Viewpoint: How can data harmonisation benefit mental health research? An example of
The Cannabis Cohorts Research Consortium (CCRC)

Running title: CCRC Harmonised cohort

Delyse M. Hutchinson
Senior Research Fellow, National Drug and Alcohol Research Centre, UNSW Australia, Sydney, NSW, Australia. Centre for Social and Early Emotional Development, Deakin University Australia, and; Murdoch Childrens Research Institute, The Royal Children's Hospital Melbourne
d.hutchinson@unsw.edu.au

Edmund Silins (author for correspondence)
Research Fellow, National Drug and Alcohol Research Centre, UNSW Australia, Sydney, NSW, 2052, Australia.
Email: e.silins@unsw.edu.au
Telephone: +61 2 9385 0333
Facsimile: +61 2 9385 0222

Richard P. Mattick
Professor, National Drug and Alcohol Research Centre, UNSW Australia, Sydney, NSW, Australia
r.mattick@unsw.edu.au
George C. Patton
Professor, Centre for Adolescent Health, Murdoch Children’s Research Institute, Royal Children’s Hospital, Melbourne, VIC, Australia; Department of Paediatrics, University of Melbourne, Melbourne, VIC, Australia
george.patton@rch.org.au

David M. Fergusson
Professor, Christchurch Health and Development Study, Department of Psychological Medicine, University of Otago, Christchurch, New Zealand
dm.fergusson@otago.ac.nz

Reza Hayatbakhsh
Senior Research Fellow, School of Population Health, University of Queensland, Brisbane, QLD, Australia
m_r_hb@yahoo.com

John W. Toumbourou
Professor, Centre for Adolescent Health, Murdoch Children’s Research Institute, Royal Children’s Hospital, Melbourne, VIC, Australia; School of Psychology, Deakin University, Geelong, VIC, Australia
john.toumbourou@deakin.edu.au

Craig A. Olsson
Professor, Centre for Social and Early Emotional Development, Deakin University, Australia, and; Centre for Adolescent Health, Murdoch Children's Research Institute, The Royal Children's Hospital Melbourne, and The University of Melbourne, Parkville, Victoria, Australia.

craig.olsson@rch.org.au

Jake M. Najman
Professor, Schools of Population Health and Social Science, University of Queensland, Brisbane, QLD, Australia.

j.najman@sph.uq.edu.au

Elizabeth Spry
Project Manager, Centre for Adolescent Health, Murdoch Children's Research Institute, Royal Children's Hospital, Melbourne, VIC, Australia

elizabeth.spry@mcri.edu.au

Robert J. Tait
Senior Research Fellow, National Drug Research Institute, Faculty of Health Sciences, Curtin University, Perth, WA, Australia

robert.tait@curtin.edu.au

Louisa Degenhardt
Professor, National Drug and Alcohol Research Centre, UNSW Australia, Sydney, NSW, Australia

l.degenhardt@unsw.edu.au
Wendy Swift

Senior Research Fellow, National Drug and Alcohol Research Centre, UNSW Australia, Sydney, NSW, Australia
w.swift@unsw.edu.au

Peter Butterworth

Associate Professor, Centre for Research on Ageing, Health and Wellbeing, The Australian National University, Canberra, ACT, Australia
peter.butterworth@anu.edu.au

L. John Horwood

Associate Professor, Christchurch Health and Development Study, Department of Psychological Medicine, University of Otago, Christchurch, New Zealand
John.horwood@otago.ac.nz

And the Cannabis Cohorts Research Consortium (CCRC)

Keywords: data harmonisation, cannabis dependence, longitudinal study, child, adolescent, young adult.
Viewpoint: How can data harmonisation benefit mental health research? An example of

The Cannabis Cohorts Research Consortium (CCRC)

Despite significant research on cannabis use, less is known about antecedents and effects of less prevalent patterns of use, such as early daily or dependent use. Cohort studies typically have insufficient samples (i.e., small cell sizes) and lack the statistical power to examine relationships between regular use and low prevalence adverse outcomes (e.g., suicide).

Consequently, individual cohort studies offer limited opportunity to explore the causes and consequences of cannabis use patterns which confer the greatest burden of disease, namely, daily use and dependence (Curran and Hussong, 2009; Hofer and Piccinin, 2009).

One approach to addressing these limitations is to invest in new, large, prospective studies, capable of providing the power needed to examine less frequent events. However, such studies take decades to mature, are extremely costly and necessarily delay important health knowledge reaching the research field, clinicians, decision makers in government and the general population.

One alternative is to make use of the available data by harmonising and pooling individual participant data across individual cohorts. A multi-cohort consortium approach provides a number of potential advantages including: a) efficiency in the use of existing data, time and resources; b) the capacity to bring together expert knowledge from across a range of disciplinary boundaries; c) increased opportunity for knowledge translation and dissemination; d) the increased generalisability afforded by combining data collected by different researchers on different samples; and e) the opportunity to combine data from
number of studies to answer questions that cannot be answered in individual cohorts. We provide an example of such an approach.

The Cannabis Cohorts Research Consortium (CCRC): An example of a multi-cohort consortium approach

The bi-national Cannabis Cohorts Research Consortium (CCRC) is an example of a research effort using integrated data analysis across multiple Australasian cohorts. The CCRC is a multi-organisational and multi-disciplinary international collaboration that brings together a number of the most mature longitudinal studies of child and adolescent development across Australia and New Zealand, including: the Australian Temperament Project (30 years/15 waves), the Christchurch Health and Development Study (37 years/23 waves), the Mater Hospital and University of Queensland Study of Pregnancy (33 years/8 waves), the Personality and Total Health Through Life Project (15 years/4 waves), the Victorian Adolescent Health Cohort Study (20 years/10 waves), and the Western Australian Pregnancy Cohort (Raine) Study (25 years/11 waves).

The CCRC was formed in 2006 and is co-ordinated by the National Drug and Alcohol Research Centre (NDARC), at UNSW Australia. The CCRC stemmed from a recognised need to better understand the link between cannabis use (particularly rarer patterns of high use) and other substance use, mental health problems and psychosocial outcomes. The CCRC has first focussed on cannabis use across adolescence and young adulthood, in particular the impact of regular patterns of use (i.e., daily and dependent use). Here, we describe an approach to developing well-powered samples for the study of infrequent exposures and outcomes.
What cohorts are involved in the CCRC?

The Consortium brings together researchers from some of the largest and longest running longitudinal studies of health and well-being internationally. Details of the cohorts presently are:

1. **The Australian Temperament Project (ATP)** (Prior et al., 2000; Vassallo and Sanson, 2013) is a longitudinal study of social and emotional development that commenced in 1983 as a sample of 2443 infants (aged 4-8 months) and their parents. Maternal and Child Health nurses at Infant Welfare Centres, which successfully contacted 94% of families with a new infant in the State of Victoria, Australia, assisted with enrolling families for the study. Nurses distributed approximately 3000 questionnaires, and 2443 (81%) usable questionnaires were returned. These 2443 formed the initial sample.

*Follow-up:* The ATP has been studied on a total of 15 occasions: in childhood (4 months, 1, 2, 3, 5, 7, 9, 11, 12 years), adolescence (13, 15, 17, 19 years) and young adulthood (24, 28 years).

2. **The Christchurch Health and Development Study (CHDS)** (Fergusson and Horwood, 2001; Fergusson et al., 1989) is a longitudinal study of a birth cohort of 1265 children born in the Christchurch, New Zealand, urban region in 1977. These children included 97% of all live births occurring during the recruitment period.

*Follow-up:* The cohort has now been studied on a total of 23 occasions (birth, 4 months, yearly between 1-16 years, and then at 18, 21, 25, 30 and 35 years).

3. **The Mater Hospital and University of Queensland Study of Pregnancy (MUSP)**

Between 1981 and 1983, 8556 consecutive pregnant women attending for their first clinic visit at average 18 weeks’ gestation at the Mater Misericordiae Mothers' Hospital in Brisbane, Australia, were invited to participate in this longitudinal birth
cohort (Keeping et al., 1989). Of those invited, 8458 (99%) agreed to participate in the study and 7223 gave birth to a live singleton child. The response proportion was 84%.

Follow-up: The cohort has now been studied on eight occasions (18 weeks gestation, birth, 4 years, 5 years, 14 years, 21 years, 27 years, 30 years).

4. The Personality and Total Health (PATH) Through Life Project (Anstey et al., 2012) is a longitudinal cohort study that commenced in 1999 with the recruitment of 7485 young (age 20–24 years at baseline, n=2404), midlife (age 40–44 years at baseline, n=2530) and older (age 60–64 years at baseline, n=2551) adults randomly sampled from the electoral roll of the Australian Capital Territory and the nearby city of Queanbeyan, Australia. To align with the focus of the Consortium, only the youngest PATH cohort was included in the CCRC. The response rate amongst those contacted who were in the target age range (20-24 years) and still resident in the area was 58.6%.

Follow-up: The cohort has been assessed on four occasions between 1999-2011 (age 20-24 years at baseline, 24-28 years, 28-32 years and 32-38 years).

5. The Victorian Adolescent Health Cohort Study (VAHCS) (Patton et al., 2007) is a longitudinal study of a representative sample of mid-secondary school adolescents in Victoria, Australia. In 1992, participants were recruited via schools at the end of Year 9 (wave 1, mean age 14.9 years) or the start of Year 10 (wave 2, mean age 15.5 years). Of the invited sample of 2032 students, 1943 (96%) were assessed at least once during the first six waves.

Follow-up: Participants were assessed at six-monthly intervals on four further occasions during adolescence: wave 3 (mean age 15.9 years) - wave 6 (mean age 17.4
years); with a further four follow-ups in young adulthood: wave 7 (mean age 20.7 years) - wave 10 (mean age 35 years).

6. The Western Australian Pregnancy Cohort (Raine) Study (Newnham et al., 1993; McNight et al., 2012) commenced in 1989. An estimated 3222 (Mountain, 2013) pregnant women at around 18 weeks gestation from King Edward Memorial Hospital in Perth, Western Australia, were invited to participate. Of these, 2900 pregnant women were recruited. These pregnancies resulted in 2868 children recruited into the cohort.

Follow-up: Information was collected from the mothers during pregnancy and parent/guardian at each follow up until the participants were 17 years of age. Participants were assessed at birth and at ages 1, 2, 3, 5, 8, 10, 14, 17, 20 and most recently 23 years of age.

How has the CCRC pooled data across cohorts?

There are two major approaches to data pooling. The first, commonly used method, combines study level results meta-analytically. This approach is relatively straightforward to apply when there is a clear research hypothesis, the exposure and outcome variables are clearly and consistently specified within and across studies, and the measure of effect size can be easily interpreted (Curran and Hussong, 2009). However, because meta-analyses are based on combining estimates at the study level, rather than at the individual participant level, investigation is limited by low sample size and power. That is, often the unit of analysis in meta-analysis is the study, when more power can be derived if the unit of analysis is the study participant.
The second approach to data pooling, pooling of participant-level data, is less commonly used because gaining access to study participant data is more complicated (ethically and practically) than combining available study-level results. It is, however, ideally suited to investigations of low frequency exposures and/or outcomes such as rarer behaviours (e.g., regular drug use) or genetic variants. The approach involves developing common scales of measurement across cohorts (harmonisation) and pooling participant-level harmonised data from multiple cohorts to create a single (large) integrated dataset of participants rather than studies. Data harmonisation and pooling allows models to be fitted directly to participant-level data and, therefore, provides a number of advantages over study-level meta-analysis. The major advantages of data harmonisation are that it (Curran and Hussong, 2009):

1. provides increased sample size which enables more powerful examination of less frequent exposures and outcomes;

2. becomes feasible to explicitly model sources of between-study heterogeneity at the level of the observed data;

3. improves the stability of model estimation and reduces the influence of outliers and study specific characteristics; and,

4. provides greater precision, particularly for tests of interaction and subgroup analyses.

The CCRC has now developed a harmonised data set with over 20 variables, across four cohorts (ATP, CHDS, MUSP, VAHCS) with more than 7000 respondents, with the purpose of further examining questions around frequent or dependent cannabis use that may not be answered by analyses of individual cohorts. Selected PATH data on cannabis and depression have also been harmonised within this data set (Horwood et al., 2012). We intend to harmonise relevant data from all six cohorts in the CCRC subject to continued funding.
support. Table 1 provides an overview of the key domains of interest for data harmonisation, the availability of these data across the four main cohorts harmonised to date (ATP, CHDS, MUSP, VAHCS), and where data are available, the measures used in each of these cohort studies.

(Table 1 here)

The annual attrition rate across the six cohorts was low; ranging from 0.5% to 3.9% per annum among the cohorts; the retained (contactable) sample rates at most recent follow-up range from 70% to 96%. Although all studies showed evidence of some selection bias attributable to sample attrition, analysis of the implications of such biases for study findings have shown these to be minimal (Silins et al., 2014). Table 2 provides descriptive data on demographic and cannabis use characteristics in the harmonised sample.

(Table 2 here)

**What has it found? Key findings and publications**

Findings from the analysis of the integrated data improved knowledge of the relationships between cannabis use, mental health, other substance use and social development in young people (Horwood et al., 2010; Horwood et al., 2012; Silins et al., 2014). For example, the age of initiation of cannabis use was found to be a significant factor in educational attainment, with early use (i.e. < 15 years) accounting for up to 17% of the failure rate (population attributable risk) in obtaining key educational milestones such as high school completion compared with those who had not used by age 18 (Horwood et al., 2010). In relation to

---

1 Excludes deceased participants and those who have permanently withdrawn.
mental health, increasing frequency of cannabis consumption was associated with increased symptoms of depression, with the association strongest among adolescents (Horwood et al., 2012). Evidence has also been found for the adverse sequelae of adolescent cannabis use across a broad range of young adult outcomes, including cannabis and other illicit drug dependence and attempted suicide (Silins et al., 2014). Recently obtained additional NHMRC funding for the existing CCRC datasets will enable extension of this harmonisation work to alcohol, including the antecedents and consequences of less prevalent patterns of youth alcohol use.

**What are the advantages and challenges?**

Through the Consortium’s experience we have identified a range of advantages and challenges of the integrated data analysis and of the collaborative process by which the dataset was formed. **Advantages of the CCRC analytic process include:** access to high-quality longitudinal studies that provide a rich pool of data across important developmental periods from childhood, through adolescence and into adulthood; increased sample size and statistical power to investigate rarer exposures and outcomes; the potential to compare model parameters to establish which effects are universal versus cohort specific; the efficient use of existing resources; the development of infrastructure and expertise in Australia and New Zealand in data harmonisation methodologies; the demonstrated feasibility of the approach and the potential applicability of the approach to a range of other mental health and substance use problems.

**Advantages to the multi-cohort collaborative approach include:** increased capacity to bring together expert knowledge from across a range of disciplines; the establishment of positive collaborative relations between study investigators and across the field; opportunities for
mentorship and training of younger researchers by senior academics, including co-authorship; and the opportunity for significant knowledge dissemination and translation, with the potential to have greater impact than any one study alone. For example, media reporting of Silins et al. (2014) included 631 media citations in the month post-publication (including online =373, radio =170, television=78, and print=10), across 29 countries (e.g., United States, United Kingdom, Australia, New Zealand, Asia, Middle East, South America, Eastern Europe, Western Europe).

Challenges: One of the principal challenges for the CCRC is the cross study heterogeneity in multi-cohort analyses. Common sources of between study heterogeneity include use of different measures, sampling variation, timing of historical events and study design characteristics (Curran and Hussong, 2009). The CCRC cohorts do, however, have a number of advantages that reduce the effect of heterogeneity. Specifically, all cohorts: (1) are population based samples; (2) were born about the same historical time and are culturally similar; (3) adopted broadly similar data collection strategies (i.e., repeated measures interview and questionnaire assessments); and, (4) have used broadly consistent measures of the primary outcomes and exposures (see Table 1).

Harmonisation of data from these contemporary population studies across Australasia, with due allowance for differences in study design and variability, augments our ability to generalise our findings to the region more realistically than is possible for any individual study (Curran and Hussong, 2009). The prevalence of cannabis use in Australia and New Zealand appears to be reasonably consistent with countries such as the United States, Canada, and Western Europe (Copeland and Swift, 2009; Murray et al., 2007) and generalisation of our findings to these high income countries would appear to be justified, allowing for the fact
that the extent of possible confounding by variation in legislation between jurisdictions is unknown.

There are also challenges related to developing and maintaining this collaborative work. These include limited funding available in Australia and New Zealand for data harmonisation; the large investment in time, resources and statistical expertise required to harmonise data; limits to the number of cohorts that can be harmonised; the challenges of bringing people together and maintaining momentum, especially as the harmonisation process can be slow; and the need for clear understanding and rules about ownership of ideas, data and work product. There are also ethical constraints around data sharing (e.g., limitations on the extent to which data are directly available to researchers conducting the harmonisation work).

Conclusion

A data harmonisation approach to combining participant level data from large cohort studies provides a unique opportunity to examine salient cannabis-related or other questions that are not possible to address in individual cohorts or using meta-analytic approaches. Although acknowledging challenges to the combined cohort approach, the ongoing CCRC work has demonstrated the capacity of the researchers to collaborate in data harmonisation and that combining participant level data from several similar and large Australasian cohorts is feasible. Results from the CCRC study will allow cross-validation of important cannabis-related research questions in Australasia and internationally. Recent funding through the NHMRC will enable extension of the harmonisation work to alcohol, including the antecedents and consequences of less prevalent patterns of youth alcohol use. The work of the CCRC will produce more comparable and robust findings about the linkages between
substance use, mental health and social development in young people. Importantly, it enables critical health related questions to be addressed which may not be adequately answered within individual cohorts.

Can I get hold of the data? Where can I find out more?

Information about the CCRC can be obtained through the National Drug and Alcohol Research Centre at the University of New South Wales (http://ndarc.med.unsw.edu.au/project/cannabis-cohort-research-consortium-ccrc). Specific enquires regarding the CCRC harmonised dataset are coordinated with Data Custodians by Dr Delyse Hutchinson (d.hutchinson@unsw.edu.au) and Dr Edmund Silins (e.silins@unsw.edu.au) at NDARC. In 2010 the Australian Research Alliance for Children and Youth (ARACY) recognised the collaborative work of the CCRC as a successful network in action. The report can be accessed online: http://www.aracy.org.au/publications-resources/area?command=record&id=51.

The Cannabis Cohorts Research Consortium

Includes the listed authors and: Steve Allsop, Carolyn Coffey, Jan Copeland, Wayne Hall, Trish Jacomb, Primrose Letcher, Kerri Little, Jake Najman, Ann Sanson, Rachel Skinner, Tim Slade, Diana Smart and Maree Teesson.

Acknowledgements

We would like to acknowledge the following who have contributed to discussions about the CCRC: Dave Allsop, Anton Clifford, Alan Clough, John Howard, Caitlin Hughes, Kylie Lee,
Eugene Mattes, Jenny Mountain, Paul Nelson, Meredith O’Connor, Andrew Percy, Marian Shanahan, Jenny Williams and Tim Windsor.

We would like to thank all individuals and families involved in the participating cohorts for their time and invaluable contribution to the study.

The ATP would like to acknowledge all who have contributed to the Australian Temperament Project, especially founding investigators Professor Ann Sanson, Diana Smart, and Professors Margot Prior and Frank Oberklaid. The ATP has received on-going financial support from many funding agencies and institutions over the years, in particular the National Health and Medical Research Council and the Australian Research Council grant.

The CHDS has been supported by funding from the Health Research Council of New Zealand, the National Child Health Research Foundation, the Canterbury Medical Research Foundation, the New Zealand Lottery Grants Board, the University of Otago, the Carney Centre for Pharmacogenomics, the James Hume Bequest Fund, the US National Institutes of Health [MH077874], and the National Institute on Drug Abuse [DA/MH11301].

The MUSP would like to thank the data collection team and Greg Shuttlewood, the data manager of the University of Queensland. The core study was funded by the National Health and Medical Research Council (NHMRC) of Australia but the views expressed in the paper are those of the authors and not necessarily those of any funding body.
We thank Kaarin Anstey, Anthony Jorm, Bryan Rodgers, Simon Easteal, Nicolas Cherbuin, Andrew McKinnon, Trish Jacomb, Karen Maxwell and the PATH interviewing team for their contribution to PATH.

The Raine Study gratefully acknowledges the contribution of the Raine Study Team for cohort management and co-ordination, and the National Health and Medical Research Council (NHMRC) of Australia for funding contributions to the study over the last 23 years. We acknowledge The University of Western Australia (UWA), Telethon Institute for Child Health Research, Raine Medical Research Foundation at UWA, UWA Faculty of Medicine, Dentistry and Health Sciences, Women’s and Infant’s Research Foundation and Curtin University for core management funding.

We thank Christina O’Loughlin, John Carlin and Helen Romaniuk for their contributions to the Victorian Adolescent Health Cohort.

Declaration of conflicting interests

The Authors declare that there is no conflict of interest

Funding

The development of the CCRC harmonised cohort was supported by the National Health and Medical Research Council (NHMRC) [APP1009381] with additional support from the Australian Research Alliance for Children and Youth (ARACY) Australian Research Council/NHMRC Research Network [Seed Funding 2009]. The National Drug and Alcohol Research Centre and the National Cannabis Prevention and Information Centre at UNSW Australia are supported by funding from the Australian Government.
Peter Butterworth is supported by ARC Future Fellowship [FT1301014440]. Louisa Degenhardt is supported by an NHMRC Principal Research Fellowship [APP1041742].

Delyse Hutchinson is supported by a Vice-Chancellor’s Postdoctoral Fellowship from the University of New South Wales. Richard P Mattick is supported by an NHMRC Principal Research Fellowship [APP1045318]. Craig Olsson is supported by an Australian Research Council Principal Research Fellowship [DP 1311459]. George Patton is supported by an NHMRC Senior Principal Research Fellowship [APP1019887].
TABLE 1: Summary of study measures and data availability among integrated cohort studies

<table>
<thead>
<tr>
<th>Measures of Cannabis Use</th>
<th>ATP</th>
<th>CHDS</th>
<th>MUSP</th>
<th>VAHCS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age of onset</td>
<td>✅</td>
<td>✅</td>
<td>✅</td>
<td>✅</td>
</tr>
<tr>
<td>Frequency of use</td>
<td>✅</td>
<td>✅</td>
<td>✅</td>
<td>✅</td>
</tr>
<tr>
<td>Problem use/abuse/dependence</td>
<td>✅</td>
<td>✅³</td>
<td>✅²</td>
<td>✅¹</td>
</tr>
</tbody>
</table>

**Other Key Measures**

**Depression/Suicide**

<table>
<thead>
<tr>
<th>Symptoms/major depression</th>
<th>✅³</th>
<th>✅⁴</th>
<th>✅⁵</th>
<th>✅⁶</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suicidal behaviours/self harm</td>
<td>✅</td>
<td>✅</td>
<td>✅</td>
<td>✅</td>
</tr>
</tbody>
</table>

**Other illicit drug use**

| Frequency of use | ✅ | ✅ | ✅ | ✅ |
| Problem use/abuse/dependence | ✅ | ✅⁷ | ✅ | ✅ |

**Tobacco**

| Age of onset | ✅ | ✅ | ✅ | ✅ |
| Frequency of use | ✅ | ✅ | ✅ | ✅ |
| Problem use/abuse/dependence | ✅ | ✅⁸ | ✅¹ | ✅ |

**Crime**

| Criminal offending | ✅ | ✅⁹ | ✅¹⁰ | ✅ |
| Official contacts | ✅ | ✅ | - | - |

**Income**

| Financial income | ✅ | ✅ | ✅ | - |
| Employment status | ✅ | ✅ | ✅ | ✅ |
| Welfare dependence | ✅ | ✅ | ✅ | ✅ |

**Conduct disorder/antisocial behaviour**

| Frequency of use | ✅¹¹ | ✅¹² | ✅¹³ | ✅¹⁴ |
| Age of onset | ✅ | ✅ | ✅ | ✅ |

---

**Note:** ATP: Australian Temperament Project; CHDS: Christchurch Health and Development Study; MUSP: Mater Hospital and University of Queensland Study of Pregnancy; VAHCS: Victorian Adolescent Health Cohort Study. Check marks indicate data availability, footnotes indicate which published measures were used (if any).

1 Composite International Diagnostic Interview (CIDI) (WHO, 1993)

2 CIDI-Auto (WHO, 1997)

3 Rutter Problem Behaviour Questionnaire (with depression items added) at age 11-12, 12-13 years (Rutter et al., 1970); Short Mood and Feelings Questionnaire at age 13-14, 15-16, 17-18 years (Angold et al., 1995); Depression Anxiety Stress Scales (DASS) at age 19-20, 23-24 years (Lovibond & Lovibond, 1995)

4 Based on Diagnostic Interview Schedule for Children (DISC) (Costello et al., 1982) and Diagnosis and Statistical Manual (DSM-III-R) (APA, 1987) criteria at ages 15, 16 years; CIDI (WHO, 1993) thereafter

5 CIDI-Auto (WHO, 1997) and Young Adult Self-Report (YASR) (Achenbach, 1997), used age 21

6 Clinical Interview Schedule- Revised (CIS-R) (Lewis & Pelosi, 1990), used wave 1-7; CIDI-Auto(ICD-10) wave 9

7 DSM used at each assessment; DSM-III-R (APA, 1997) used at 15, 16 years; CIDI/DSM-IV (APA, 1994) thereafter

8 DSM-III-R (APA, 1987) at age 15 and 16 years; DSM-IV (APA, 1994) thereafter

9 Self-report at ages 15-30 years; Self-reported Early Delinquency Inventory (SRDI) (Elliott & Huizinga, 1989) used thereafter

10 Assessed using 6 item scale adapted from Mak (1993) at age 21 years

11 Adapted from Rutter Problem Behaviour Questionnaire at age 11-12, 12-13 years (Rutter et al., 1970); Antisocial behaviour adapted from Elliott & Ageton (1980), used age 19-20, 23-24 years

12 Participant reports on the SRED (Moffitt & Silva, 1988) at age 15, 16 years were used to classify according to DSM-III-R (APA, 1987) symptom criteria for conduct disorder over the period 14-16 years

13 Externalising behaviours assessed with YASR (Achenbach, 1997) at age 21 years

14 Self-report of early delinquency at wave 1-6, antisocial behaviour assessed wave 9
TABLE 2: Selected demographic and cannabis use characteristics of the CCRC harmonised cohort (N=7253) in young adulthood1

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>N (%)</th>
<th>ATP</th>
<th>CHDS</th>
<th>MUSP</th>
<th>VAHCS</th>
<th>Harmonised cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>390</td>
<td>488</td>
<td>1780</td>
<td>696</td>
<td>3354</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(39.0)</td>
<td>(48.7)</td>
<td>(47.2)</td>
<td>(45.8)</td>
<td>(46.2)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>610</td>
<td>515</td>
<td>1950</td>
<td>824</td>
<td>3899</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(61.0)</td>
<td>(51.4)</td>
<td>(52.3)</td>
<td>(54.2)</td>
<td>(53.8)</td>
<td></td>
</tr>
<tr>
<td>High school completion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>1129</td>
<td>444</td>
<td>2953</td>
<td>1276</td>
<td>5802</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(90.0)</td>
<td>(42.2)</td>
<td>(79.1)</td>
<td>(84.1)</td>
<td>(76.8)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>122</td>
<td>609</td>
<td>779</td>
<td>242</td>
<td>1752</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(9.8 )</td>
<td>(57.8)</td>
<td>(20.9)</td>
<td>(15.9)</td>
<td>(23.2)</td>
<td></td>
</tr>
<tr>
<td>Degree attainment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>440</td>
<td>256</td>
<td>157</td>
<td>889</td>
<td>1742</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(44.3)</td>
<td>(25.5)</td>
<td>(4.2)</td>
<td>(58.5)</td>
<td>(23.9)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>554</td>
<td>747</td>
<td>3603</td>
<td>631</td>
<td>5535</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(55.7)</td>
<td>(74.5)</td>
<td>(95.8)</td>
<td>(41.5)</td>
<td>(76.1)</td>
<td></td>
</tr>
<tr>
<td>Marital status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married/de-facto</td>
<td>332</td>
<td>477</td>
<td>788</td>
<td>525</td>
<td>2122</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(34.2)</td>
<td>(47.6)</td>
<td>(21.2)</td>
<td>(31.9)</td>
<td>(28.9)</td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>526</td>
<td>2938</td>
<td>1122</td>
<td>5226</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(65.8)</td>
<td>(74.5)</td>
<td>(78.9)</td>
<td>(71.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>883</td>
<td>827</td>
<td>2850</td>
<td>1401</td>
<td>5961</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(88.3)</td>
<td>(82.5)</td>
<td>(76.6)</td>
<td>(86.3)</td>
<td>(81.1)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>117</td>
<td>176</td>
<td>871</td>
<td>223</td>
<td>1387</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(11.7)</td>
<td>(17.6)</td>
<td>(23.4)</td>
<td>(13.7)</td>
<td>(18.9)</td>
<td></td>
</tr>
<tr>
<td>Age of cannabis use onset</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;15 years</td>
<td>99</td>
<td>92</td>
<td>450</td>
<td>56</td>
<td>697</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(11.0)</td>
<td>(9.2)</td>
<td>(12.2)</td>
<td>(3.68)</td>
<td>(9.8)</td>
<td></td>
</tr>
<tr>
<td>15-17 years</td>
<td>274</td>
<td>310</td>
<td>1029</td>
<td>411</td>
<td>2024</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(30.3)</td>
<td>(30.9)</td>
<td>(28.0)</td>
<td>(27.0)</td>
<td>(28.5)</td>
<td></td>
</tr>
<tr>
<td>Never before age 18 years</td>
<td></td>
<td>530</td>
<td>600</td>
<td>2197</td>
<td>1053</td>
<td>4380</td>
</tr>
<tr>
<td></td>
<td>(58.7)</td>
<td>(59.9)</td>
<td>(59.8)</td>
<td>(69.3)</td>
<td>(61.7)</td>
<td></td>
</tr>
<tr>
<td>Frequency of cannabis use2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>376</td>
<td>203</td>
<td>1878</td>
<td>366</td>
<td>2823</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(38.7)</td>
<td>(21.2)</td>
<td>(50.4)</td>
<td>(27.2)</td>
<td>(40.3)</td>
<td></td>
</tr>
<tr>
<td>Less than monthly</td>
<td>463</td>
<td>444</td>
<td>454</td>
<td>48</td>
<td>774</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(47.2)</td>
<td>(56.7)</td>
<td>(25.9)</td>
<td>(55.5)</td>
<td>(38.8)</td>
<td></td>
</tr>
<tr>
<td>Monthly or more</td>
<td>79</td>
<td>76</td>
<td>419</td>
<td>45</td>
<td>619</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(8.1 )</td>
<td>(7.9)</td>
<td>(11.3)</td>
<td>(3.4)</td>
<td>(8.8)</td>
<td></td>
</tr>
<tr>
<td>Weekly or more</td>
<td>39</td>
<td>81</td>
<td>257</td>
<td>94</td>
<td>471</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(4.0 )</td>
<td>(8.5)</td>
<td>(6.9)</td>
<td>(7.0)</td>
<td>(6.7)</td>
<td></td>
</tr>
<tr>
<td>Daily</td>
<td>19</td>
<td>55</td>
<td>206</td>
<td>94</td>
<td>374</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.0 )</td>
<td>(5.7)</td>
<td>(5.5)</td>
<td>(7.0)</td>
<td>(5.3)</td>
<td></td>
</tr>
<tr>
<td>Cannabis dependence3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>43</td>
<td>38</td>
<td>288</td>
<td>102</td>
<td>471</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(4.4 )</td>
<td>(3.8)</td>
<td>(11.4)</td>
<td>(6.7)</td>
<td>(7.8)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>927</td>
<td>965</td>
<td>274</td>
<td>1416</td>
<td>3335</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(95.6)</td>
<td>(96.2)</td>
<td>(88.6)</td>
<td>(93.3)</td>
<td>(92.2)</td>
<td></td>
</tr>
</tbody>
</table>

Note: ATP: Australian Temperament Project; CHDS: Christchurch Health and Development Study; MUSP: Mater Hospital and University of Queensland Study of Pregnancy; VAHCS: Victorian Adolescent Health Cohort Study. 1Among: MUSP sample at age 21 years, ATP sample at age 24 years, VAHCS sample at age 24 years, CHDS sample at age 25 years. For information on the strategies used to harmonise variables across multiple cohorts see Horwood et al. (2010, 2012) and Silins et al. (2014). 2Reporting period: Past month in ATP, MUSP; past year in CHDS, VAHCS. 3Reporting period: Past 12 months in ATP, CHDS, VAHCS; past month in MUSP.
References


Keeping JD, Najman JM, Morrison J, et al. (1989) A prospective longitudinal study of social, psychological and obstetric factors in pregnancy: Response rates and demographic


Mountain J. (2013) personal communication.


