are already fragile in countries with few resources, whereas others have asserted that weak health systems prevent progress in meeting disease-specific targets” (WHO Maximizing Positive Synergies Collaborative Group, 2009, p. 2137).

Like this thesis, these studies also indicate the need for the global health initiative approach in a fragile situation but also highlight the areas of improvement for such an approach.
According to the current research, the weighted scores for IPL on the aspect of compliance with the Paris principles (before converting the scores to equal scale) were the lowest among the three projects. However, IPL obtained good scores on most aspects of economy, efficiency, effectiveness and sustainability. Due to a small budget and a clear exit plan for handing over the project to the MoH, IPL scored highest among the three projects on the aspect of sustainability. However, the cost effectiveness analysis and comparing its achievements against the targets showed that technical efficiency of IPL’s performance was the lowest among the three projects.

Considering the country’s fragile situation, unexplained poor performance in the districts not covered by IPL intervention, and a relatively short period of engagement by IPL, the appropriateness of the project modality used by IPL needs to be carefully weighed against its short term benefits and long term impact on the country’s health system. As Witter (2012) put it:

“Contracts for services to non-state actors may produce short-term benefits in terms of enhanced service delivery, but there are problems in terms of building sustainable service-delivery systems for the long term” (p.2374).

The results from the comparative evaluation of the three projects and findings from the other studies indicate that with their own strengths and challenges different aid management modalities need to work together in a fragile situation to provide necessary synergy towards meeting short term needs and achieving long term development. In their evaluation of the Paris principles, Wood and Betts (2013) also had a similar opinion:

“With respect to aid modalities, the evaluation found that no single modality (e.g., budget or sector support, program[me]s, or projects) will automatically produce better development results, and a mix of aid modalities has continued to make sense for partner countries and donors” (p.113).

Other than a dichotomous approach to complying with the Paris principles, this research found that different aid management modalities have different degrees of compliance with the different aspects of the Paris principles of aid effectiveness. The relationship between the Paris principles of aid effectiveness and development outcomes of following these principles in aid funded interventions have largely remained as a theory both in academic studies and in development practice. This is probably because of the fact that so far the Paris principles have not been fully implemented anywhere and a
The great majority of aid effectiveness evaluations have been dominated by macro-economic cross-country regression methods.

Although there are underlying programme theories and good indication of pathways between implementation of the Paris principles and better development results, no studies have demonstrated a linear causal link between them. Therefore, the comment still remains that the Paris principles of aid effectiveness is “presumed” rather than proven to be a “vehicle of development effectiveness” (Brown, 2012; Stern et al., 2008).

Comparing the scores of each project on their compliance with the Paris principles with their effectiveness scores in the balanced scorecards in this research did not clearly prove that better compliance with the Paris principles resulted in better programmatic outcomes and health systems strengthening. HSSP-SP showed weak compliance with the Paris principles and lower economy and effectiveness. The NAP, on the other hand, showed strong compliance with the Paris Principles and higher economy and effectiveness but lower sustainability. However, IPL, even with weak compliance with the Paris Principles, demonstrated higher economy, effectiveness, and sustainability. This indicates that a SWAp or the design of an aid-funded intervention does not automatically guarantee better compliance with the Paris principles and programmatic outcomes. Rather, it is also important to see how the recipient country perceives the intent of a programme design and engages with the donors in the process of implementation of the intervention.

The design and amount of aid might also have an effect on the sustainability of the aid funded programme especially in a fragile state context. Although the NAP demonstrated greater effectiveness, its sustained effect would require continuous funding from the Global Fund creating an aid dependency culture. Achievement of the longer term objectives of HSSP-SP would also require continuous funding support.

Bourguignon and Sundberg (2007) pointed out:

“The causality chain [between aid effectiveness and intervention outcomes] has been largely ignored and as a consequence the relationship between aid and development has been handled mostly as a kind of “black box”. Making further progress on aid effectiveness requires opening that box” (p.316).
Despite the mixed trend of the compliance with the Paris principles, effectiveness, and sustainability of the interventions perceived by the stakeholders, the actual process of their implementation and interactions with the government mechanisms seemed to be influenced largely by the aid instruments and degree of their compliance with the Paris principles. In this context, the attempt to run a correlation evaluation in this research provides an optimistic approach, which will be more feasible if increasing numbers of projects are included in the analysis.

11.4 Rigour and limitations of this research

The theoretical framework and research methods followed in this thesis offer methodology and protocols for a theory-driven and realist evaluation of aid effectiveness employing both qualitative and quantitative techniques. This approach attempted to address issues around both experimental and non-experimental research designs.

As Victora et al. (2011) said, design of health-programme evaluations “has been dominated traditionally by experimental approaches used in medicine, in which specific individuals or clusters of people receive an intervention whereas others do not. Studies tend to be undertaken in controlled environments in which the influence of external factors is kept to a minimum or eliminated” (p.85). Experimental evaluation designs, such as RCTs, often try to mimic a ‘counterfactual’ by comparing a group with a hypothetical situation that would occur in the absence of an intervention and tend to assign participants randomly to a ‘case’ and a ‘control’ group such as ‘before and after intervention groups’ or ‘with and without intervention’ groups in order to estimate an unbiased and internally valid estimate of outcomes (Khandker et al., 2010).

While RCTs are believed to be the ‘gold standard’ in clinical outcome efficacy trials where “the causal chain between the agent and the outcome is relatively short and simple and where results may be safely extrapolated to other settings” (Victora et al., 2004), use of RCTs in evaluating large scale national level interventions is not often practical due to the feasibility, ethical, and often political and economic challenges associated with having a control group (Prowse, 2007). For example, for the large-scale national level interventions such as the NAP or HSSP-SP, it would not be possible to find
an identical health system or a homogeneous control group without the presence of an intervention.

In general, the positivist approaches, such as the RCTs, focus on the agent and outcome and are able to tackle ‘what’ and ‘where’ questions. But RCTs cannot capture the context or conditions of an intervention and; therefore, cannot answer the ‘why’ and ‘how’ questions (Ellis, 2000). Since all contexts are different, for the large-scale public health interventions there is no guarantee that what worked in one context would work the same way in another context. Therefore, findings from the RCTs cannot provide all necessary information for replication or scaling up large-scale public health and international development interventions.

As described in chapter 3, this research employed qualitative, interpretivist and social-constructivist approach in analyzing context and mechanisms of interventions. It then employed some quantitative techniques and a quasi-experimental design in estimating counterfactuals with mathematical and epidemiological modeling for conducting outcome evaluations with cost-effectiveness analysis. Use of some other analytical tools such as balanced scorecards and DEA helped deeper and broader constructions of understanding of the process and outcomes. Thus the mixed methods employed in this research shed light on the question to what extent a policy was followed in implementation of an intervention and if following the policy resulted in desired outcomes.

The advantages of employing this theory based evaluation and mixed methods approach to this research are fairly comparable with the list of advantages of the realist evaluation approach as identified by Marchal et al. (2012). First, this research provided a generic framework to systematically deconstruct an intervention into its different components and guiding principles and then to reconstruct them with the causal chains that led to the observed outcomes. This allowed comparison of different components of an intervention with those of other interventions. Thus, the evaluation focused not only on the outcomes, but also on the policies, context, implementation process, and mechanisms.
Second, the assessment of the intermediate steps and processes between the context, activities, and outcomes of an intervention backed up by complementary qualitative and quantitative data, and triangulation process could better address the attribution issues (Weiss, 1997).

Third, the detailed analysis of what worked, how, and under what conditions may enhance the transferability of the research findings to other settings and, thus, the findings can be more useful for the policy makers than the positivist answer to the question “whether aid works” (Kernick & Mannion, 2005; Stame, 2010; van Belle et al., 2010).

Fourth, the comparison of policies, processes, mechanisms, and outcomes of different interventions in the same context provides useful insight about comparative advantages of selection and implementation of an aid management modality in a fragile state context.

White (2009) argued that while a theory-based evaluation approach sheds light on the ‘why’ question, application of such an approach often remains weak. In order to overcome the weakness, White identified six principles to successful application of the theory-based evaluation approach: (1) mapping out the causal chain (programme theory); (2) understanding context; (3) anticipating heterogeneity; (4) rigorous evaluation of impact using a credible counterfactual; (5) rigorous factual analysis; and (6) use of mixed methods. As can be seen in Chapter 3, the research described in this thesis mapped out the underlying programme theories of each intervention under evaluation; tried to understand the context by analyzing the country context, state fragility and health systems; anticipated heterogeneity of policy, objectives, mechanism, and outcomes of each intervention; conducted outcome evaluation with cost effectiveness analysis by using counterfactual scenario projected with mathematical and epidemiological modeling; conducted rigorous factual analysis with programmatic data and results; and used mixed methods for research and analysis.

Apart from these rigorous criteria to ensure the quality of evaluation, this evaluative research also meets “the agreed DAC criteria for evaluating development assistance:
relevance, efficiency, effectiveness, impact and sustainability” (OECD-DAC, 2010). According to the OECD/DAC, relevance refers to the “extent to which the aid activity is suited to the priorities and policies of the target group, recipient and donor”. Effectiveness refers to a “measure of the extent to which an aid activity attains its objectives”; whereas efficiency measures “the outputs -- qualitative and quantitative -- in relation to the inputs”. Impact, on the other hand, refers to the “positive and negative changes produced by a development intervention, directly or indirectly, intended or unintended”. Finally, sustainability is concerned with “measuring whether the benefits of an activity are likely to continue after donor funding has been withdrawn” (http://www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm).

Along with investigating the compliance with the Paris principles by each intervention under evaluation, this research attempted to evaluate the relevance, economy, efficiency, effectiveness, and sustainability of each intervention through the analysis of their context, objectives, mechanism, outputs, and outcomes. Thus, this research covered the agreed OECD-DAC (2010) criteria for evaluation.

Despite advantages, there are several limitations of this study. First, the sample size for finding the correlations in this research was very small. In fact, data from the three projects constituted the absolute minimum number for a sample size to find any correlations at two degrees of freedom. Second, the balanced scorecards and comparative ranking of the projects based on the stakeholder interviews might have perception bias. As all stakeholders were not equally knowledgeable on all three projects, the number of respondents varied for each project. So, there might be a chance that the ranking given by some stakeholders were uninformed and based on the respondent’s perception or impression about a project.

Third, the cost effectiveness analysis used a number of assumptions to calculate the DALYs saved by each project. Although the median values were used in the analysis, there was always a range of uncertainties in the results. Moreover, in the absence of real control groups, the counterfactuals such as what would happen in absence of an
intervention, were, in fact, estimated by the tools and models used in the cost effectiveness analysis.

Fourth, in order to decide total DALYs saved attributable to each project, this research used “contribution analysis” with best guessed estimates as espoused by Mayne (2001). Although this approach reduces the uncertainty about the contribution made by an intervention, it cannot ascertain the attribution to an intervention. As Mayne (2001) pointed out “in most cases what we are doing is measuring with the aim of reducing uncertainty about the contribution made, not proving the contribution made” (p.21).

Fifth, the DEA technique used in Chapter 9 assumed a linear relationship among the outputs and inputs considered in the model and a constant return to scale. However, these assumptions may not be always true in reality, especially when different projects with different objectives are compared. The results of DEA model also depend on the selection (or omission) of inputs and outputs and therefore cannot be taken as absolute.

Sixth, in a few cases for IPL and HSSP-SP where a detailed expenditure report was not available, the calculation used alternative cost information such as total disbursement received by the project. This disbursement could be slightly higher than the actual expenditure as the project would have some cash balance left at the end of the evaluation period. However, since the project IPL was closed it was assumed that IPL utilized 100% of its allocated budget by the end of the project period.

Finally, the research did not consider all possible factors that might have confounding effects on the results. As Chapman and Vaillant (2010) commented:

“Monitoring and evaluation in fragile states is different: the non-linear nature of peace processes may confound conventional logic models, while data availability is compounded by weak statistical systems, limited field access, and the risk of “doing harm” in collecting data” (p.xiv).

This research found indications that fragile situation and health system's weaknesses affected the performance of IPL due to the shortage of vaccines and deteriorated performance in non-IPL focus districts. There was also indication that increased road accidents and injury associated with construction of roads and increased number of vehicles affected HSSP-SP forecast of reduced burden of diseases in Timor-Leste. These
indications provide examples of possible confounding factors that would require more sophisticated statistical modeling to consider.

These limitations, in fact, represent the challenges in the real life evaluation scenarios especially in a fragile situation where all required information is not always available and it may not be possible to find or set up a control group (for example, for a country-wide national programme). Nevertheless, this study tried to estimate the best-informed cases by utilizing, analyzing and triangulating available information from multiple sources and by employing multiple analytic techniques including both qualitative and quantitative analysis.

11.5 Conclusion

In a fragile state’s health sector context such as the health sector of Timor-Leste this research attempted to pilot a theory driven realist evaluation approach for a comparative analysis of the effectiveness of three different management modalities of external funding to see which modality worked better for complying with the Paris principles of aid effectiveness, achieving intervention objectives, and strengthening the health systems. This research also wanted to see if there could be a feasible approach to see if the compliance with the Paris principles was correlated with achievement of intervention outcomes and health system strengthening outcomes.

It was evident that the country context, state fragility, and health systems influenced the outcomes of interventions. However, it appears from the findings and discussion that aid management modality matters. It seems in the same context and for the same objectives different mechanisms of aid management would produce different results. Therefore, finding the right match between context, mechanism, and expected outcome of an aid-funded intervention is important.

The design of an aid instrument or aid management modality by themselves cannot ensure compliance with the principles of aid effectiveness and guarantee best possible results. While the SWAp, by definition, was expected to be more compliant with the Paris principles (Chansa et al., 2008), stakeholders actually viewed it less compliant with the Paris principles than the NAP. It seems that the adherence to the aid effectiveness
principles is not determined by the design of an aid-funded intervention; rather, it is largely determined by the intervention’s interaction with the context and how the intervention is implemented.

It seems that there are inherent tensions between two sub-sets of the Paris principles. In one sub-set there are the principles of country ownership, alignment with recipient country systems, and harmonization of donor efforts. In the other sub-set includes the principles of managing for results, and mutual accountability. Resolving the tension between these two sub-sets implies the need for the readiness by the recipient country in terms of having adequate capacity, country systems, and policies in place to be able to lead development initiatives, allow donors to align and harmonize their efforts, and meet requirements for result based management and mutual accountability.

Absence of such readiness creates a vicious cycle in fragile situations where donors tend to employ multiple aid mechanisms and impose increasing conditions on the recipient country to avoid risks, and thus contribute to the increased burden on the limited capacity of the recipient country. It was found that none of the interventions fully adhered to the Paris principles in Timor-Leste’s health sector. It seems that the simultaneous presence of a health SWAp and multiple other aid management modalities outside the umbrella of SWAp at the same time defeated the purpose of a SWAp and made it more difficult for the recipient country to lead alignment and harmonization.

However, it appears that even in a fragile situation with capacity constraints, compliance with the Paris principles of aid effectiveness results in better efficiency and cost effectiveness outcomes.

11.5.1 Recommendations for aid projects

The following are the key recommendations arising from this research.

1. HSSP-SP can improve its programme design by simplifying and clarifying the stated objectives and results framework linked to its activities. An independent evaluation of Australian aid to Timor-Leste recommended for a “limited set of clear and compelling development objectives” (Office of Development Effectiveness, 2014, p. 5) for aid-funded interventions in Timor-Leste. It looks
like HSSP-SP learnt its lessons and in its amended continuation project known as NHSSP-SP, it revised its overall objective as “to support Government of Timor-Leste to get more resources to where they are needed most to improve the delivery of health services in districts and sub-districts” (World Bank, 2013k). This revised objective better links the additional inputs to address the equity issue of the health sector performance. Accordingly, NHSSP-SP also revised their performance framework with inclusion of process indicators that can be more readily attributed to NHSSP-SP funding. Still, there is room for improving the underlying programme theory of NHSSP-SP to link its activities with the outcome and impact indicators.

2. As the manager of the MDTF for HSSP-SP, the World Bank can also improve its compliance with the Paris principles, especially on the aspects of country ownership and alignment. This will require revisiting and simplifying World Bank’s administrative requirements and ways of engaging with the government in implementing this project.

3. The research findings imply that along with addressing the issue of sustainability, the NAPs can do better by paying attention to the “doing no harm” principle for their engagement in a fragile situation by avoiding actions that might have negative impacts on the country’s health systems. Manning and Trzeciak-Duval (2010) offered a list of such actions that might have negative impacts on the country including “substituting for partner capacity instead of building on existing capacity; imposing overly complex procedures and reporting requirements that burden partner country administrations; creating parallel administrative structures and project implementation units” (p. 128).

4. So far there have been limited efforts for harmonization and alignment from the NAPs in Timor-Leste and the Global Fund did not take part in any joint assessments with other donors in Timor-Leste. This stand-alone and vertical nature of the NAP might have achieved short-term results. But with fragmentation and missing synergies this relief-type approach may not produce the best results for the country in the long run. Therefore, the Global Fund may
need to modify its approach to suit to the country’s transitional stage from relief to development. In its new funding model, the Global Fund, in fact, is looking for enhanced engagement with the ongoing country dialogue. It is expected that this enhanced engagement by the Global Fund means greater appreciation of the country context, taking part in joint country assessments, and greater efforts on policy coherence, harmonization and alignment rather than an enhanced focus on the micro-management of project level activities.

5. It is recommended that interventions like IPL have a longer period of engagement, and a wider scope of work for addressing the country’s fragile situation including the health systems with greater compliance with the Paris principles of aid effectiveness. It can be seen from the results of IPL that the state fragility context, which was beyond IPL’s scope to address, came into play in achieving its results. Principle 9 of the Principles of good international engagement in fragile states recommends to “stay engaged long enough to give success a chance” (OECD, 2007d). However, IPL’s project period was only for 30 months and IPL was only contracted with limited scope for the 7 selected districts.

6. It appears from the research that long-term effectiveness of all three projects could be enhanced by looking at their contributions towards greater ‘policy coherence’. Policy coherence refers to “the systematic promotion of mutually reinforcing policies across government departments and agencies creating synergies towards achieving the defined objective” (Lockhart, 2011). In a fragile situation, policy coherence among different sectors aligned to the country’s peace-building and state-building goals is very important for achieving certain policy goals (Lockhart, 2011). The New Deal for engagement in fragile states (International Dialogue on Peacebuilding and Statebuilding, 2011) also agreed to use the peace-building and state-building goals as an important foundation towards achieving MDGs in the fragile states. None of the comparing projects, however, attempted to connect their objectives to the peace-building and state-

building goals of Timor-Leste to achieve greater policy coherence. It is recommended that aid funded interventions in a fragile state consciously link their objectives to the peace-building and state-building goals of the country and put efforts for greater policy coherence towards a whole of government approach.

11.5.2 Implications for further research

The following recommendations for further research arise from this thesis:

1. As discussed earlier, this research offers a generic framework and protocols for comparative evaluation of aid effectiveness. Although conducted in a fragile situation context, with modifications suggested in Chapter 10, this evaluation approach can be employed and tested in other situations also.

2. As this research was done with a limited sample of three projects, the statistical confidence to infer if there is any correlation between the compliance with the principles of aid effectiveness and development outcomes would be improved with a larger sample. Further research can be done by randomly selecting a sizable number of aid funded projects from a single country context and then compare their degree of compliance with the aid effectiveness principles and intervention outcomes. Another approach of addressing the sampling issue will be to undertake a good number of single country comparative evaluation studies such as the current one, and then synthesize and analyze the findings from all the single country studies.

3. The current research tracked the results of aid-funded interventions for a limited period of time and assesses the likely sustainability of results based on stakeholders’ opinions. It will be useful to conduct a follow up study to investigate the actual sustainability of results and see if this is associated with the principles of aid effectiveness or aid management modality. The current version of the research framework and methods can be employed for such a repeat study.

4. Like a dashboard, the findings of the current research only highlighted the broad indicators and units of performance measurements. These indicators can only
identify the areas of relative performance of an intervention. However, more in-depth investigation will be required to identify the factors behind the relatively good or poor performance by an intervention and if lessons learned from such performance variations can be replicated and transferred in other contexts.

5. Although a list of recommendations is offered in this thesis to improve the effectiveness of each intervention, implementation of these recommendations can be tested through operational research and further evaluation of their outcomes.

6. This research indicated possible effects of state fragility, health system weaknesses, and other confounding factors (such as increased road accidents associated with post-conflict development surge) on the performance of aid funded projects. Research can be undertaken to develop more sophisticated statistical modeling and scenario analysis tools to consider such confounding factors for outcome effectiveness trials. Such tools can also focus on finding if outcomes of an intervention were influenced by weak design and underlying programme theory, or weak implementation, or by the external confounding factors.
12 References


Dujardin, B. (2009). Sector-wide approach and health policy reforms: SWAP is the answer, but what is the question. Tropical Medicine and International Health, 14, 17-18. doi: http://dx.doi.org/10.1111/j.1365-3156.2009.02352.x


Easterly, W. (2006). The white man’s burden : why the West’s efforts to aid the rest have done so much ill and so little good. Oxford: Oxford University Press.


Grap-Pa Sante. (2011). What results can be reasonably expected from the implementation of the Paris Declaration in the health sector by 2011? Paper presented at the OECD Technical Team on Health as a Tracer Sector.


## Annex 1: Questionnaire for stakeholder interviews

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Criteria</th>
<th>Scoring Description</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input policy and mechanism</strong></td>
<td>1. Degree of ownership by recipient country</td>
<td>The project is not guided by a country-owned long-term vision linked to a medium-term strategy, and there is little to no effort within the country to develop or update these strategic instruments. A medium-term strategy is under preparation, but may not yet be derived from a country-owned long-term vision. Sector related project strategies are few, and may not yet be tied into a medium-term strategy. A strategic framework may be guiding short-term project activities. There is a long-term vision and a medium-term strategy or strategies that may not be linked. Project strategies may not yet be integrated into sector strategy and national development strategy. The role of different strategy instruments in guiding policy is unproven, unclear, or provisional. Where they exist, efforts to align local with national strategy are preliminary.</td>
<td>1.</td>
</tr>
<tr>
<td><strong>Input policy and mechanism</strong></td>
<td>2. Alignment: degree to which there is a serious effort to connect aid programmes with country policies and processes</td>
<td>Total volume of project aid is not recorded in country’s annual sector budget. Project aid is not considered in sector planning and budget preparation but not in Total volume of project aid is recorded separately but not in country’s annual sector budget. Project aid is sometimes considered in sector planning and budget preparation. Total volume of project aid is recorded separately and mentioned in country’s annual sector budget. Project aid is often considered in sector planning and budget preparation. Total volume of project aid is recorded in country’s annual sector budget. Project aid is considered in sector planning and budget preparation following some systematic procedure. Total volume of project aid has been recorded and integrated in country’s annual sector budget. Project aid is fully integrated in sector planning and budget preparation.</td>
<td>5.</td>
</tr>
</tbody>
</table>

1
<table>
<thead>
<tr>
<th>Dimension</th>
<th>Criteria</th>
<th>Scoring Description</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input policy and mechanism</td>
<td>3. Degree of alignment of donor technical cooperation with the partner country’s capacity development objectives and strategies</td>
<td>Project technical cooperation did not consider the country’s capacity development objectives and strategies. There are no clearly identifiable inputs and activities in the project specifically targeted to health systems strengthening.</td>
<td></td>
<td>a systematic way</td>
<td>preparation and follows some systematic procedure</td>
<td></td>
<td></td>
<td>HSSP-SP</td>
</tr>
<tr>
<td>Input policy and mechanism</td>
<td>4. Alignment: Avoid a dedicated management units designed to support development projects or programmes</td>
<td>The project uses a separate and dedicated management unit designed to implement the project activities.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NAP</td>
</tr>
<tr>
<td>Input policy and mechanism</td>
<td>5. Alignment: Use of country’s own financial</td>
<td>The project does not use country’s own financial management.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>IPL</td>
</tr>
</tbody>
</table>

- **Input policy and mechanism 3.** Degree of alignment of donor technical cooperation with the partner country’s capacity development objectives and strategies:
  - Project technical cooperation did not consider the country’s capacity development objectives and strategies. There are no clearly identifiable inputs and activities in the project specifically targeted to health systems strengthening.
  - Project technical cooperation considered the country’s capacity development objectives and strategies but not yet aligned to the country’s capacity development objectives and strategies. There are very few clearly identifiable inputs and activities in the project specifically targeted to health systems strengthening.
  - Project technical cooperation is aligned to the country’s capacity development objectives and strategies most of the time. There are considerable number of clearly identifiable inputs and activities in the project specifically targeted to health systems strengthening.
  - Project technical cooperation is consistently aligned to the country’s capacity development objectives and strategies. There are adequate number of fully aligned and clearly identifiable inputs and activities in the project specifically targeted to health systems strengthening.

- **Input policy and mechanism 4.** Alignment: Avoid a dedicated management units designed to support development projects or programmes:
  - The project uses a separate and dedicated management unit designed to implement the project activities.
  - The project uses a dedicated facilitation unit set up within the Ministry of Health designed to support and manage project activities.
  - The project uses regular implementation mechanisms of the Ministry of Health with some facilitation provided by a common facilitation unit.
  - The project avoids a dedicated management units and project activities are fully carried out by regular mechanism of the Ministry of Health.

- **Input policy and mechanism 5.** Alignment: Use of country’s own financial:
  - The project does not use country’s own financial management.
  - The project uses country’s own financial management.
  - The project uses country’s own financial management.
  - The project uses country’s own financial management.
<table>
<thead>
<tr>
<th>Dimension</th>
<th>Criteria</th>
<th>Scoring Description</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>management, procurement and monitoring and evaluation systems</td>
<td>management, procurement and monitoring and evaluation systems, and no efforts are made to strengthen these systems so they can be used to a greater extent over time</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>management, procurement and monitoring and evaluation systems, but some efforts are made to strengthen these systems so they can be used to a greater extent over time</td>
<td>procurement and monitoring and evaluation systems to a greater extent, and significant efforts are made to strengthen these systems so they can be used fully over time</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>management, procurement and monitoring and evaluation systems to a limited extent, and significant efforts are made to strengthen these systems so they can be used to a greater extent over time</td>
<td>procurement and monitoring and evaluation systems, and significant efforts are made to further strengthen these systems over time</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>management, procurement and monitoring and evaluation systems, and no efforts are made to strengthen these systems so they can be used to a greater extent over time</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>management, procurement and monitoring and evaluation systems to a limited extent, and significant efforts are made to strengthen these systems so they can be used to a greater extent over time</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>management, procurement and monitoring and evaluation systems to a greater extent, and significant efforts are made to strengthen these systems so they can be used fully over time</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>management, procurement and monitoring and evaluation systems, and significant efforts are made to further strengthen these systems over time</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Input policy and mechanism</td>
<td>6. Harmonization: Takes part in joint mission and joint country analytic work</td>
<td>The project never takes part in joint mission and joint country analytic work with other development partners. Project activities are never harmonized with other similar initiatives in the country</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>The project rarely takes part in joint mission and joint country analytic work with other development partners. Project activities are rarely harmonized with other similar initiatives in the country</td>
<td>The project sometimes takes part in joint mission and joint country analytic work with other development partners. Project activities are sometimes harmonized with other similar initiatives in the country</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>The project often takes part in joint mission and joint country analytic work with other development partners. Project activities are sometimes harmonized with other similar initiatives in the country</td>
<td>The project always takes part in joint mission and joint country analytic work with other development partners. Project activities are always harmonized with other similar initiatives in the country</td>
<td>5</td>
</tr>
<tr>
<td>Input policy and mechanism</td>
<td>7. Managing for results: Extent of using a results-based monitoring and evaluation (M&amp;E) system that, building on sound statistical data and open access to</td>
<td>Data collection is sporadic and outdated. Data have little relation to tracking the goals and targets in the long-term vision and medium-term strategy.</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Data collection is improving but largely restricted to limited geographic or sectoral areas. Data may not cover key goals and targets in the long-term vision and medium-term strategy.</td>
<td>Data collection has become more systematic and efforts to extend its geographic or sectoral scope are underway. Data are increasingly related to tracking goals and targets in the long-term vision and medium-term strategy.</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Data are generally timely and comprehensive, and directly related to tracking the achievement of country goals and targets identified in the long-term vision and medium-term strategy.</td>
<td>There are no warning signs of possible deterioration, and there is widespread expectation that the progress achieved is sustainable.</td>
<td>5</td>
</tr>
<tr>
<td>Dimension</td>
<td>Criteria</td>
<td>Scoring Description</td>
<td>Score</td>
</tr>
<tr>
<td>-----------</td>
<td>----------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------</td>
<td>-------</td>
</tr>
<tr>
<td></td>
<td>information, produces data on progress toward desired inputs, outputs, and outcomes of the project</td>
<td>medium-term strategy.</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>coordinated and systematic data gathering and analysis.</td>
<td>2</td>
</tr>
<tr>
<td>Input policy and mechanism</td>
<td>8. Mutual accountability: undertake mutual assessments of progress in implementing agreed commitments</td>
<td>There are limited or no evidence that the donor and recipient undertake mutual assessments of progress in implementing agreed commitments but not on a regular basis</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>There are some evidence that the donor and recipient undertake mutual assessments of progress in implementing agreed commitments but not on a regular basis</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>There are growing evidence that the donor and recipient sometimes undertake mutual assessments of progress in implementing agreed commitments</td>
<td>5</td>
</tr>
<tr>
<td>Activities</td>
<td>9. Economy: Unit costs</td>
<td>There are no warning signs of possible deterioration, and there is widespread expectation that the progress achieved is sustainable.</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>There are no warning signs of possible deterioration, and there is widespread expectation that the progress achieved is sustainable.</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>There are no warning signs of possible deterioration, and there is widespread expectation that the progress achieved is sustainable.</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>There are no warning signs of possible deterioration, and there is widespread expectation that the progress achieved is sustainable.</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>There are no warning signs of possible deterioration, and there is widespread expectation that the progress achieved is sustainable.</td>
<td>5</td>
</tr>
<tr>
<td>Activities</td>
<td>10. Economy: Management and transaction costs</td>
<td>There are no warning signs of possible deterioration, and there is widespread expectation that the progress achieved is sustainable.</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>There are no warning signs of possible deterioration, and there is widespread expectation that the progress achieved is sustainable.</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>There are no warning signs of possible deterioration, and there is widespread expectation that the progress achieved is sustainable.</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>There are no warning signs of possible deterioration, and there is widespread expectation that the progress achieved is sustainable.</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>There are no warning signs of possible deterioration, and there is widespread expectation that the progress achieved is sustainable.</td>
<td>5</td>
</tr>
<tr>
<td>Dimension</td>
<td>Criteria</td>
<td>Scoring Description</td>
<td>Score</td>
</tr>
<tr>
<td>-----------</td>
<td>----------</td>
<td>---------------------</td>
<td>-------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>explain and justify additional cost; Cost exceeds BM by wide margin and represents poor return</td>
<td>adequate returns</td>
</tr>
<tr>
<td>Activities</td>
<td>11. Economy: Procurement</td>
<td>No discernable use of procurement to manage or reduce costs</td>
<td>Some identifiable management of costs through procurement; Ongoing monitoring of procurement costs not identified; Little or no assessment of effect of procurement savings on outputs/outcomes;</td>
</tr>
<tr>
<td>Outputs</td>
<td>12. Project outputs: Percentage of achievement of different project outputs against their targets</td>
<td>Unacceptable (&lt;30%)</td>
<td>Inadequate but potential demonstrated (30% to 59%)</td>
</tr>
<tr>
<td>Outputs</td>
<td>13. Health system outputs: Percentage of achievement of health system outputs (related</td>
<td>None or very insignificant outputs</td>
<td>A few outputs; Little significance</td>
</tr>
<tr>
<td>Dimension</td>
<td>Criteria</td>
<td>Scoring Description</td>
<td>Score</td>
</tr>
<tr>
<td>-----------</td>
<td>----------</td>
<td>---------------------</td>
<td>-------</td>
</tr>
<tr>
<td></td>
<td>to service delivery, workforce, information, supplies and commodities, financing, and governance) against the targets established through participatory impact pathway analysis</td>
<td></td>
<td>HSSP-SP</td>
</tr>
<tr>
<td>Outputs</td>
<td>14. Efficiency: Productivity measure</td>
<td>Cost of activities/outputs higher than similar programmes and no mitigation factors identified; No evidence that value of outputs optimised; Weak or no activity schedule and milestones; Weak, no efficiencies and very poor input-output ratios</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Cost of activities/outputs higher than similar programmes and few mitigation factors identified; Little evidence that value of outputs is optimised; Activity schedule and milestones insufficiently well-planned for delivering timely outputs; Poor productivity, with no efficiencies achieved</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cost of activities/outputs comparable with similar programmes; Some evidence that value of some outputs is optimised (e.g. through timing of delivery, increase in proportion of output; decrease in proportion of input); Activities planned in integrated, sequenced way but milestones poor on timing and delivery; Adequate productivity with some efficiencies</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cost of activities/outputs comparable with similar programmes; Good evidence that value of some outputs is optimised (e.g. through timing of delivery, increase in proportion of output; decrease in proportion of input); Integration and sequencing of activities supports delivery and measurement of productivity (actual + planned); Efficient with good</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cost of activities/outputs comparable with similar programmes; Strong evidence that value of critical outputs is optimised (e.g. through timing of delivery, increase in proportion of output; decrease in proportion of input); Integration and sequencing of activities supports delivery and measurement of productivity (actual + planned); Very efficient with</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dimension</td>
<td>Criteria</td>
<td>Scoring Description</td>
<td>Score</td>
</tr>
<tr>
<td>-----------</td>
<td>---------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------</td>
<td>--------------------------------------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Outputs</td>
<td>15. Efficiency: Business lead time—Regular activities</td>
<td>Very long business lead time compared with benchmarked lead time (BM); No mitigating factors identified which explain and justify additional time requirement; Lead time exceeds BM by wide margin and represents poor added value</td>
<td>Business lead time is above BM; Few mitigating factors explained which justify additional time requirement; Lead time exceeds BM and is not adding adequate value returns</td>
</tr>
<tr>
<td>Outputs</td>
<td>16. Efficiency: Business lead time—Exceptional activities</td>
<td>Very long business lead time compared with benchmarked lead time (BM); No mitigating factors identified which explain and justify additional time requirement; Lead time exceeds BM by wide margin and represents poor added value</td>
<td>Business lead time is above BM; Few mitigating factors explained which justify additional time requirement; Lead time exceeds BM and is not adding adequate value returns</td>
</tr>
<tr>
<td>Outcomes</td>
<td>17. Project outcome indicators: Percentage of achievement of strategic</td>
<td>Unacceptable (&lt;30%)</td>
<td>Inadequate but potential demonstrated (30% to 59%)</td>
</tr>
<tr>
<td>Dimension</td>
<td>Criteria</td>
<td>Scoring Description</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>--------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>objectives of the intervention</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outcomes</td>
<td>18. Health system outcome: Extent of improvement of related health service delivery system (access, coverage, quality and equity) Extent of improvement of health system’s responsiveness on the related issue Extent of improved efficiency of the health system on the related issue</td>
<td>None, very insignificant, or negative outcomes</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Very little outcomes; Little significance</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Some outcomes; Moderately significant</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Considerable outcomes; With significance</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Highly significant outcomes</td>
<td></td>
</tr>
<tr>
<td>Effectiveness</td>
<td>19. Effectiveness: Relevance and Robustness (Relevance= clear, planned, causally linked, and system-wide. Robust=data to support indicators are available, accessible,</td>
<td>Indicators and achievements are neither relevant nor robust.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Indicators and achievements have many significant weaknesses in terms of relevance and robustness.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Indicators and achievements have some significant weaknesses in terms of relevance and robustness.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Indicators and achievements are mostly relevant and robust</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Indicators and achievements are relevant and robust</td>
<td></td>
</tr>
<tr>
<td>Dimension</td>
<td>Criteria</td>
<td>Scoring Description</td>
<td>Score</td>
</tr>
<tr>
<td>-----------</td>
<td>----------</td>
<td>---------------------</td>
<td>-------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Effectiveness</td>
<td></td>
<td>credible, and disaggregate-able)</td>
<td></td>
</tr>
<tr>
<td>Sustainability</td>
<td></td>
<td>Little or no likelihood outputs will deliver purpose; Too little information on assumptions to assess effects on outcomes; Risk of not achieving purpose very high</td>
<td></td>
</tr>
<tr>
<td>Sustainability</td>
<td></td>
<td>Outputs do not meet ‘necessary and sufficient’ rule; Assumptions are questionable and insufficiently detailed; Risk of not achieving purpose high</td>
<td></td>
</tr>
<tr>
<td>Sustainability</td>
<td></td>
<td>Outputs are necessary and sufficient to deliver purpose; Some assumptions about externalities realistic and credible; some questions about coverage and/or depth; Some risk of underachieving but managed to enable achievement of purpose</td>
<td></td>
</tr>
<tr>
<td>Sustainability</td>
<td></td>
<td>Outputs are necessary and sufficient to deliver purpose; Realistic and credible assumptions about externalities, good coverage and depth; Low risk of underachieving; likely will achieve purpose</td>
<td></td>
</tr>
<tr>
<td>Sustainability</td>
<td></td>
<td>Outputs are necessary and sufficient to deliver purpose; Realistic and credible assumptions, analyzing key externalities in sufficient depth; Probable will achieve or exceed purpose</td>
<td></td>
</tr>
<tr>
<td>Sustainability</td>
<td></td>
<td>Leverage/Replication</td>
<td></td>
</tr>
<tr>
<td>Sustainability</td>
<td></td>
<td>No leverage or wider effects identified; No or very low potential for additional benefits (e.g. scale-up, multiplier or replication) identified</td>
<td></td>
</tr>
<tr>
<td>Sustainability</td>
<td></td>
<td>Some leverage of other activities/investment and wider effects identified; Limited potential for additional benefits (e.g. scale-up, multiplier or replication) identified</td>
<td></td>
</tr>
<tr>
<td>Sustainability</td>
<td></td>
<td>Leverage of other activities/investments described and supported by some evidence; Some potential for additional benefits (e.g. scale-up, multiplier or replication) identified</td>
<td></td>
</tr>
<tr>
<td>Sustainability</td>
<td></td>
<td>Leverage of other activities/investments and wider effects described, and supported by strong evidence; Considerable potential for additional benefits (e.g. scale-up, multiplier or replication) identified</td>
<td></td>
</tr>
<tr>
<td>Sustainability</td>
<td></td>
<td>Leverage of other activities/investments and wider effects described with evidence that shows significant potential for expansion or replication; Very high potential for additional benefits (e.g. scale-up, multiplier or replication) identified</td>
<td></td>
</tr>
<tr>
<td>Sustainability</td>
<td></td>
<td>It is very unlikely that outputs delivered through the project will be</td>
<td></td>
</tr>
<tr>
<td>Sustainability</td>
<td></td>
<td>It is unlikely that outputs delivered through the project will be sustained by</td>
<td></td>
</tr>
<tr>
<td>Sustainability</td>
<td></td>
<td>It is somehow likely that outputs delivered through the project will be</td>
<td></td>
</tr>
<tr>
<td>Sustainability</td>
<td></td>
<td>It is likely that outputs delivered through the project will be sustained by national</td>
<td></td>
</tr>
<tr>
<td>Sustainability</td>
<td></td>
<td>It is highly likely that outputs delivered through the project will be sustained by</td>
<td></td>
</tr>
<tr>
<td>Dimension</td>
<td>Criteria</td>
<td>Scoring Description</td>
<td>Score</td>
</tr>
<tr>
<td>-----------</td>
<td>--------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------</td>
<td>-------</td>
</tr>
<tr>
<td></td>
<td>through the project be sustained by national ownership and capacities, after the end of the project duration</td>
<td>sustained by national ownership and capacities, after the end of the project duration; No visible indications of sustainability</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>national ownership and capacities, after the end of the project duration; Very few indications of sustainability</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>sustained by national ownership and capacities, after the end of the project duration; Supported by few indications</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ownership and capacities, after the end of the project duration; Supported by moderate indication</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>national ownership and capacities, after the end of the project duration; Supported by strong indication</td>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>HSSP-SP</th>
<th>NAP</th>
<th>IPL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Annex 2: Informed consent form

Aid Effectiveness and Health Systems Strengthening in a Fragile State: A Comparative Analysis of Aid Management Modalities in Timor-Leste’s Health Sector

INFORMATION SHEET AND CONSENT FORM FOR PARTICIPANTS

Thank you for showing an interest in this project. Please read this information sheet carefully before deciding whether or not to participate. If you decide to participate we thank you. If you decide not to take part there will be no disadvantage to you and we thank you for considering our request.

What is the aim of the project?

This project is undertaken as part of Doctoral research at the Department of Preventive and Social Medicine of the University of Otago, New Zealand. This project aims to analyze and compare the effects of three externally funded interventions in Timor-Leste’s health sector in terms of their project outcomes and impacts on the health system and tries to identify if different outcomes are correlated with internationally agreed aid effectiveness principles.

Considering the specific fragile state context, intervention funding and management mechanisms, and their interactions with the health system, this research attempts to see what works for whom under what conditions with the following objectives—

- Define aid effectiveness and develop an evaluation framework in the context of the health system of Timor-Leste
- Identify and analyze different aid management modalities in Timor-Leste’s health sector in relation to their particular contexts, intervention plans and possible effect on the health system
- Investigate for each aid management modality the extent of following aid effectiveness principles, interactions with health systems, and their short term and long term outcomes
- Analyze the relationship between aid effectiveness principles, aid management modalities and project and health system outcomes and compare effectiveness of different health aid management modalities for their efficiency, transaction costs and development outcomes

What type of participants are being sought?

We are involving key stakeholders from different sections of Timor-Leste’s health sector, development partners, implementers and beneficiaries in the process of this evaluation. Based on the role and relevant experience we are inviting a limited number of people to take part in this study and be part of a research reference group.

Participant contribution will be duly acknowledged in the research report and publications and draft findings will be shared with the participants for their opinion before finalizing the report. Where applicable, participants will also be compensated for their local transportation costs incurred for taking part in this study.
What will participants be asked to do?

Should you agree to take part in this project, you may be asked to:

- Take part in two or three meetings of the stakeholders at different stages of the research. Each meeting may take between 60 minutes to 90 minutes.
- Take part in an interview in a place and time according to your convenience and with a questionnaire shared with you before the interview. Completion of the questionnaire may take about 40 to 60 minutes of your time.
- Occasionally give your comment and feedback by email or telephone on the areas of your knowledge, experience or expertise.

Please be aware that you may decide not to take part in the project at any stage without any disadvantage to yourself of any kind.

What data or information will be collected and what use will be made of it?

This project will ask for information on and your opinions about activities and outcomes of three particular interventions such as the Global Fund funded National HIV/AIDS Program, AusAID-World Bank funded NHSSP-SP and USAID funded MCHIP. It will ask you to rate particular inputs, outputs and outcomes of those interventions according to your opinion. Your response will be recorded by written notes and if required by audio-taping. However, no individually identifying or personal questions will be asked.

Information collected in this project will be used for scoring and evaluating the interventions for their comparative analysis. The information may be accessed by the researcher, researcher supervisors and the research reference group. There will be no commercial use of the data collected for this project.

The data collected will be securely stored in computers with hardcopy and electronic file back-ups with limited access permission by the research team only. At the end of the project any personal information will be destroyed immediately except that, as required by the University's research policy, any raw data on which the results of the project depend will be retained in secure storage for five years, after which it will be destroyed.

The research report and publications will reflect summary information and analysis and will not normally identify any person individually. On some occasions the readers of the research report might be able to identify an individual for an opinion, in such cases, on the Consent Form please indicate your preference of anonymity. Please be aware that should you wish we will make every attempt to preserve your anonymity. However, with your consent, there are some cases where it would be preferable to attribute contributions made to individual participants. It is absolutely up to you which of these options you prefer. The results of the project may be published and will be available in the University of Otago Library (Dunedin, New Zealand).

Along with a structured questionnaire, this project may also involve an open-questioning technique. The general line of questioning includes your rating of particular result indicators and associated contextual information to interpret your rating. The precise nature of the questions which will be asked have not been determined in advance, but will depend on the way in which the interview develops.

In the event that the line of questioning does develop in such a way that you feel hesitant or uncomfortable you are reminded of your right to decline any particular question(s) and also that you may withdraw from the project at any stage without any disadvantage to yourself of any kind. The draft study report will be shared with you for your opinion and you will be given chance to review and correct any opinion, which can be attributed to you, appears there.
This proposal has been reviewed and approved by the Department of Preventive and Social Medicine of the University of Otago. This has also been submitted for review and approval by the Research Ethics Committee of the Ministry of Health of Timor-Leste.

Can participants change their mind and withdraw from the project?

Again, please be assured that you may withdraw from participation in the project at any time and without any disadvantage to yourself of any kind.

What if participants have any questions?
If you have any questions about our project, either now or in the future, please feel free to contact either:
Hasibul Haque and/or Prof. Robin Gauld
Department of Preventive and Social Medicine Department of Preventive and Social Medicine
Telephone: +64 3 479 4011 Telephone: +64 3 479 8632
Email: hasibul.haque@yahoo.com Email: robin.gauld@otago.ac.nz

This study has been approved by the Department stated above. If you have any concerns about the ethical conduct of the research you may contact the Committee through the Human Ethics Committee Administrator (ph 03 479-8256). Any issues you raise will be treated in confidence and investigated and you will be informed of the outcome.
Aid Effectiveness and Health Systems Strengthening in a Fragile State: A Comparative Analysis of Aid Management Modalities in Timor-Leste’s Health Sector

CONSENT FORM FOR PARTICIPANTS

I have read the Information Sheet concerning this project and understand what it is about. All my questions have been answered to my satisfaction. I understand that I am free to request further information at any stage.

I know that:

1. My participation in the project is entirely voluntary;

2. I am free to withdraw from the project at any time without any disadvantage;

3. Personal identifying information recorded in emails and audio-tapes will be destroyed at the conclusion of the project but any raw data on which the results of the project depend will be retained in secure storage for at least five years;

4. This project may involve an open-questioning technique. The general line of questioning includes rating of result indicators of particular interventions and contextual information to interpret the rating. The precise nature of the questions which will be asked have not been determined in advance, but will depend on the way in which the interview develops and that in the event that the line of questioning develops in such a way that I feel hesitant or uncomfortable I may decline to answer any particular question(s) and/or may withdraw from the project without any disadvantage of any kind;

5. There will be no commercial use of data collected in this project;

6. The results of the project may be published and available in the University of Otago Library (Dunedin, New Zealand) but every attempt will be made to preserve my anonymity.

☐ I wish to keep complete anonymity of my identity in the research reports. (Please tick if applicable to you)

I agree to take part in this project.

............................................................................................................. ................................................
(Signature of participant) (Date)
Annex 3: Ethical approval from Timor-Leste

Dear Haque,

Re: Aid effectiveness and health strengthening in a fragile state: A comparative analysis of aid management modalities in Timor-Leste’s health sector

Thank you very much for your first presentation on Friday 13th July 2012 to the Cabinet of Health Research and Development Technical and Ethical Review Committee. The Technical and Ethical Committee gives full approval for this research project to be carried but you must consider all input. We recognize this topic is important for your study and these results will directly benefit Ministry of Health Timor-Leste.

We recommended that all input by team Technical and Ethical Committee Timor-Leste you have to consider and your finding research could you submit one book and electronic copy to the Cabinet of Health Research and Development Ministry of Health.

We would be grateful if you will present the findings of this research to the Cabinet of Health Research and Development on completion.

Yours sincerely,

Mr Valente da Silva, SKM, MPH
Director of CHRD Ministry of Health
Annex 4: Ethical approvals from the University of Otago

26 June 2012

Mr Hasibul Haque,
Department of Preventive and Social Medicine,
University of Otago

Dear Hasibul,

Re: Proposal to RAC

On Thursday 21st June, the Research Advisory Committee considered your proposal for a study of Aid Effectiveness and Health Systems Strengthening in a Fragile State: A Comparative Analysis of Aid Management Modalities in Timor Leste’s Health Sector.

I am pleased to confirm that the Committee has approved your project. We found your proposal to be very well written and very clear, and there was an impressive level of conceptual thinking demonstrated in the proposal. The project is an interesting and challenging one.

Part of the RAC’s role is to provide feedback on the project to support the research. One of your supervisors, Professor Robin Gauld is a member of the RAC. He will be able to provide more details about the discussion. One of the Committee’s thoughts was that the size of the project had the potential to be very large and we were concerned that the project be kept to a manageable size. We fully understand that this is an emerging project but we felt reassured that you had the ability to remain flexible in further development of the project.

Best wishes for the project.

Rob McGee
Convenor

Research Advisory Committee

cc. Jennie Connor, HoD
    Robin Gauld
    Phillip Hill
Dear Professor Gauld,

I am again writing to you concerning your proposal entitled “Aid Effectiveness and Health Systems Strengthening in a Fragile State: a Comparative Analysis of Aid Management Modalities in Timor-Leste’s Health Sector”, Ethics Committee reference number D12/240.

Thank you for your comprehensive response to our letter. We acknowledge that you have provided us with the following documentation: a letter outlining the documentation and assurances held by the Department relating to Hasibul Haque’s travel arrangements, insurance details and medical preparations; the NZ Government travel advisory for Timor-Leste and the WHO’s health information for travellers; and documentation from Timor-Leste governmental authorities Providing approval for the research project and Hasibul Haque’s special stay authorisation. We note your advice regarding Hasibul Haque’s experience working for the UN in Timor-Leste; and the amended Information Sheet and Consent Form which gives participants an estimation of the time involved in participating in the questionnaire and focus group.

On the basis of this response, I am pleased to confirm that the proposal has full ethical approval to proceed.

Yours sincerely,

[Signature]

Mr Gary Witte
Manager, Academic Committees
Tel: 479 8256
Email: gary.witte@otago.ac.nz

c.c. Professor J L Connor  Head  Department of Preventive and Social Medicine
Annex 5: List of unpublished project documents used as information sources

A. HSSP-SP:

1. Final minutes of Timor-Leste health donor harmonization meeting. Date and Time: 28 July 2011, 9.30 – 10:30
11. Timor-Leste: National Health Sector Strategic Plan Support Program (NHSSP-SP)


**B. NAP:**

   External Prinit Version Last Updated on 5 November 2012.

**IPL:**

1. GAVI Alliance Annual Progress Report 2012. Submitted by the Government of Timor-Leste
2. UNICEF Annual Report 2012 for Timor-Leste
11. IPL. (2011). PMP data compilation sheet and quarterly reporting format (Results Framework)
14. IPL. (2013). Project Asset Inventory Report as at 28 February 2013
15. IPL. (2013). Inventory Report: Durable Expendable Equipment (Exp <$500) as at 31 March 2013
20. JSI Timor-Leste. (2013). Staff List for Immunizasaun Proteje Labarik (IPL)