A 40-year study of child maltreatment over the early life course predicting psychiatric morbidity, using linked birth cohort and administrative health data: protocol for the Childhood Adversity and Lifetime Morbidity (CALM) study

Steve Kisely, Stuart Leske, Urska Arnautovska, Dan Siskind, Nicola Warren, Korinne Northwood, Shuichi Suetani and Jake Moses Najman

Background

Child maltreatment is a major public health issue worldwide. Retrospective studies show a strong association between self-reported child maltreatment and poor mental and physical health problems. Prospective studies that use reports to statutory agencies are less common, and comparisons of self- and agency-reported abuse in the same cohort even rarer.

Aims

This project will link state-wide administrative health data with prospective birth cohort data ($N = 7223$) from Brisbane in Queensland, Australia (including notifications to child protection agencies), to compare psychiatric outcomes in adulthood of agency- and self-reported child maltreatment while minimising attrition bias.

Method

We will compare people with all forms of self- and agency-reported child maltreatment to the rest of the cohort, adjusting for confounding in logistic, Cox or multiple regression models based on whether outcomes are categorical or continuous. Outcomes will be hospital admissions, emergency department presentations or community/out-patient contacts for ICD-10 psychiatric diagnoses, suicidal ideation and self-harm as recorded in the relevant administrative databases.

Conclusions

This study will track the life course outcomes of adults after having experienced child maltreatment, and so provide an evidence-based understanding of the long-term health and behavioural consequences of child maltreatment. It will also consider health outcomes that are particularly relevant for adolescents and young adults, especially in relation to prospective notifications to statutory agencies. Additionally, it will identify the overlap and differences in outcome for two different sources of child maltreatment identification in the same cohort.

Keywords

Childhood experience; epidemiology; anxiety disorders; depressive disorders; post-traumatic stress disorder.

Copyright and usage

Child maltreatment is a major public health issue worldwide, with serious and often debilitating long-term consequences for physical and mental health. Long-term adverse effects include anxiety and post-traumatic stress disorder (PTSD), depression, schizophrenia spectrum disorders, addiction disorders, conversion disorders, borderline personality disorder, eating disorders, somatoform disorders, obesity, HIV, chronic pain and chronic physical disease.[1–4] Understanding the distinctive long-term impact of differing types of maltreatment may help clinicians better connect early childhood adversity with current health-related morbidities, to both provide more holistic care and identify public health targets for primary prevention efforts. Prospective studies are important as retrospective self-report of child maltreatment may be subject to memory bias and omission.[5,6] To our knowledge, this is the first planned study to link administrative health data with prospective birth cohort data (including notifications to child protection agencies) to compare outcomes in adulthood of agency- and self-reported child maltreatment of all types in the same sample.

There are four different types of child maltreatment: physical abuse, sexual abuse, emotional abuse and neglect.[2] Information on child maltreatment comes from two main sources: retrospective self-reports and prospectively gathered notifications to statutory child protection agencies. The general focus has been on sexual or physical abuse rather than neglect, emotional abuse and the co-occurrence of different child maltreatment types.[7] Most existing prospective studies do not control for relevant confounding factors, with some controlling for none. Lack of data prevented the most recent meta-analysis of longitudinal cohort studies assessing childhood trauma and adult mental disorder from providing estimates for separate mental disorders and conducting adequate publication bias and sensitivity analyses, highlighting the need for more of these studies.[3] In addition, in a recent large-scale review of child maltreatment and depressive and anxiety disorders, 64% of studies were cross-sectional, in which self-reported child maltreatment could be either a cause or consequence of pathology, and 28.3% were longitudinal studies with retrospective reporting. Both are subject to recall bias. This left only eight longitudinal studies with prospective child maltreatment assessment. Few studies have controlled for factors that may predispose to both child maltreatment and later adverse outcomes.

In contrast to cross-sectional designs, birth cohort studies can provide robust evidence of causation if they fully adjust for confounding, given the ability to establish preceding risk factors and subsequent outcomes. The Mater-University of Queensland Study

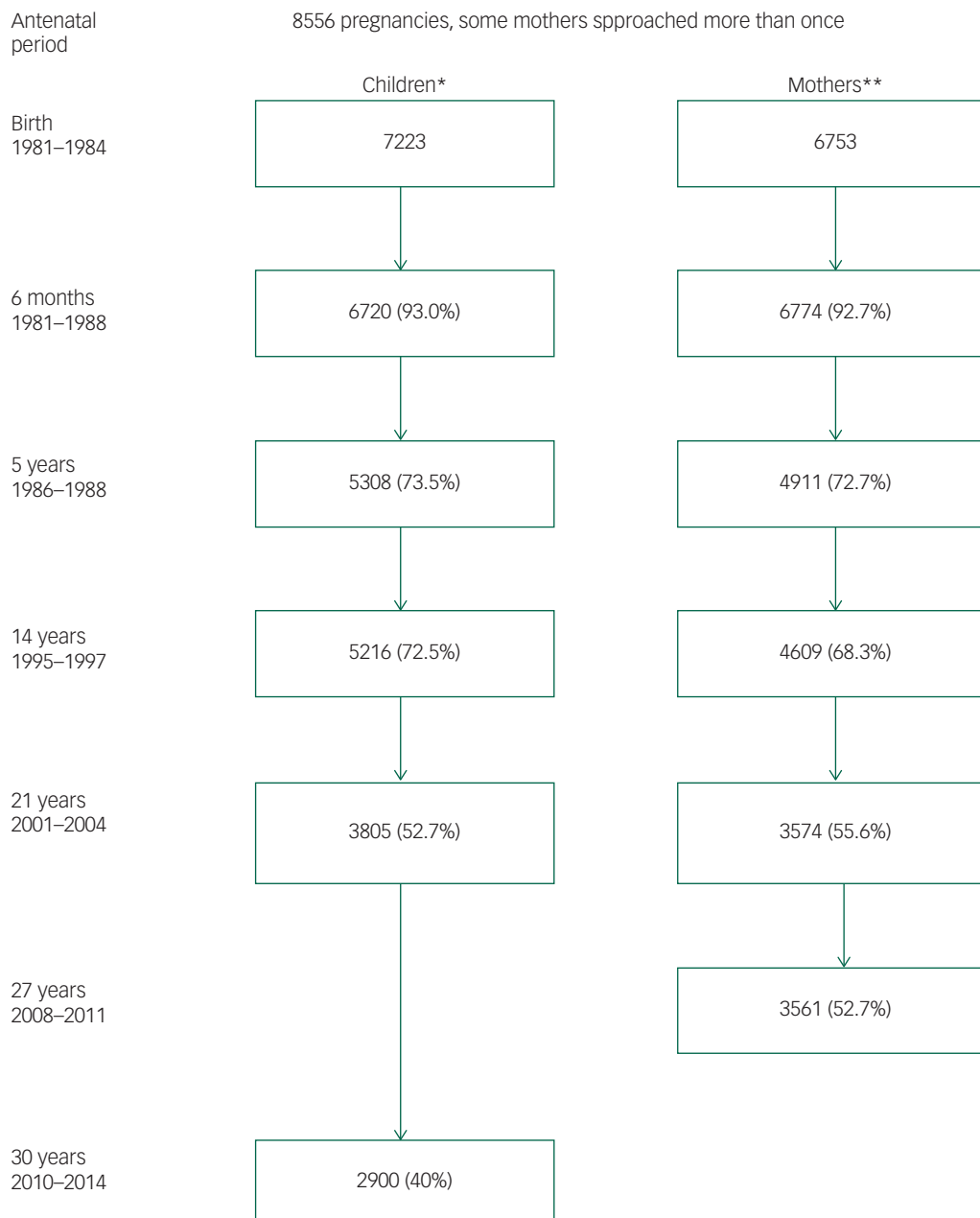| Antenatal period	8556 pregnancies, some mothers spproached more than once	
	Children*	Mothers**
Birth 1981–1984	7223	6753
6 months 1981–1988	6720 (93.0%)	6774 (92.7%)
5 years 1986–1988	5308 (73.5%)	4911 (72.7%)
14 years 1995–1997	5216 (72.5%)	4609 (68.3%)
21 years 2001–2004	3805 (52.7%)	3574 (55.6%)
27 years 2008–2011		3561 (52.7%)
30 years 2010–2014	2900 (40%)	

Fig. 1 Flowchart of the number (%) retained in the Mater-University of Queensland Study of Pregnancy at each phase (age) of data collection (reproduced with permission from reference 8). Permission to replicate this figure has been given to the authors by Oxford University Press. *Child singleton cohort includes 520 sets of sibling pairs. **Maternal cohort includes 50 mothers who had only multiple birth deliveries in the study.

of Pregnancy (MUSP) is a 40-year longitudinal birth cohort study from Brisbane, Australia, which has examined a variety of environmental risk factors for cognitive, psychological and behavioural health outcomes for over 7000 women and their children, going on to include the children of those children.[8]

Between 1981 and 1983, 8556 consecutive pregnant women attending the Mater Misericordiae Mothers' Hospital for their first prenatal visit were invited to participate. The final cohort numbered 7223 mother and infant pairs, which included only consenting participants who delivered live singleton infants at the study hospital (Fig. 1).[8] Unique features of the MUSP include the collection of a wide range of both maternal and child sociodemographic, health and behavioural variables, as well as prospectively substantiated child maltreatment through linkage to reports to child

protection agencies and self-reported child maltreatment in adulthood. Findings to date have demonstrated multiple links between early maternal/childhood adversity and poor outcomes. Specifically, agency-reported child maltreatment was associated with internalising/externalising symptoms in adolescence and anxiety, depression, psychosis or PTSD in adulthood, as well as substance use disorder, obesity, asthma, health anxiety and poor oral health.[9–14] The associations with self-reported child maltreatment and many of the psychosocial outcomes were even stronger.[10,15,16]

Of the 8556 consecutive women invited to participate in the study, 8458 accepted the invitation (98.9% response rate). The major limitation, however, has been the effect of attrition. Children were interviewed at the 14-, 21- and 30-year follow-ups. Of the original cohort of mothers who gave birth to a live singleton

baby (6753 mothers gave birth to 7223 children during the 3-year recruitment period), only 39.6% of children ($n = 2861$) participated in the 30-year follow-up (Fig. 1).[8] Importantly, loss to follow-up was associated with measures of socioeconomic disadvantage, raising the possibility of attrition bias.[17] Furthermore, only a limited number of health outcomes have been explored, and there has been no information on common morbidities found in adolescents and young adults, such as trauma and self-poisonings.

This protocol presents a novel approach whereby birth cohort data can be linked to administrative health data through a series of steps that preserve the anonymity of the original survey participants, thereby allowing the use of data where it would be impossible to obtain consent through loss to follow-up.

Aims and objectives

We aim to link administrative health data to the MUSP longitudinal, population-based birth cohort to compare health outcomes for the entire sample of child maltreatment (including the two reporting sources and different subtypes) with those of the rest of the cohort. Use of state-wide administrative health data will mean that it is possible to capture all health service contacts throughout the state in situations where cohort participants have moved from the South Metropolitan area.

Hypotheses

Several specific hypotheses will be tested in this study: (a) that children exposed to agency- or self- reported child maltreatment are more likely to experience several negative mental health outcomes, as measured by administrative health data, in terms of admissions, emergency department visits and community mental health service contacts, after adjusting for confounders and covariates at different levels of child ecology; (b) that different and multiple forms and/or recurrent incidents of child maltreatment independently predict later mental health outcomes; (c) that there will be statistically significant associations with subsequent health service use for both agency- and self- identified child maltreatment, although different types of maltreatment may lead to different patterns (odds ratios will range from small to large, depending on the prevalence of the outcome, the type of exposure and type of report (self or agency)); and (d) that an independent association between some types of child maltreatment and some mental disorders will remain after adjusting for all possible confounding factors, thereby indicating potential causation.

Method

Ethics approval and consent to participate

This project has been approved by the following ethics committees and data custodians, including a waiver to seek participant consent from The University of Queensland Human Research Ethics Committee (HREC) (approval number 2021/HE001925) and the Metro South Health HREC (approval number HREC/2022/QMS/83690). Access to the administrative health data-sets was approved by the Director-General of Queensland Health as Chief Executive under the Public Health Act 2005 (PHA 83690). The study is registered with the Australian and New Zealand Clinical Trials Registry (ACTRN12622000870752).

Study design

This is an observational, longitudinal birth cohort study linked to administrative health data. We will report findings per the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guideline (Supplementary Table 1)[18] and the Reporting of Studies Conducted using Observational Routinely-Collected Health Data (RECORD) Statement (Supplementary Table 2).[19]

Setting

The original study setting was a large hospital that accounted for around 50% of the births in Brisbane, the capital city of the state of Queensland, servicing both public and private patients.[20]

Data sources/measurement

This study will link data to the MUSP cohort (Fig. 1) from the following Queensland-wide databases, up until 1 January 2020 (Figs 2 and 3): (a) the Queensland Hospital Admitted Patients' Data Collection (QHAPDC) of all admissions for all diagnoses to all hospitals in the public (from 1 January 2000) and private sectors (from 1 July 2007), (b) the Emergency Data Collection (EDC) of all emergency department presentations state-wide (from 2008) and (c) the Consumer Integrated Mental Health Application (CIMHA) of all contacts with state-run community-based or out-patient mental health services in the State (from 2010).

All three databases have used diagnostic codes from the ICD-10[21] for the entire period of the proposed study. We will investigate the possibility of using a cross-walk to convert these codes to ICD-11 diagnoses.

Participants

As noted previously, the cohort consisted of 7223 mother–child pairs recruited from the Mater Misericordiae Mothers' Hospital, the main obstetrics unit for Brisbane. The recruitment started in January 1981, using a sampling frame involving all consecutive pregnant women attending a booking in clinic that usually occurred in the first trimester of pregnancy. The sampling frame excluded emergency transfers and admissions referred from other hospitals.[20] This study will focus on the children's health outcomes. Data were collected from mothers at their first clinic visit, and from mother–child pairs when children were 6 months and 5, 14 and 21 years of age (Fig. 3).[8]

Mothers' and children's data were collected separately at the 27-year (mothers) and 30-year follow-ups (children). All follow-ups of the MUSP have received ethical clearance. In September 2000, cases of child abuse and neglect that were reported to the Department of Families, Youth and Community Care (DFYCC) were accessed and confidentially linked to the longitudinal database following approval from the ethical review committees of both the Mater Misericordiae Children's Hospital and The University of Queensland.[22] The DFYCC (renamed in 2001) was Queensland's statutory child protection agency. Information included the date of referral, whether the referral was substantiated on subsequent investigation, and the type of maltreatment (physical, emotional, sexual and neglect) up until the child was 15 years of age. Confidentiality was preserved by using an identification number to link the two databases anonymously.

Data linkage

There are two groups of researchers (Professor Najman and colleagues, as well as the MUSP data custodian) who are based at the School of Public Health, and the other team members are based in the Medical School at the Princess Alexandra Hospital. The Princess Alexandra Hospital researchers who will do the analysis
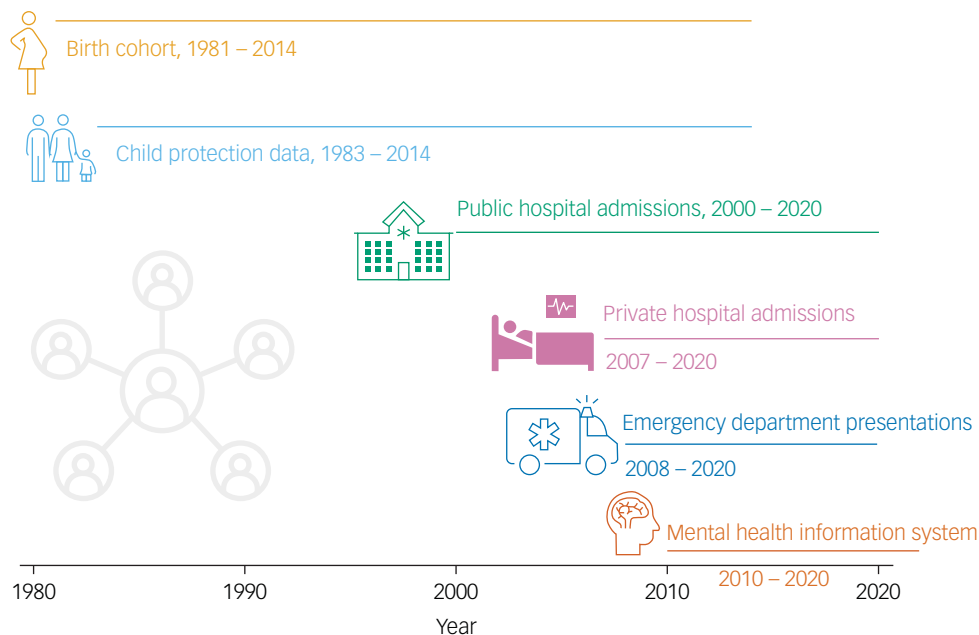
Fig. 2 Contributing data-sets.

will not have access to identifiable administrative health, MUSP or child protection data, and will only be able to view anonymised data extracts.

Confidentiality will be preserved by using an identification number to link the two person-level databases anonymously. The researchers analysing the maltreatment data will have no access to identifying information. Ethical approval for the anonymous database matching at the time was obtained from the ethical review committees of both the Mater Misericordiae Children's Hospital and The University of Queensland.
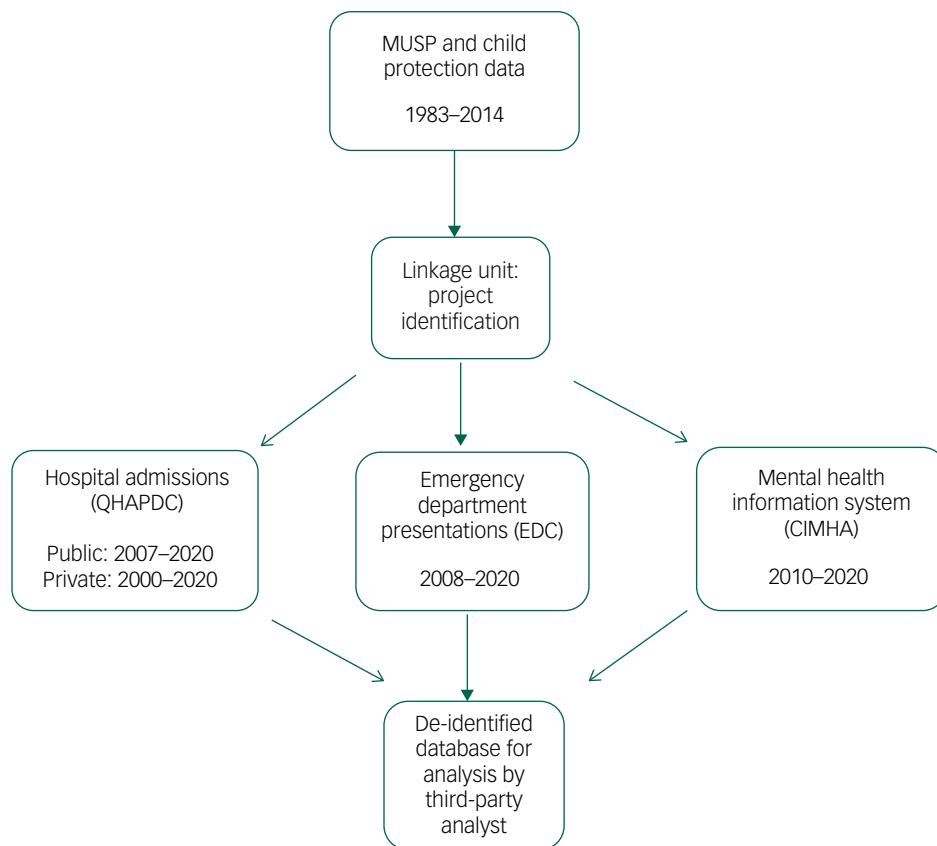

Fig. 3 Database linkages. CIMHA, Consumer Integrated Mental Health Application; EDC, Emergency Data Collection; MUSP, Mater-University of Queensland Study of Pregnancy; QHAPDC, Queensland Hospital Admitted Patients' Data Collection.

Data linkage will be done by the Statistical Services Branch (SSB) of Queensland Health. The MUSP data custodian will send a data-set to the SSB that is limited to the name, date of birth and gender of birth cohort participants (Fig. 3). This data-set will contain no other information so the SSB staff will only know that participants were born in Brisbane between 1981 and 1983. Within the SSB, these details will be matched to the Queensland Health master linkage file (MLF), using Oracle SQL Developer, version 22.2.1 for Windows (Oracle Corporation, Redwood Shores, CA, USA; www.oracle.com/database/sqldeveloper/technologies/download/). This involves a thorough review of all possible matches rather than using match thresholds, meaning the linkage is of very high accuracy. Every linkage process is cross-checked by the project linkage team manager. The MLF uses several custom ChoiceMaker (version 2.7 for Windows; ChoiceMaker LLC, Princeton, NJ, USA; www.choicemaker.com/Download) 'models' to build the linkage map. To determine linkage decisions for pairs of records, models use a series of weighted tests to calculate a match probability, apply decision thresholds (probability cut-offs) to assign a differ/review/match decision (all machine learning), and then in certain conditions apply rules to override the probability-assigned decision. To assure quality and efficiency, SSB's proprietary linkage models are developed for different types of data sources. Each model can comprise hundreds of tests, with test weights assigned through supervised machine learning, using robust training data with fine-tuned decision thresholds and deterministic rules. The error rates are approximately 99.5% for false positives and 98% for false negatives, with more detail about the processes provided online.[23]

Each person is assigned a unique project identification number that is returned to the MUSP data custodian, as well as the custodians for each administrative data-set. The data custodian attaches the relevant clinical or service information to the unique project identification number, but no identifying information. These anonymised data extracts are then directly released to a third-party analyst (in this case, the research team at the Princess Alexandra Hospital), but not returned to the MUSP data custodian. The third-party analyst uses the project IDs from the anonymised data to link the four data-sets. The third-party analyst then does the analysis, but only releases summary tables, not individual-level data.

Data and linkage quality

Administrative data are not collected for research purposes and therefore subject to recording bias. However, linking several databases from the same population permits cross-checking and validation of information. As noted previously, there are several quality assurance checks.[23] Although weights assigned by the linkage software process can largely determine true positive and true negative links, there is an area of uncertainty that requires human judgement. Paired sets of identifying information are therefore examined manually to inform whether they are to be considered to be true matches. Following this, identifying variables are cleaned and standardised, and quality assurance checks are performed for missing or duplicate data. Previous work has reported that demographic features such as Indigenous status were correctly identified in 89% of cases when checked with other sources.[24]

In keeping with standard practice, researchers will also perform basic frequency analyses following receipt of the data to include checks for duplicate records before and after merging data-sets, and whether the gender distribution, or years covered, seems appropriate.[23]

Variables

The following are predictor variables and possible covariates that will come from the MUSP data-set. We cross-checked these

against all the confounders included in a 2021 systematic review and meta-analysis of prospective longitudinal cohort studies in this area.[25]

(a) The number of notifications, if any, received by DFYCC in relation to each participant in the cohort before the collection in 2000, at which time the youngest children in the cohort were approximately 16.5 years of age.

(b) The number and proportion of these notifications that were substantiated by the DFYCC after investigations.

(c) The number and proportions of notifications and substantiations attributable to the recognised subtypes of abuse and neglect, i.e. physical abuse, sexual abuse, emotional abuse and neglect.

(d) Self-reported child maltreatment as measured by the Childhood Trauma Questionnaire Short-Form (CTQ)[26] at the 30-year follow-up. This has been used extensively and been found to have reliability, sensitivity and discriminant validity. There are five subscales with five items each: physical abuse, sexual abuse, emotional abuse, physical neglect and emotional neglect. Scores across the five items in each subscale are aggregated (producing an ordinal measure) and then recoded to four levels: none (or minimal), low (to moderate), moderate (to severe) and severe (to extreme). To ensure consistency with previous work on the same cohort, we will dichotomise the variables into none/low and moderate/severe, but also perform sensitivity analyses of the effect of splitting the scores into 'severe' versus all other categories.[16,27] This approach will enable comparisons between self- and agency-reported maltreatment in line with our earlier work on which the CTQ designer was a co-author.[27] Self-reported data on sexual abuse were also collected at 21-year follow-up with a different questionnaire based on five non-exclusive abuse scenarios.[15]

(e) Pregnancy-related characteristics of both the mothers of the original MUSP participants and when they, in turn, had children. These include the following: prenatal intention of pregnancy, feelings about becoming pregnant, number of previous pregnancies, number of previous pregnancies and live births, maternal dyad and relationship status during pregnancy, perceptions of pregnancy (easy/hard), problems with delivery, prenatal stressful life events, attitudes toward their newborn and social support during or after pregnancy. Pregnancy-related characteristics have previously been shown to be associated with a range of adverse outcomes in the MUSP cohort.[8]

(f) Sociodemographic variables on entry to the study, including year of birth, gender of the child, parental racial origin, maternal age, parental relationship status, family income at study entry (first prenatal visit), chronic socioeconomic disadvantage and parental education or employment.

(g) Sociodemographic variables up until the 30-year-old follow-up, including the offspring's employment status, income level at 30 years, educational achievement and marital status, maternal or paternal absence or both, and out of home placement. The administrative data will give employment and marital status at 40 years.

(h) Behavioural and environmental factors, including parental lifestyle, intimate partner conflict and violence, internalising or externalising behaviours, victim of bullying at 14 years, maternal mental health and family instability or dysfunction.

Outcomes from linked administrative data

These will be health service contacts as measured by QHAPDC (from 1 January 2000 for public hospitals and 1 July 2007 for

Table 1 ICD-10 codes

Broad diagnostic group	Detailed diagnostic categories	ICD-10 codes
Severe mental disorders	Schizophrenia, schizoaffective and other psychotic disorders	F20, F22, F23, F24, F25, F25.0, F25.1, F25.2, F25.8, F25.9, F28, F29
	Severe or psychotic affective disorders	F30, F31, F32.2, F32.3
	Psychotic disorders related to substance use	F10.5, F11.5, F12.5, F13.5, F14.5, F15.5, F15.50, F15.51, F15.59, F15.70, F16.5, F17.5, F18.5, F19.5, F19.7
Common mental disorders	Depressive and other mood disorders (e.g. recurrent depressive disorder, cyclothymia, dysthymia)	F32.0, F32.1, F32.8, F32.9, F33.0, F33.1, F33.4, F33.8, F33.9, F34, F38, F39
	Phobic anxiety disorders	F40, F40.1, F40.2, F40.8, F40.9
	Reaction to severe stress (e.g. acute stress reaction, post-traumatic stress disorder)	F43.0, F43.1, F43.8, F43.9
	Adjustment disorders	F43.2
	Other anxiety disorders (e.g. obsessive-compulsive, dissociative and somatoform disorders)	F41, F42, F44, F45, F48
Personality disorders	Cluster A	F21, F60.0, F60.1
	Cluster B	F60.2, F60.3, F60.30, F60.31, F60.4
	Cluster C	F60.5, F60.6, F60.7
	Other personality disorders	F60.8, F60.09, F60.9, F61, F62, F62.0, F62.1, F62.8, F62.9, F68.0, F68.1, F68.8, F69
Substance use disorders	Mental and behavioural disorders due to use of alcohol	F10
	Mental and behavioural disorders due to use of other substances	F11, F12, F13, F14, F15 F16, F17, F18, F19
Other adult-onset disorders	Organic disorders (e.g. dementia, disorders due to brain damage or dysfunction, amnesic syndrome)	F00, F01, F02, F03, F04, F05, F06, F07, F09
	Eating disorders	F50
	Self-harm and suicidal ideation	R45.8, X60-X84
	Other adult-onset disorders (e.g. sleep disorders, sexual dysfunction, postnatal and abuse of non-dependence-producing substances, habit and impulse disorders, gender identity disorders, sexual preference disorders; unspecified disorders)	F51, F52, F53, F54, F55, F59, F63, F64, F65, F66, F99
Other childhood-onset disorders	Mental retardation	F70, F71, F72, F73, F78, F79
	Disorders of psychological development (e.g. disorders of speech and language, pervasive developmental disorders)	F80, F81, F82, F83, F84, F88, F89
	Childhood behavioural (e.g. conduct and hyperkinetic disorders, mixed disorders of conduct and emotion)	F90, F91, F92
	Other childhood onset disorders (e.g. emotional disorders, disorders of social functioning, tic disorders)	F93, F94, F95, F98
Self-harm	Suicidal ideation and self-harm	R45.8, X60 to X84

private hospitals), the EDC (from 2008) and CIMHA (from 2010) up until 31 December 2019, in terms of numbers of visits and diagnoses as outlined by the relevant ICD-10 codes (Table 1). Participants will be classified as having a psychiatric disorder if they have ever had a hospital admission, emergency department presentation or community contact with a primary diagnosis of a psychiatric disorder (ICD-10 codes F00 to F99), suicidal ideation (code R45.8) or self-harm (codes X60 to X84). Psychiatric diagnoses will be divided into six broad categories, with each broad category further subdivided into more detailed subcategories (Table 1): severe mental disorders (non-affective psychoses and bipolar disorder), common mental disorders (anxiety and depressive disorders and post-traumatic stress disorder), personality disorders, substance use disorders (excludes substance-induced psychoses, which are included under severe mental disorders), other adult-onset disorders and other childhood onset disorders.

Diagnostic categories and subdivisions will be mutually exclusive sets of ICD-10 codes and coded as binary indicators (present/ not present) for each individual. We expect that drug, alcohol and mental health problems will be common reasons for presentation, given that the cohort will be approximately 40 years of age, based on pilot data from the decision support team at the Princess Alexandra Hospital.

Analyses

We will compare people with all forms of self- and agency-reported child maltreatment to the rest of the cohort, adjusting for potential confounders with a series of logistic, Cox and multiple regressions as follows: (a) logistic regression analyses to explore the association with a discrete outcome (e.g. admission for a particular psychiatric diagnosis), and control for relevant demographic and clinical covariates – this will be the primary outcome for each psychiatric diagnosis; (b) survival analysis to identify a terminal event (i.e. admission) and the time to event; (c) multiple regression analyses to allow predictions of continuous outcomes (e.g. number of presentations, bed days) and (d) where appropriate, we will apply marginal structural models, inverse probability weighting of marginal structural cox models and the g-computation formula for causal inference.[28,29]

Each form of childhood maltreatment will be sequentially entered into the models. We will run separate analyses for each administrative data-set (QHAPDC, CIMHA and EDC) and the psychiatric outcome of primary interest. Given our pilot data (see above), we will focus on the following diagnoses: self-harm, severe mental disorders (non-affective psychoses, bipolar disorder), common mental disorders (depression, anxiety or post-traumatic stress disorder), substance use (intoxication, hazardous/harmful use or dependence) and alcohol use (intoxication, hazardous/ harmful use or dependence). Many of these outcomes were investigated in our earlier analyses of just the MUSP cohort, so allowing comparison with our previous work. Other psychiatric disorders will be explored if there are sufficient numbers.

In the case of self-reported child maltreatment, we will undertake propensity score analyses that considers baseline covariates across the entire at-risk cohort rather than multiple imputation, as we cannot assume the data will be missing at random. In addition, previous work has multiple imputation of data from this cohort changes neither the estimates nor their precision, and that rate of loss to follow-up, for both dependent and independent variables, has little effect on estimates of association.[8]

Power calculations are based on dichotomous variables (our primary outcome), and used the WINdows Programs for EPIdemiologists (WINPEPI) program, version 11.65 for Windows (Joe Abramson, Jerusalem, Israel; www.brixtonhealth.com/

pepi4windows.html), for logistic regression, titled LOGISTIC, which allows for adjustment for covariates, with specified multiple correlation coefficients between covariates and the tested predictor.[30] Data from studies that were restricted to the MUSP cohort suggest minimum differences in rates of adverse outcomes between agency- or self-reported child maltreatment of between 10 and 15%, with odds ratios ranging from 1.34 to 5.53. Power calculations therefore indicate we would require a minimum of 6500 participants to have an 80% chance of detecting a significant difference at $P < 0.05$, meaning that our current sample is sufficient for commonly expected outcomes. We will handle missing data following 'Treatment And Reporting of Missing Data in Observational Studies' framework, and clean data following the guidance to detect, diagnose and edit data abnormalities.[31,32]

Feasibility

As outlined above, no new data will be needed. The MUSP data have already been collected and previously linked with the child protection data, whereas the administrative health data are available from Queensland Health. The QHAPDC, CIMHA and EDC data custodians have approved the project, and linkage and data cleaning are underway. In terms of notified maltreatment, 789 out of the 7223 participants (10.8%) had some type of notification by 16 years of age. In order of frequency, these were neglect ($n = 500$), emotional abuse ($n = 446$), physical abuse ($n = 475$) and sexual abuse ($n = 259$). At the 30-year follow-up, 600 out of the 2816 participants (24.7%) self-reported maltreatment of any type, with 326 (13.4%) rating this as severe. Neglect was the most frequent maltreatment ($n = 382$), followed by emotional ($n = 225$), sexual ($n = 198$) and physical abuse ($n = 197$).

Patient and public involvement

Patients and the public were not involved in the design of the study. However, as described in our dissemination activities, MUSP participants are regularly informed that information about studies arising from MUSP can be found on the MUSP website. This includes a full list of publications, as well as user-friendly descriptions of research. In addition, two of the authors are investigators in the National Research Translation Centre to Implement Mental Health Care at Scale in Primary Care and Community Settings (ALIVE) in Australia. This is a nationwide programme to co-design and evaluate approaches to improve mental health through evidence-based strategies for addressing physical, behavioural, psychological and other determinants, in partnership with people with lived experience. Early identification of risk factors and prevention of mental illness was identified in consultation with people with lived experience and their carers as a research priority, and this has informed the design of the present programme.

Plans for dissemination and publication of project outcomes

The immediate focus of the knowledge translation component of our work will be on those providing services to families with experiences of adversity. These service providers will primarily include nurses, social workers and other staff employed in hospitals or the Department of Child Safety of the Queensland Government. In addition, MUSP participants are also regularly informed of research arising out of the MUSP study (see above). Finally, findings will be submitted for presentation at academic conferences and publication in peer-reviewed journals.

Discussion

Retrospective studies show a strong association between self-reported child maltreatment and mental health problems.[2,3] Prospective studies using reports to statutory agencies are less common, and comparisons of self- and agency-reported abuse in the same cohort even rarer.[3] Where there have been comparisons of the two reporting sources, outcomes have been limited to single disorders.[33] In addition, even the most successful prospective studies have been subject to attrition. For instance, the follow-up rates at 30–40 years of age of two New Zealand birth cohorts were 76 and 89%, respectively.[34,35] Both relied on retrospective measures of child maltreatment.[34,35]

This project will link state-wide administrative health data with prospective birth cohort data (including notifications to child protection agencies), to compare a wide range of psychiatric and substance use outcomes in adulthood of both agency- and self-reported child maltreatment while minimising attrition bias. Comparing these two different sources of child maltreatment is important as each identifies different groups of individuals with comparatively little overlap between the two.[27,36] For instance, although subject to recall bias, self-reported data may be more comprehensive and representative than agency-reported, which may be restricted to the most severe cases. On the other hand, although prospective child maltreatment notifications are not subject to recall bias, they may focus on the most extreme or severe examples with supporting physical evidence. As a result, abuse and neglect may be underreported, as they may be less easily recognisable by others. By using both child maltreatment sources to cover the widest possible range of cases, our findings could inform both clinical care and child health policy, nationally and internationally.

The design is also innovative in that it illustrates a process whereby a birth cohort can be linked to administrative health data where it is impossible to obtain consent because of loss to follow-up, thereby minimising attrition bias. In the absence of participant consent, the MUSP data custodian will send a data-set to the SSB that is limited to the name, date of birth and gender of all birth cohort participants. This data-set will contain no other information so the SSB staff will only know that participants were born in Brisbane between 1981 and 1983. These are linked to the relevant administrative health databases to extract a de-identified subset to be supplied to the third-party analysts. Participant names are only required for the linkage of the database extracts, and are not given to the analysts who are from separate from both the SSB and the MUSP data custodian. The third-party analysts are the only people who will see the clinical data from the various linked data sources. However, by this stage, the data will only include a unique project identifier and no personal identifiers.

This study has several limitations. It relies on self-reported child maltreatment, using retrospective structured instruments, or prospective reports to statutory child protection agencies. Both sources have their weaknesses. In the case of the former, respondents may be reporting on events that occurred years, or even decades, previously. Unlike agency-reported child maltreatment, these reports are also subject to attrition bias in that they are only available for those who were successfully followed up into adulthood. As a result of these possible sources of recall or attrition bias, the child maltreatment cases may not be entirely representative.

The major limitation of agency notifications is that they may be a significant underestimation of the actual prevalence of child maltreatment with a possible bias toward certain populations, such as those experiencing sexual abuse.[25,37,38] Some researchers have argued that prospective and retrospective reports of child sexual abuse should be treated as distinct populations. For instance, a recent review found that on average, 52% of individuals with prospective reports of childhood maltreatment did not report it retrospectively, whereas 56% of individuals who retrospectively reported childhood maltreatment did not have any prospective reports.[36]

Furthermore, an even smaller number of reports are substantiated, thereby potentially representing a further underestimate of true prevalence. For instance, these cases may represent the most extreme or severe examples with supporting physical evidence. These underestimates may therefore mean that the study may be of insufficient power to detect statistical significance for some of the less common subtypes such as agency-reported sexual abuse, or rarer outcomes such as schizophrenia or bipolar affective disorder. In addition, the notified or substantiated cases may reflect the reporting practices of 15–20 years ago.

There may also be bias owing to follow-ups occurring too infrequently to capture changes in time-varying confounders or not all confounders being measured and available for inclusion in analyses.[39–41] These biases may limit the capacity for causal inferences.

Finally, the outcomes are based on health service contacts as recorded in administrative data. These may be subject to recording bias and only include people who have presented for treatment. The data do not cover people who have undiagnosed and untreated disorders, and those who are treated by general practitioners, or private psychiatrists and psychologists.

Despite these limitations, this study can highlight and compare the wide range of harms from both self- and agency-reported child maltreatment of all types in the same population sample while minimising attrition bias. A greater awareness of these outcomes may help clinicians to understand the importance of asking about early-life trauma and arrange referral, where appropriate, to relevant services. This may be particularly relevant to early-life neglect and emotional abuse as these child maltreatment types are often less obvious and more difficult to establish.[42] At a societal level, further recognition of these issues could inform interventions, including primary and secondary prevention, thereby changing the focus from immediate safety to a public health approach.

Steve Kisely (iD), Princess Alexandra Hospital Southside Clinical Unit, Greater Brisbane Clinical School, Medical School, The University of Queensland, Australia; Departments of Psychiatry, Community Health and Epidemiology, Dalhousie University, Canada; and Addiction and Mental Health Services, Metro South Addiction and Mental Health Service, Australia; **Stuart Leske**, Princess Alexandra Hospital Southside Clinical Unit, Greater Brisbane Clinical School, Medical School, The University of Queensland, Australia; **Urska Arnautovska**, Princess Alexandra Hospital Southside Clinical Unit, Greater Brisbane Clinical School, Medical School, The University of Queensland, Australia; and Addiction and Mental Health Services, Metro South Addiction and Mental Health Service, Australia; **Dan Siskind** (iD), Princess Alexandra Hospital Southside Clinical Unit, Greater Brisbane Clinical School, Medical School, The University of Queensland, Australia; and Addiction and Mental Health Services, Metro South Addiction and Mental Health Service, Australia; **Nicola Warren**, Princess Alexandra Hospital Southside Clinical Unit, Greater Brisbane Clinical School, Medical School, The University of Queensland, Australia; and Addiction and Mental Health Services, Metro South Addiction and Mental Health Service, Australia; **Korinne Northwood**, Princess Alexandra Hospital Southside Clinical Unit, Greater Brisbane Clinical School, Medical School, The University of Queensland, Australia; and Addiction and Mental Health Services, Metro South Addiction and Mental Health Service, Australia; **Shuichi Suetani**, Institute for Urban Indigenous Health, Australia; **Jake Moses Najman**, School of Public Health, The University of Queensland, Australia; and School of Social Sciences, The University of Queensland, Australia

Correspondence: Steve Kisely. Email: s.kisely@uq.edu.au

First received 28 Sep 2022, final revision 2 Jan 2023, accepted 4 Feb 2023

Supplementary material

Data availability

Data availability is not applicable to this article as no new data were created or analysed in this study.

Author contribution

Funding

Declaration of interest

None.

References

1 Jaffee SR. Child maltreatment and risk for psychopathology in childhood and adulthood. *Annu Rev Clin Psychol* 2017; **13**(1): 525–51.

2 Hailes HP, Yu R, Danese A, Fazel S. Long-term outcomes of childhood sexual abuse: an umbrella review. *Lancet Psychiatry* 2019; **6**(10): 830–9.

3 Gardner MJ, Thomas HJ, Erskine HE. The association between five forms of child maltreatment and depressive and anxiety disorders: a systematic review and meta-analysis. *Child Abuse Negl* 2019; **96**: 104082.

4 Dragioti E, Radua J, Solmi M, Arango C, Oliver D, Cortese S, et al. Global population attributable fraction of potentially modifiable risk factors for mental disorders: a meta-umbrella systematic review. *Mol Psychiatry* 2022; **27**(8): 3510–9.

5 Colman I, Kingsbury M, Garad Y, Zeng Y, Naicker K, Patten S, et al. Consistency in adult reporting of adverse childhood experiences. *Psychol Med* 2016; **46**(3): 543–9.

6 Widom CS, Morris S. Accuracy of adult recollections of childhood victimization: part 2, childhood sexual abuse. *Psychol Assess* 1997; **9**(1): 34–46.

7 Stoltenborgh M, Bakermans-Kranenburg MJ, Alink LR, van IJzendoorn MH. The prevalence of child maltreatment across the globe: review of a series of meta-analyses. *Child Abuse Review* 2015; **24**(1): 37–50.

8 Najman JM, Alati R, Bor W, Clavarino A, Mamun A, McGrath JJ, et al. Cohort profile update: the Mater-University of Queensland Study of Pregnancy (MUSP). *Int J Epidemiol* 2015; **44**(1): 78–78f.

9 Strathearn L, Giannotti M, Mills R, Kisely S, Najman J, Abajobir A. Long-term cognitive, psychological and health outcomes associated with child abuse and neglect. *Pediatrics* 2020; **146**(4): e20200438.

10 Kisely S, Strathearn L, Najman JM. Self-reported and agency-notified child abuse as a contributor to health anxiety in a population-based birth cohort study at 30-year follow-up. *J Acad Consult Liaison Psychiatry* 2022; **63**(5): 445–53.

11 Kisely S, Strathearn L, Najman JM. A comparison of oral health outcomes of self-reported and agency-notified child maltreatment in a population-based birth cohort at 30-year-old follow-up. *Psychosom Med* 2022; **84**(2): 179–87.

12 Abajobir AA, Kisely S, Scott JG, Williams G, Clavarino A, Strathearn L, et al. Childhood maltreatment and young adulthood hallucinations, delusional experiences, and psychosis: a longitudinal study. *Schizophr Bull* 2017; **43**(5): 1045–55.

13 Kisely S, Strathearn L, Najman J. The influence of child maltreatment on substance or alcohol use in 30-year-old adults: a birth cohort study. *Drug Alcohol Rev* 2021; **40**(4): 673–80.

14 Kisely S, Strathearn L, Najman JM. A comparison of psychosis-like symptoms following self-reported and agency-notified child abuse in a population-based birth cohort at 30-year-follow-up. *Schizophr Res* 2022; **239**: 116–22.

15 Mills R, Kisely S, Alati R, Strathearn L, Najman J. Self-reported and agency-notified child sexual abuse in a population-based birth cohort. *J Psychiatr Res* 2016; **74**: 87–93.

16 Kisely S, Strathearn L, Mills R, Najman JM. A comparison of the psychological outcomes of self-reported and agency-notified child abuse in a population-based birth cohort at 30-year-follow-up. *J Affect Disord* 2021; **280**(Pt A): 167–72.

17 Kisely S, Strathearn L, Najman JM. Child maltreatment and mental health problems in 30-year-old adults: a birth cohort study. *J Psychiatr Res* 2020; **129**: 111–7.

18 von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol* 2008; **61**(4): 344–9.

19 Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, et al. The Reporting of Studies Conducted using Observational Routinely-Collected Health Data (RECORD) statement. *PLoS Med* 2015; **12**(10): e1001885.

20 Najman JM, Bor W, O'Callaghan M, Williams GM, Aird R, Shuttlewood G. Cohort profile: the Mater-University of Queensland Study of Pregnancy (MUSP). *Int J Epidemiol* 2005; **34**(5): 992–7.

21 World Health Organization. *International Classification of Diseases* (10th edn). WHO, 1994.

22 Strathearn L, Mamun AA, Najman JM, O'Callaghan MJ. Does breastfeeding protect against substantiated child abuse and neglect? A 15-year cohort study. *Pediatrics* 2009; **123**(2): 483–93.

23 Queensland Government. *Queensland Data Linkage Framework*. State of Queensland (Queensland Health), 2020 (https://www.health.qld.gov.au/__data/assets/pdf_file/0030/150798/qlddatalinkframework.pdf).

24 Kisely S, Pais J. Can administrative data provide insights into the mental health of Indigenous Queenslanders? *Australas Psychiatry* 2011; **19**(suppl 1): S12–6.

25 McKay MT, Cannon M, Chambers D, Conroy RM, Coughlan H, Dodd P, et al. Childhood trauma and adult mental disorder: a systematic review and meta-analysis of longitudinal cohort studies. *Acta Psychiatr Scand* 2021; **143**(3): 189–205.

26 Bernstein DP, Stein JA, Newcomb MD, Walker E, Pogge D, Ahluvalia T, et al. Development and validation of a brief screening version of the Childhood Trauma Questionnaire. *Child Abuse Negl* 2003; **27**(2): 169–90.

27 Najman JM, Kisely S, Scott JG, Strathearn L, Clavarino A, Williams GM, et al. Agency notification and retrospective self-reports of childhood maltreatment in a 30-year cohort: estimating population prevalence from different data sources. *Child Abuse Negl* 2020; **109**: 104744.

28 Hernán MÁ, Brumback B, Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology* 2000; **11**(5): 561–70.

29 Robins J. A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. *Math Model* 1986; **7**(9): 1393–512.

30 Abramson JH. WINPEPI updated: computer programs for epidemiologists, and their teaching potential. *Epidemiol Perspect Innov* 2011; **8**(1): 1.

31 Lee KJ, Tilling KM, Cornish RP, Little RJA, Bell ML, Goetghebeur E, et al. Framework for the treatment and reporting of missing data in observational studies: the treatment and reporting of missing data in observational studies framework. *J Clin Epidemiol* 2021; **134**: 79–88.

32 Van den Broeck J, Argeseanu Cunningham S, Eeckels R, Herbst K. Data cleaning: detecting, diagnosing, and editing data abnormalities. *PLoS Med* 2005; **2**(10): e267.

33 Scott KM, McLaughlin KA, Smith DA, Ellis PM. Childhood maltreatment and DSM-IV adult mental disorders: comparison of prospective and retrospective findings. *Br J Psychiatry* 2012; **200**(6): 469–75.

34 Bell CJ, Foulds JA, Horwood LJ, Mulder RT, Boden JM. Childhood abuse and psychotic experiences in adulthood: findings from a 35-year longitudinal study. *Br J Psychiatry* 2019; **214**(3): 153–8.

35 Breslau N, Koenen KC, Luo Z, Agnew-Blais J, Swanson S, Houts RM, et al. Childhood maltreatment, juvenile disorders and adult post-traumatic stress disorder: a prospective investigation. *Psychol Med* 2014; **44**(9): 1937–45.

36 Baldwin JR, Reuben A, Newbury JB, Danese A. Agreement between prospective and retrospective measures of childhood maltreatment: a systematic review and meta-analysis. *JAMA Psychiatry* 2019; **76**(6): 584–93.

37 Gilbert R, Widom CS, Browne K, Fergusson D, Webb E, Jansen S. Burden and consequences of child maltreatment in high-income countries. *Lancet* 2009; **373**(9657): 68–81.

38 Newbury JB, Arseneault L, Moffitt TE, Caspi A, Danese A, Baldwin JR, et al. Measuring childhood maltreatment to predict early-adult psychopathology: comparison of prospective informant-reports and retrospective self-reports. *J Psychiatr Res* 2018; **96**: 57–64.

39 Penning de Vries BBL, Groenwold RHH. Bias of time-varying exposure effects due to time-varying covariate measurement strategies. *Pharmacoepidemiol Drug Saf* 2022; **31**(1): 22–7.

40 Young JG, Vatsa R, Murray EJ, Hernán MA. Interval-cohort designs and bias in the estimation of per-protocol effects: a simulation study. *Trials* 2019; **20**(1): 552.

41 Chiu Y-H, Rifas-Shiman SL, Kleinman K, Oken E, Young JG. Effects of intergenerational exposure interventions on adolescent outcomes: an application of inverse probability weighting to longitudinal pre-birth cohort data. *Paediatr Perinat Epidemiol* 2020; **34**(3): 366–75.

42 Dubowitz H, ed. Neglect of children's health care. In *Neglected Children: Research, Practice, and Policy*: 109–31. SAGE Publications, 1999.