

ORIGINAL CONTRIBUTION

Open Access



Quantifying sources of bias in longitudinal data linkage studies of child abuse and neglect: measuring impact of outcome specification, linkage error, and partial cohort follow-up

Jared W. Parrish^{1,5*}, Meghan E. Shanahan^{2,5}, Patricia G. Schnitzer³, Paul Lanier⁴, Julie L. Daniels² and Stephen W. Marshall^{2,5}

Abstract

Background: Health informatics projects combining statewide birth populations with child welfare records have emerged as a valuable approach to conducting longitudinal research of child maltreatment. The potential bias resulting from linkage misspecification, partial cohort follow-up, and outcome misclassification in these studies has been largely unexplored. This study integrated epidemiological survey and novel administrative data sources to establish the Alaska Longitudinal Child Abuse and Neglect Linkage (ALCANLink) project. Using these data we evaluated and quantified the impact of non-linkage misspecification and single source maltreatment ascertainment use on reported maltreatment risk and effect estimates.

Methods: The ALCANLink project integrates the 2009–2011 Alaska Pregnancy Risk Assessment Monitoring System (PRAMS) sample with multiple administrative databases through 2014, including one novel administrative source to track out-of-state emigration. For this project we limited our analysis to the 2009 PRAMS sample. We report on the impact of linkage quality, cohort follow-up, and multisource outcome ascertainment on the incidence proportion of reported maltreatment before age 6 and hazard ratios of selected characteristics that are often available in birth cohort linkage studies of maltreatment.

Results: Failure to account for out-of-state emigration biased the incidence proportion by 12% (from 28.3%_w to 25.2%_w), and the hazard ratio (HR) by as much as 33% for some risk factors. Overly restrictive linkage parameters biased the incidence proportion downwards by 43% and the HR by as much as 27% for some factors. Multi-source linkages, on the other hand, were of little benefit for improving reported maltreatment ascertainment.

(Continued on next page)

* Correspondence: jared.parrish@alaska.gov

¹Alaska Division of Public Health, Section of Women's, Children's, and Family Health, 3601 C St., Suite 358, Anchorage, AK 99503, USA

⁵The University of North Carolina at Chapel Hill, Injury Prevention Research Center, 137 East Franklin St, CB# 7505, Chapel Hill, NC 27599, USA

Full list of author information is available at the end of the article

(Continued from previous page)

Conclusion: Using the ALCANLink data which included a novel administrative data source, we were able to observe and quantify bias to both the incidence proportion and HR in a birth cohort linkage study of reported child maltreatment. Failure to account for out-of-state emigration and low-quality linkage methods may induce bias in longitudinal data linkage studies of child maltreatment which other researchers should be aware of. In this study multi-agency linkage did not lead to substantial increased detection of reported maltreatment. The ALCANLink methodology may be a practical approach for other states interested in developing longitudinal birth cohort linkage studies of maltreatment that requires limited resources to implement, provides comprehensive data elements, and can facilitate comparability between studies.

Keywords: Child maltreatment, Record linkage, Birth cohort, Longitudinal study, Bias, Health informatics, PRAMS

Background

Child maltreatment, which includes all forms of abuse, neglect, and mental injury of a child by a parent or other caregiver, is under-studied relative to its public health significance, impact on children, and contribution to adult health outcomes (Butchart et al. 2006; Leeb et al. 2008). Given the complex etiologies contributing to maltreatment, it is important to focus and evaluate prevention efforts using analytic models that utilize population representative longitudinal data sources (Cicchetti and Carlson 1989; Cicchetti 1994). Current nationally available data on maltreatment such as those collected and reported by the National Child Abuse and Neglect Data System (NCANDS) and National Incidence Study (NIS) do not allow for longitudinal assessment but provide annual snapshots (US Department of Health and Human Services et al. 2016; Sedlak et al. 2010a). Studying maltreatment, especially at the population level and over time is challenging but necessary to quantify risk (Rothman et al. 2008). Due to the many conceptual and logistical challenges of conducting population level longitudinal maltreatment research, traditional prospective cohort studies are often limited to subset populations known to child welfare (e.g. The Longitudinal Studies on Child Abuse and Neglect) (Runyan et al. 1998; Bertolli et al. 1995; Brownell and Jutte 2013). Large population representative longitudinal cohort studies are expensive, time-consuming, and require extensive administrative support to conduct participant follow-up (Rothman et al. 2008). Due to these and other challenges, alternative methods for generating population-representative longitudinal studies to examine child maltreatment are necessary.

Accordingly, linkage projects combining statewide birth records with child protective services (CPS) records have emerged as a health informatics approach (Putnam-Hornstein et al. 2013a; Jonson-Reid and Drake 2008; Wu et al. 2004; Jutte et al. 2011; Stanley et al. 1994). Birth cohort linkages in Australia, New Zealand and Canada have demonstrated the benefit of studying maltreatment through administrative record linkages (Brownell and Jutte 2013; Holman et al. 1999; Tonmyr et al. 2014). In the US, linkage studies of full birth

cohorts in California, Florida, Texas, and Alaska have highlighted the promise of this approach to study many child health outcomes and measure the incidence proportion of maltreatment over time (Wu et al. 2004; Van Horne et al. 2015; Putnam-Hornstein and Needell 2011; Gessner et al. 2004).

The use of entire statewide birth cohorts typically results in good statistical precision; these results however, are still subject to systematic error (Bertolli et al. 1995). Bias may result from a number of factors (Bohensky 2016), including 1) the influence of unknown selection factors with registration on administrative databases (e.g. institutional racism that can lead to biased reporting, or regional variation in applying screening policies), 2) pragmatic difficulties associated with accurately tracking all subjects over time using routine administrative databases (e.g. no access to or source of annual or regularly updated population level administrative data) resulting in unmeasured loss-to-follow up, 3) incomplete covariate adjustment for predictive and etiologic assessments due to limitations in availability and scope of data elements (e.g. birth certificates provide limited prenatal, social, and behavioral information and can limit etiologic and predictive modeling), 4) reliance on official reports of maltreatment to capture the outcome (Official reports to child welfare agencies are known to under-represent the magnitude of the problem due to under-reporting) (Sedlak et al. 2010b; Ewigman et al. 1993; Drake and Zuravin 1998), and 5) linkage misspecifications (e.g. using restrictive linkage assumptions when integrating data or having limited capacity for manual review of partial matches, no access to name change records, and large subpopulations with differential linkage patterns due to name homogeneity). Limited research to date has assessed the influence of these sources of bias on population based child maltreatment data linkage studies (Bohensky 2016; Greene et al. 2011).

We recently piloted a novel data linkage approach based on the methodology suggested by Bertolli et al. nearly 2 decades ago for studying child maltreatment (Bertolli et al. 1995). Our pilot project integrated the Pregnancy Risk Assessment Monitoring System

(PRAMS) sample in Alaska with CPS reports occurring after birth (Parrish et al. 2011) and demonstrated comparable associations to those published in the literature using full birth cohorts (Wu et al. 2004; Putnam-Hornstein and Needell 2011).

This paper expands greatly upon the initial pilot studies and describes the creation of the Alaska Longitudinal Child Abuse and Neglect Linkage (ALCANLink) project that integrates epidemiologic survey and multi-sector administrative data (Calderwood and Lessof 2009) to create a comprehensive longitudinal birth cohort study. We highlight the benefit of the ALCANLink methodology by documenting the bias in incidence and hazard ratios that can arise in birth cohort linkage studies due to incomplete data linkages, nonlinkage assumptions, and single source outcome ascertainment.

Methods

The ALCANLink project integrates the 2009–2011 PRAMS respondent births (hereafter referred to as

PRAMS births) with a core set of sources to follow the PRAMS cohort prospectively, which include vital records, child death review, and Alaska Permanent Fund Dividend (PFD) records (Alaska Department of Revenue: Permanent Fund Dividend Division 2016). Additional administrative sources and a three-year follow up survey to PRAMS capture additional factors. We limited our in-depth assessment described in this paper to a single PRAMS year (2009 births) to minimize the manual review and classification processes required, and allow for easier presentation of cohort details (Fig. 1).

Cohort establishment

Alaska PRAMS uses a representative stratified systematic sample of annual resident live births. It oversamples Alaska Native mothers and low birth weight (<2500 g) infants for reasons of statistical precision. Oversampling and nonresponse are reflected in post-stratification sample weights. Alaska PRAMS samples approximately 1 in every 6 live births occurring to resident mothers, and

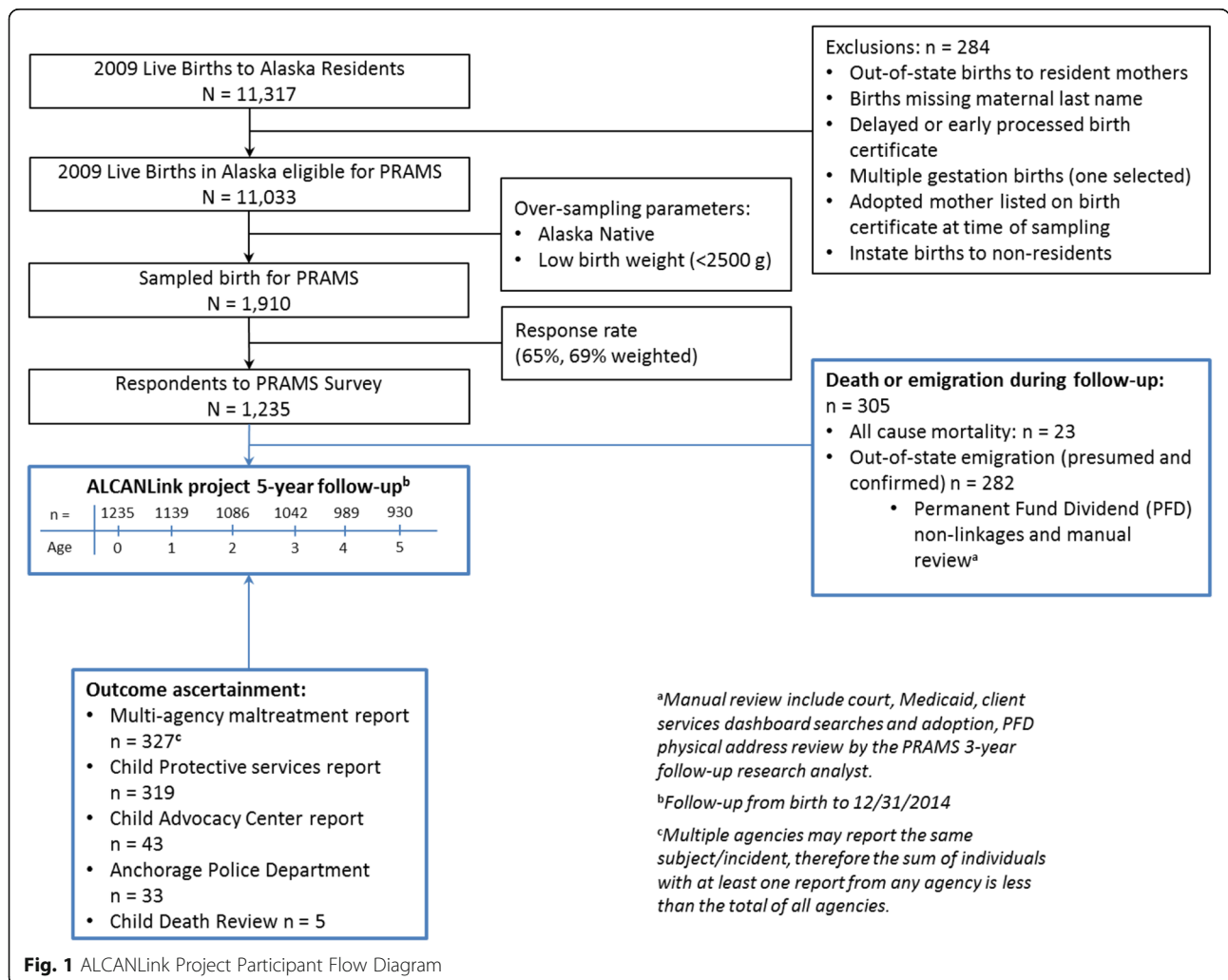


Fig. 1 ALCANLink Project Participant Flow Diagram

maintains about a 65% weighted response rate. Complete PRAMS methodology is described in detail elsewhere (Shulman et al. 2006).

In 2009, 11,317 live births occurred among Alaska resident mothers, with 11,033 meeting PRAMS inclusion criteria. Alaska PRAMS attempted to survey 1910 (17.3%) of these eligible mothers of newborns, with 1235 (64.7%) responding to the survey (69% weighted response rate).

We used the PRAMS cohort as the basis of ALCAN-Link opposed to the full birth cohort for the following reasons: a) PRAMS provides population-representative exposure measures that are extensively more comprehensive and detailed than those available only on birth records, b) PRAMS respondents provide consent to have their responses linked to other information the department has about them facilitating data linkages with multiple administrative databases, c) PRAMS is conducted in nearly all other states potentially allowing for standardization and expansion of these methods, and d) The complex sampling enables population estimation while reducing the resources required for exhaustive data linkages which can ultimately make additional administrative data linkages unfeasible.

Cohort follow-up

The 2009 PRAMS respondent births ($n = 1235$) were followed prospectively by linking PRAMS birth children to death certificates, Alaska Child Death Review (CDR) program records, and the Alaska Permanent Fund Dividend database with the most recent complete and available year (2014) at the time of this analysis. We used these follow-up sources to censor subjects for competing cause mortality and out-of-state emigration from Alaska.

In order to identify maltreatment-related child fatalities, we cross-checked all identified fatalities with the Alaska CDR program (see outcome ascertainment section below). All PRAMS births were subsequently linked to the annual PFD database. Adopted by constitutional amendment in 1976, Alaska established the Permanent Fund to invest a portion of the revenue earnings generated from petroleum production (Goldsmith 2002). The dividend is available, upon application, to all legal

Alaskan residents with strict eligibility requirements. Infants born on or before December 31st of a qualifying year are eligible for a PFD. Since 2009, an average of 92.2% of the state population has applied for, and 86.0% approved for a dividend annually (Alaska Department of Revenue: Permanent Fund Dividend Division 2016). The PFD essentially serves as an annual census and therefore provides a unique source for conducting historical cohort studies using the Alaska population. We know of no other comparable epidemiologic resource for all residents within any other US state, which enables us to explicitly quantify the potential bias associated with linkage misspecification in longitudinal birth cohort linkage studies. PRAMS respondent children that failed to link with either death records or PFD records underwent an extensive manual review using multiple administrative state sources which included, child and parental searches in PFD, court, Medicaid and WIC records, and a state based master client index.

Censoring & Competing Causes

PRAMS births censored due to competing causes of death (deaths not classified as maltreatment-related by the Alaska CDR committee) were followed and censored at the date of death. To detect out-of-state emigration, the PFD data was used. This source only allows for annual interval censoring, which required us to develop a set of rules for systematic classification (Table 1). For a visual depiction of these censoring rules please see Additional file 1: Figure S1.

The PFD provides only a crude annual estimation of censorship and even with manual review we are unable to identify exact dates, we examined multiple different censoring rules and compared the: 1) person-time estimation, 2) number of outcomes (reports of maltreatment) excluded based on the rule specification (recognizing that such exclusions may reflect lack of precision in the PFD and CPS dates), and 3) impact on the incidence estimate, crudely approximated as the number of events divided by the total person-time at risk (data presented in Additional file 1: Table S1). Based on the evaluated rule sets, our original *a priori* definition captured all outcomes within our observation window

Table 1 Censoring rules for ALCANLink cohort

Linkage observation	Censoring rule ^a
Linkage with vital death records and not classified by Alaska CDR as maltreatment related.	Censor at date of death
PRAMS births failing to link with any PFD year, or only linking with PFD records with only a non-Alaska address	Censor at 3-months from the date of birth
Last PFD indicates out-of-state residence or non-linkage with subsequent PFDs	Censor at the mid-year of the subsequent calendar year.
Linkage with last available PFD (2015 PFD year, 2014 calendar year)	Administrative censor at 12/31/2014

^aFor these rules we assumed a uniform distribution of out-of-state emigration during a calendar year. We also reasoned that (for example) as a 2011 PFD year application reflects residence during the 2010 calendar year, to be eligible for the 2010 PFD the subject must have physically resided in Alaska (aside for a few exceptions) the majority of that year and lost eligibility sometime during the subsequent year (2011)

and censored for out-of-state emigration using a conservative rule to maximize accrual of person-time and was thus utilized as a reasonable expectation. Finally, we scrutinized, using the manual review process described above, a subset of cases that responded to the three-year follow-up survey but had a survey date greater than the last PFD linkage date ($n = 23$). Among these 23 cases, 18 had either moved or died, 3 were either adoptions or had substantial name changes and missed as initial linkages, and 2 had no documented history with PFD but remained in the state. Based on our censoring rule for these 23 cases we estimated a contribution of 126 person-months compared to 120 person-months calculated when using the most probable departure dates identified through manual review.

Outcome ascertainment

Identifying and classifying maltreatment is problematic because it requires capturing an event that often occurs out of sight (US Department of Health and Human Services et al. 2016; Putnam-Hornstein and Needell 2011; Socolar et al. 1995; Dubowitz et al. 2005; Runyan et al. 2005; Hussey et al. 2005). Official reports, survey, medical records review, and multi-source data linkages have all been used to detect and classify maltreatment (Fluke et al. 2008). Official reports to child welfare agencies are known to under-represent the magnitude of the problem due to under-reporting (Ewigman et al. 1993; Drake and Zuravin 1998; Sedlak et al. 2010b). The process of screening and confirming maltreatment (substantiation) is influenced by policy, adequacy of information, and other external processes (Drake 1996). Although substantiations or confirmations are important, public health research has begun to shift towards the use of all recorded maltreatment reports by CPS agencies regardless of determination (Kohl et al. 2009). Studies document that children confirmed for maltreatment by child welfare experience similar negative health outcomes as those that are recorded but unconfirmed as well as those that are only reported but not evaluated for maltreatment (Parrish et al. 2011; Runyan et al. 2005; Hussey et al. 2005; Drake 1996; Putnam-Hornstein 2011; Leiter et al. 1994).

For the ALCANLink project and based the public health definitions proposed by CDC for classifying maltreatment (Leeb et al. 2008), we attempted to improve upon sole reliance on CPS records by broadening the range of agencies contributing maltreatment reports. In Alaska, state statute mandates that specified professionals (e.g. medical provider, education instructor), must report suspected maltreatment to the state child welfare agency. We developed a combined multi-agency reported maltreatment outcome measure to account for suspected non-reported maltreatment to CPS. The

multi-agency measurement includes child welfare records (including both screened in and screened out reports), 8 of the 10 active Child Advocacy Center (CAC) agencies reports, the Anchorage Police Department (APD) which covers nearly 50% of Alaska's population, and the Alaska Maternal Child Death Review (MCDR) maltreatment committee determinations. The Alaska MCDR committee reviews all child deaths occurring in Alaska and for each death classifies if any form of omission or commission caused or contributed to the death. Due to know underestimation of death certificate classifications and to be consistent with our sensitive reported maltreatment definition we included all deaths that the committee indicated abuse, neglect, or negligence "yes" or "yes probably" caused or contributed to the death. The CDC definitions provide a framework for quantifying potential maltreatment from a public health perspective and allow for a more sensitive cross-jurisdictional qualification of incidents (Jack 2010). For a more detailed description of the reported maltreatment classification see Additional file 1: Table S2.

Linkage methods

We implemented both deterministic and probabilistic methods to link PRAMS births with each dataset. Prior to all linkages we conducted systematic record set cleaning, including date, character, and case equalization, standardization of missing data and treatment of special characters, and removal of leading/trailing spaces. Using iterative linkages (deterministic followed by probabilistic) we reduced the amount of suspected matches requiring manual review. For probabilistic linkages we developed comparison patterns based on a Joarowinkler distance metric to account for typos, spelling errors, transpositions, and other edits or deletions between two strings or set of strings and dates. The probabilistic linkage approach automatically accepted matches when the first, last, and alias names, date of birth and sex were identical. Suspected matches that returned a probability match score between 0.85 and 0.99 were manually reviewed, while those below 0.85 were automatically rejected. For complete linkage details and methods on establishment of these thresholds for review please see Data linkages in the Additional file 1. The RecordLinkage package (Sariyar and Borg 2010) in the R environment (R Core Team 2014) was used for all data linkages.

Statistical analysis

We calculated the incidence proportion ("cumulative risk") of first multi-source report of maltreatment before age six years. We estimated the survivorship function $S(t)$ using a weighted Aalen hazard-based estimation (Klein and Moeschberger 2005) and 95% confidence interval on the log survival scale (Link 1984). We

calculated the weighted cumulative distribution function $F(t)$ from the weighted survivorship function $S(t)$ [$F(t) = 1 - S(t)$]. We used weighted $F(t)$ to estimate the incidence proportion of a multi-source maltreatment report before age six in the birth population. Frequency counts are presented as actual participant responses and weighted proportions from the complex sampling design are noted as %_w.

We created a dichotomous variable for censorship (yes or no) to assess the probability of censorship for a limited number of selected covariates obtained from both the birth certificate and PRAMS responses using logistic regression. The limited set of covariates selected for investigation to assess this potential bias included: as a proxy for military families if the birth was paid by Tri-care (yes, no); sex of the child (male, female); years of maternal education completed at delivery of child (<12 year, 12 + years); marital status at birth (married; unmarried); any maternal alcohol use during pregnancy as indicated on the birth certificate or PRAMS (yes, no); any maternal smoking during pregnancy as indicated on the birth certificate or PRAMS (yes, no); maternal race (Asian/Pacific Islander, Black, Native, White); birth defect indicated on the birth certificate (yes, no); mother or child on Medicaid at birth (yes, no); fathers name listed on birth certificate (yes, no); maternal age at birth (continuous); multi-agency maltreatment report (yes, no); mother reported being divorced/separated 12 months before pregnancy (yes, no); mother reported moving 12 months before pregnancy (yes, no); mother reported losing a job 12 months before pregnancy (yes, no); mother reported partner/husband losing a job 12 months before pregnancy (yes, no). These covariates were selected due to either being previously documented in the literature to be associated with maltreatment and hypothesized to potentially have differential population movement (Wu et al. 2004; Putnam-Hornstein and Needell 2011; Rentz et al. 2006; Putnam-Hornstein et al. 2013b). We then calculated and compared the incidence proportion and hazard ratio with and without out-of-state emigration to measure the impact of systematic bias on these selected values. We followed this same methodology to estimate the impact on both incidence proportion and hazard ratios assuming only deterministic linkages and reliance on CPS reported cases only, and in combination. All analyses were conducted in R 3.1.0 (R Core Team 2014) using the survey package (Lumley 2012).

Results

We successfully matched 1162 (94.1%) of the 1235 PRAMS births to at least one PFD record with an Alaska residence before the age of 6 years. Among the 73 non-matching births, 15 were deaths occurring during the first year of life. On average, deterministic linkages

captured 93.7% of all correct matches with annual PFD data. The PRAMS sample consistently linked with between 9% and 10% of PFD, CPS, APD, and CAC records (see Additional file 1: Table S3 for linkage rate details for ALCANLink project).

Outcome ascertainment

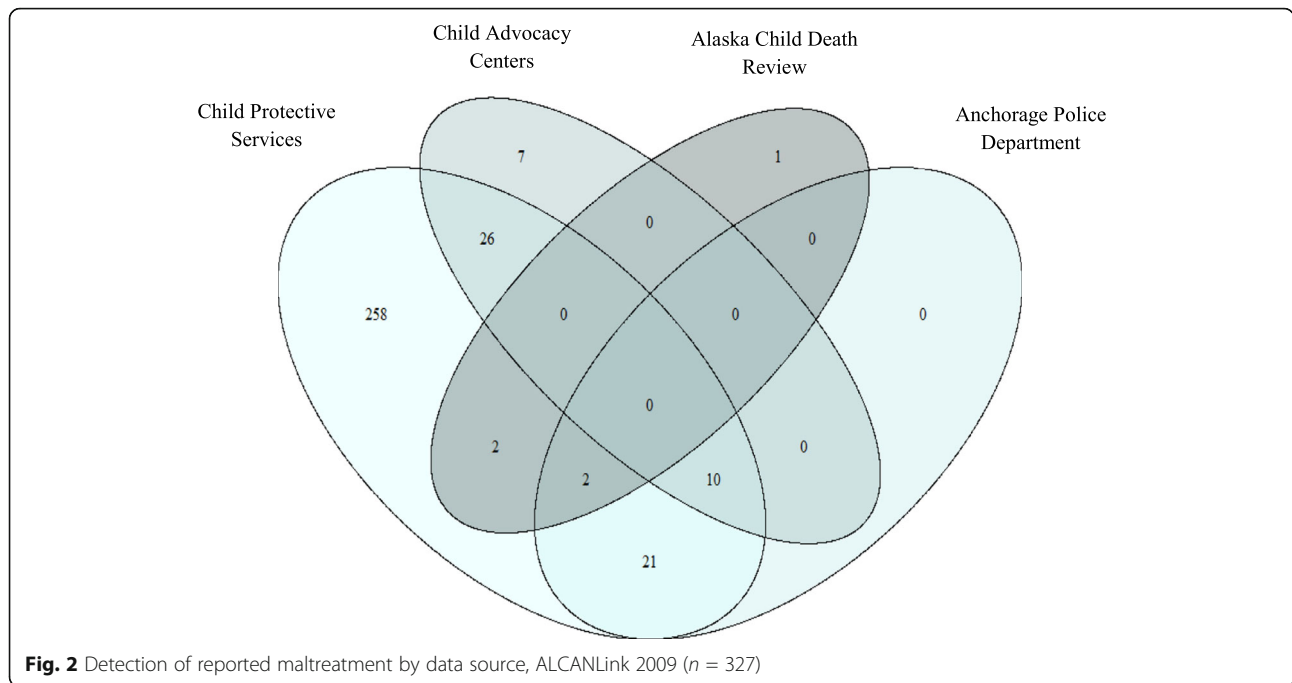
Among the 1235 PRAMS births, 327 (24.2%_w) had at least one multi-source report of maltreatment during the follow-up period. Of the 327 multi-source reports detected, CPS captured the overwhelming majority ($n = 319$, 98%), CAC captured 43 (13%), APD captured 33 (10%), and the CDR captured five (2%) fatalities (Fig. 2). The preponderance of reports occurred prior to age 1 year (39.1%_w), and monotonically decreased to 10.6% through age 5 years. Among the 1235 PRAMS births and considering only documented reports by CPS, 2.7%_w (95% CI: 1.7%_w, 6.5%_w) were reported for alleged sexual abuse, 5.1%_w (95% CI: 3.6%_w, 6.5%_w) were reported for alleged physical abuse, 9.1%_w (95% CI: 7.2%_w, 11.0%_w) were reported for alleged mental injury, and 21.0%_w (95% CI: 18.4%_w, 23.6%_w) were reported for alleged neglect among the birth population. The majority reported to CPS were due to neglect (88.7%_w), followed by mental injury (38.5%_w), physical abuse (21.4%_w), and sexual abuse (11.6%_w; totals sum to greater than 100% due to children being reported for multiple types of maltreatment).

Cohort follow-up

The cohort was followed for 5812.7 (86.9%) of the 6690.9 total potential person-years. Among the 1235 PRAMS births, 930 (75.3%) had complete cohort follow-up through the first 5 years of life. Approximately 4% of the births were lost-to-follow up annually. Among the 305 births lost-to-follow up during the project period regardless of outcome, 32% were lost prior to age 1 year and 49% prior to age 3 years. There were 23 total deaths, with 78% occurring prior to age 1 year. Cohort follow-up details are available in Table 2.

A total of 162 (14.5%_w) PRAMS births were paid by TRICARE (crude proxy for military births). Military paid births had substantially more out-of-state emigration before age six (73.2%_w vs 17.0%_w, $p < 0.001$), to such an extent that military paid births accounted for 42.5%_w of all emigration movements. Among military paid births only 54.6%_w of total potential person time was captured, compared with 91.9%_w among non-military births before age six. The proportion of first reported multi-agency events was slightly lower among military paid births compared to non-military births (18.4%_w vs 25.2%_w, $p = 0.183$).

Among the selected covariates assessed, the odds of out-of-state emigration censorship was higher among military paid births, married mothers at birth, maternal



Black race (relative to White), birth or pregnancy not being covered by Medicaid as indicated on the birth certificate, and maternal self-reporting husband or partner losing a job or reporting moving to a new address during the 12 months before the child’s birth. The odds of out-of-state emigration censorship was lower among children of Alaska Native mothers (relative to White) (Table 3).

Incidence proportion estimates and hazard ratios

We observed that before the age of 6 years 28.3%_w (95% CI: 23.6%_w, 33.0%_w) of the 2009 births to Alaska residents were the subject of at least one multi-source maltreatment report. Under the non-linkage assumption for out-of-state

emigration (assuming all non-linkages to any of the multi-source outcome agencies remained in the cohort outcome free) the incidence proportion calculated attenuated from 28.3%_w to 25.2%_w, an absolute difference of 3.1%. When we restricted our analysis to deterministic linkages only, the incidence proportion calculated attenuated from 28.3%_w to 20.1%_w, an absolute difference of 8.2%. Combining both sources of non-linkage error, the incidence proportion further attenuated to 18.5%_w, an absolute difference of 9.8%. Finally, the incidence proportion calculated when restricted to using only child welfare reports (27.7%_w; 95% CI: 23.0%_w, 32.4%_w) was nearly equivalent to the multi-source maltreatment report outcome (Fig. 3).

Table 2 Cohort follow-up by age in years to event, death, or censor, ALCANLink 2009 (n = 1235)

	[0–1)		[1–2)		[2–3)		[3–4)		[4–5)		[5–6)	
	N	% _w	N	% _w	N	% _w	N	% _w	N	% _w	N	% _w
Follow-Up												
Subjects entering age interval outcome free	1235	(100)	1017	(84.7)	917	(75.1)	831	(68.5)	747	(61.8)	674	(56.5)
Person-years accrued in interval	1118.4	(92.6)	968.3	(94.5)	871.3	(95.7)	788.7	(94.9)	708.6	(95.5)	325.2	(49.2)
Censoring												
Competing deaths	16	(77.8)	0	0	0	0	1	(11.1)	1	(11.1)	0	(0)
Loss to follow-up	67	(26.5)	45	(25.3)	37	(14.0)	41	(17.9)	36	(13.4)	11	(2.9)
Outcome Ascertainment												
First multi-source report ^a	135	(39.1)	54	(18.7)	48	(15.5)	42	(12.8)	36	(10.6)	12	(3.4)
CPS only report	131	(38.2)	53	(18.9)	48	(16.0)	41	(13.1)	34	(10.3)	12	(3.5)

^amulti-source report is the first report of maltreatment to either Child Protective Services, Child Advocacy Center, Anchorage Police Department, or the Maternal Child Death Review

Table 3 Unadjusted Odds of censorship among the PRAMS respondents, ALCANLink ($n = 1235$)

	Unweighted OR (95% CI)	pvalue	Weighted OR (95% CI)	pvalue
Birth Paid by Tricare (ref = no)	12.3 (8.5, 18.1)	<0.001	13.3 (8.4, 21.8)	<0.001
Male Sex (ref = Female)	1.2 (0.9, 1.5)	0.301	1.3 (0.9, 1.8)	0.087
Maternal education 12 + years (ref = <12 years)	1.3 (0.9, 1.9)	0.246	1.2 (0.8, 2.1)	0.387
Married (ref = unmarried)	2.2 (2.9, 1.7)	<0.001	2.3 (1.6, 3.4)	<0.001
Mom drink during pregnancy (ref = no drinking)	0.8 (0.3, 1.6)	0.545	0.9 (0.4, 2.6)	0.827
Mom Smoke during pregnancy (ref = no smoking)	0.8 (0.6, 1.1)	0.236	1.2 (0.8, 1.9)	0.469
Maternal Race (ref = White)				
Black	3.2 (1.7, 6.0)	<0.001	3.2 (1.3, 8.7)	0.013
Native	0.3 (0.2, 0.4)	<0.001	0.3 (0.2, 0.4)	<0.001
Asian/PI	0.9 (0.5, 1.5)	0.602	0.7 (0.3, 1.4)	0.332
No Birth Defect on BC (ref = Birth defect)	0.7 (0.3, 1.4)	0.248	0.9 (0.4, 2.6)	0.889
Birth or pregnancy covered by Medicaid (ref = yes)	2.4 (1.8, 3.1)	<0.001	2.3 (1.7, 3.2)	<0.001
Father Name on BC (ref = no father listed)	0.8 (0.5, 1.3)	0.385	0.7 (0.3, 1.3)	0.209
Maternal age at birth (continuous)	1.0 (0.9, 1.0)	0.281	1.0 (0.9, 1.0)	0.591
Maltreatment Report (ref = no maltreatment report)	0.6 (0.5, 0.9)	0.003	0.8 (0.5, 1.2)	0.275
Divorced/Separated 12 months before pregnancy (ref = no)	1.6 (1.1, 2.4)	0.018	1.3 (0.7, 2.1)	0.360
Reported moving 12 months before pregnancy (ref = no)	1.7 (1.3, 2.2)	<0.001	1.4 (1.0, 1.9)	0.048
Reported losing job 12 months before pregnancy (ref = no)	1.4 (0.9, 2.2)	0.111	1.9 (1.1, 3.2)	0.021
Reported partner/husband lost job 12 months before pregnancy (ref = no)	0.8 (0.5, 1.3)	0.426	1.0 (0.6, 1.8)	0.883

OR Odds Ratio

The hazard ratios for multiple risk and demographic factors were also influenced by failing to correctly account for censoring and/or restrictive data linkage (Table 4). Failing to account for out-of-state emigration underestimated the HR by 33% for military paid births (0.7 vs 1.1), and overestimated the HR by 11% for Alaska Native mothers (3.3 vs 3.0), and 10% for Medicaid births (4.1 vs 3.7). Limiting linkages to deterministic matches also resulted in biased HRs, with unmarried mothers (3.1 vs 3.8), and low maternal education (2.3 vs 3.1) all reporting underestimated HRs, and maternal smoking (3.6 vs 2.9) overestimating the HR. Combining both forms of error (failing to account for censoring and restrictive linkages), multiple factors and characteristics were both over and underestimated by 10% or more and include: military paid births, Alaska Native mothers, marital status, low education, child sex, young maternal age, maternal smoking during the 2 years before pregnancy, and reporting moving 12 months prior to birth.

Discussion

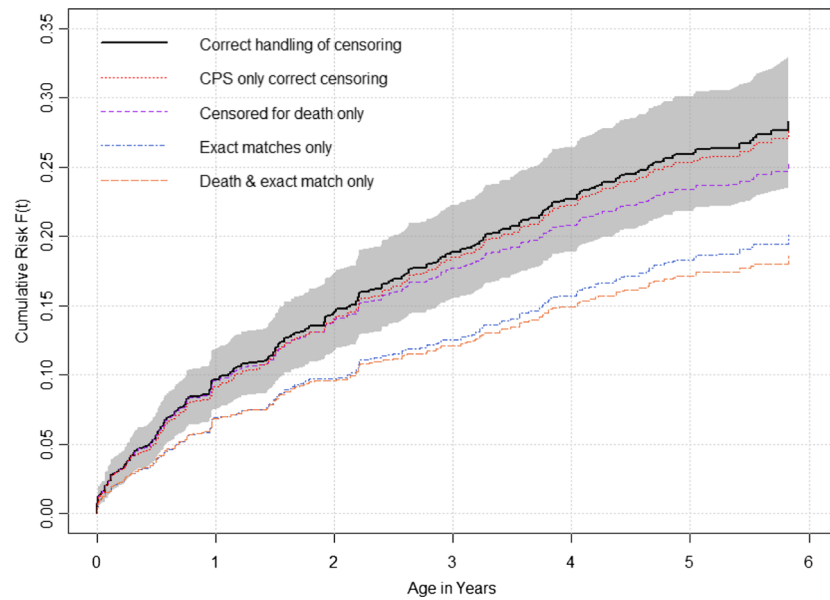
We documented that failing to account for out-of-state emigration and/or using restrictive linkage methods in longitudinal birth linkage studies will bias both the incidence proportion and effect estimates. Integrating unique data resources in the state of Alaska enabled us to examine these sources of bias. The manageable sample size facilitated comprehensive high confidence data

linkages and total cohort follow-up using the PFD. Furthermore, we demonstrated the utility of linking the PRAMS sampled child of a respondent mother with administrative data to effectively measure the incidence proportion of reported maltreatment over time in a representative birth population.

Outcome ascertainment data sources

All administrative studies using official reports of maltreatment (reports to CPS) are affected by potential detection bias (Hussey et al. 2006; McGee et al. 1995). It is important to note that not all maltreatment occurring in this population is reported, and that not all reports are substantiated by child welfare. It is assumed that many cases of maltreatment are never reported for a wide variety of reasons, including failure to seek care, stigmatization, minimal contact with mandatory reporters, missed diagnosis, among other reasons (Gilbert et al. 2009; Delaronde et al. 2000; Gunn et al. 2005). We attempted to improve upon reliance on CPS records alone by including reports to Child Advocacy Centers, Anchorage Police Department, and Child Death Review records. However, in this sample, we found that CPS reports captured nearly all (98%) of the ascertained maltreatment reports, and these additional administrative sources had essentially no influence on incidence proportion estimates of any maltreatment.

Future linkage studies, when any reports are the outcome, may gain little utility by linking additional sources



	Age interval (Years)					
	[0,1)	[1,2)	[2,3)	[3,4)	[4,5)	[5,6)
Correct handling of censoring ^a	0.097	0.146	0.189	0.227	0.259	0.283
CPS only ^b report correct censoring	0.092	0.141	0.185	0.222	0.254	0.277
Censored for death only	0.095	0.140	0.177	0.208	0.234	0.252
Exact matches only	0.069	0.097	0.125	0.157	0.183	0.201
Death and exact match only	0.068	0.095	0.121	0.149	0.171	0.185

^aCensored for competing cause mortality and out of state emigration
^bLimited to only Child Protective Services (CPS) reports, all other estimates use a multi-source report that include first report of maltreatment to either CPS, Child Advocacy Center, Anchorage Police Department, or the Maternal Child Death Review

Fig. 3 Risk of first maltreatment report, ALCANLink 2009 (n = 1235)

beyond CPS when all allegations of reports, regardless of screening determination and type are recorded and available through child welfare. However, it is clear that CPS records alone are an imperfect source of data for measuring child maltreatment and these conclusions may not apply to states with different types of child welfare agency structures (e.g. non-centralized) (Fallon et al. 2010). Other sources (not included in our study) may still be beneficial for increased detection of reports, for example medical records and self or proxy reported maltreatment obtained through survey (Robinson et al. 1997; Schnitzer et al. 2011; Turner et al. 2010; Finkelhor et al. 2009). A benefit of the ALCANLink methodology is that self/proxy reported maltreatment through survey can in theory be implemented through follow-up survey fairly easily. Alaska currently has a three-year follow-up survey to PRAMS and in 2016 (2013 PRAMS cohort) began asking questions about maltreatment experiences. Additional follow-up could also be done later in life for improved serial detection, and combined with administrative records would maximize ascertainment (Calderwood and Lessof 2009). Finally, because detection and reporting may be differential by maltreatment type additional

research is needed to determine if maltreatment type produce the same patterns of bias as seen with any maltreatment, and if particular sources increase/decrease detection of specific maltreatment types.

In addition to increased detection, improving outcome ascertainment and classification is also needed. Consensus review by expert panels is a standardized process that could be used to improve the reliability and consistency of maltreatment classification (Schnitzer et al. 2004). Such panels are already used for child death review processes, and could be extended to non-fatalities and unlike full birth cohort studies are potentially feasible for PRAMS based maltreatment linkage studies that have a manageable sample size.

Bias in incidence proportion

This study was able to achieve a high rate of follow-up through the first five years of life (especially for non-military paid births). Three quarters (75%) of the 2009 PRAMS births, representing 86% of the person-time of follow-up, had complete follow-up from birth to administrative censoring. High completeness of follow-up on the entire baseline population minimizes the potential

Table 4 Comparing the hazard ratios of first maltreatment report before age six years accounting for out of state emigration and linkage parameters, ALCANLink ($n = 1235$)

Factor	Accounting for emigration (probabilistic)			Not accounting for emigration (probabilistic)			Accounting for emigration (deterministic)			Not accounting for emigration (deterministic)					
	HR	95% CI		HR	95% CI		HR	95% CI		HR	95% CI				
Military paid births (ref = nonmilitary)	1.1	0.7, 1.8		0.7	0.4, 1.2		0.333	1.2	0.6, 2.1		0.081	0.7	0.4, 1.4		0.309
Race (ref = White)															
Alaska Native	3.0	2.2, 3.9		3.3	2.5, 4.4		0.114	3.0	2.2, 4.3		0.019	3.6	2.6, 5.0		0.204
Other	2.3	1.4, 3.8		2.3	1.4, 3.8		0.000	2.3	1.3, 4.3		0.006	2.4	1.3, 4.3		0.039
Unmarried marital status ref. = Married)	3.8	2.9, 5.1		4.1	3.1, 5.5		0.079	3.1	2.3, 4.4		0.177	3.1	2.2, 4.3		0.186
Medicaid birth (ref = no)	3.7	2.6, 5.2		4.1	2.9, 5.7		0.104	3.7	2.4, 5.6		0.005	4.0	2.6, 6.1		0.088
<12 years maternal education (ref = 12+ years)	3.1	2.3, 4.2		3.3	2.5, 4.5		0.058	2.3	1.6, 3.3		0.267	2.5	1.7, 3.6		0.212
Female (ref = Male)	1.3	1.0, 1.7		1.2	0.9, 1.6		0.024	1.2	0.9, 1.6		0.024	1.5	1.1, 2.0		0.152
Maternal age (ref = 25+ years)															
20–24 years	2.2	1.6, 3.0		2.2	1.6, 3.0		0.009	2.2	1.6, 3.0		0.009	2.2	1.6, 3.0		0.009
< 20 years	3.5	2.5, 5.1		3.7	2.5, 5.3		0.034	3.7	2.5, 5.3		0.034	3.1	2.1, 4.7		0.121
Birth weight < 2500 g (ref = 2500 + g)	1.7	1.4, 2.0		1.6	1.3, 1.9		0.060	1.7	1.3, 2.1		0.005	1.5	1.2, 1.9		0.096
Any Maternal smoking 2 years prior to birth (ref = No)	2.9	2.2, 3.9		3.0	2.2, 4.0		0.015	3.6	2.6, 5.0		0.225	3.5	2.5, 4.9		0.193
Any Birth defect listed on birth certificate (ref = No)	1.2	0.5, 2.8		1.2	0.5, 2.7		0.026	1.1	0.3, 3.3		0.125	1.1	0.4, 3.2		0.071
Mother Lost job ^b (ref = No)	2.4	1.7, 3.5		2.4	1.6, 3.4		0.025	2.1	1.3, 3.4		0.124	2.3	1.5, 3.7		0.034
Moved to a new address ^b (ref = No)	1.6	1.2, 2.1		1.6	1.2, 2.0		0.019	1.6	1.2, 2.2		0.012	1.4	1.0, 2.0		0.103

HR Hazard Ratio, 95% CI 95% Confidence Interval, ref reference level

^a% change compared to HR accounting for emigration and using probabilistic linkages

^bDuring the 12 months before the child was born

for bias in estimating incidence proportion and effect estimates over time (Rothman 2012). Using the PFD and death data allowed us to investigate the assumption made in nearly all birth cohort linkages studies that subjects who do not link with CPS records remain in the cohort outcome free. As we detected an increasing bias with length of follow-up, longitudinal birth cohort linkage studies without an annual census equivalent to the PFD and with follow-up beyond 3-years may need to adjust their estimates by a scale factor to produce unbiased estimates. Clearly, out-of-state emigration likely varies from state to state which could lead to differences in the impact of the non-linkage assumption bias. One possible way to address this issue and estimate a scale factors would be to derive inverse-probability-of-censoring weights from the Alaska data. Although a state may have differential out-of-state emigration patterns, with a sufficiently large predictor set, the inverse probability of censorship weights from the Alaska data may be transferable and allow for improved accuracy in subgroup comparisons of the incidence proportion over time.

By limiting to the PRAMS population-based subsample (as opposed to the entire Alaskan 2009 birth cohort) we were able to set liberal manual review ranges and only automatically accept linkages with perfect

matches on all linkage elements (first and last name, date of birth, sex, and residence). This resulted in high overall linkage success between sources with minimal effort, resources, and time. Studies linking entire birth cohorts may limit manual review and rely heavily on probabilistic cut points as a product of limited resources and data size resulting in unquantified sensitivity and specificity (Qayad and Zhang 2009). Variation in a state's capacity and ability to integrate data could impact comparability of estimates produced through large scale data linkage projects. Deterministic linkages alone underestimate the incidence proportion of maltreatment, thus probabilistic methods are needed. Birth population studies that are unable to extensively manually review probabilistic linkages should consider quantifying the impact of mismatches within the probabilistic linkages, and adjust estimates accordingly. Furthermore, publishing full linkage methodology in supplemental material can allow other researchers to replicate methods and develop comparable estimates. The benefit of the ALCANLink methodology to conduct a longitudinal birth cohort linkage study is reflected in the manageable population representative and standardized PRAMS sample methodology utilized. These methods may be a viable option for states to consider and can be implemented in a largely

systematic method allowing for improved comparability, regional, and even national assessments. Further development is needed to create a transferable platform for other PRAMS jurisdictions to utilize.

Bias in hazard ratios

We confirmed that the hazard ratio will be biased for some estimates if out-of-state emigration is unaccounted for, or linkages are made overly restrictive (as in the extreme case of exact matches only). The direction and magnitude of the error associated with the bias depends on the three-way association between the exposure, outcome, and factor influencing linkage and therefore can produce estimates that over or underestimate the true effect. We detected that the bias associated with linkage method can be strong enough to “pull” the effect across the null (as in the case for military paid births). Because the direction and magnitude of the bias is not readily predictable, results produced without addressing these forms of bias could result in erroneous conclusions, especially when comparing subgroups.

Comparison with prior published research

No national estimate is available for comparison of the incidence proportion estimate generated in the ALCANLink study and bias estimates measured. Researchers in California however, did observe that 14.8% of children born during 2006–2007 in the state were reported to child welfare before age 5 years (Putnam-Hornstein et al. 2014). They also reported that relative to White children, Native American children had 2.7 (95% CI: 2.6, 2.8) times the incidence proportion of being reported before age five (36.5% vs 13.7%). Using similar methodology to California’s estimates (not accounting for out-of-state emigration) we observed that 25.1%_w (95% CI: 21.0%_w, 29.1%_w) were reported to child welfare before age 5 years. This crude estimate is 1.7 times that of California. Similarly to California however, we observed that relative to White children, Alaska Native/American Indian (AN/AI) children had 2.8 (95% CI 2.3, 4.1) times the hazard of being reported before age five, with similar stratum specific estimates (41.3%_w; 95% CI: 33.2%_w, 49.3%_w) for AN/AI children vs 15.8%_w (95% CI: 11.4%_w, 20.2%_w) for White children. Although Alaska indicated a crude elevated estimate relative to California, variations in population movement between these states could impact any direct comparison. Further, the observed similarity in the stratum specific estimates indicate confounding by race and that race standardization may be needed to account for large differences in underlying population distributions to facilitate state-by-state comparisons.

Limitations

This study has a few notable limitations. 1) PRAMS respondents may be differential from the total sampled population resulting in selection bias. We conducted a post-hoc comparison with the full 2009 birth cohort and found a similar raw percentage of births reported to CPS suggesting a minimal impact on overall estimates. 2) This study accounted for censorship using the PFD based on a crude mid-year interval specification which may have led to erroneous or imprecise exclusions which could result in an overestimation of out-of-state emigration. The impact of this on our person-time estimation is unknown but could in theory result in an overestimation. However, we feel that the overall impact is likely minimal as we conducted extensive data mining from all available systems for respondents that failed to match with the PFD. Further, for those that had “breaks in PFD applications” for example applied in 2010 and again in 2011 but not in 2009 we assumed they remained in the state even though we were able to document for some cases intermittent movement (e.g. attendance at out-of-state school). Thus our conservative censoring rule may in fact still overestimate actual eligible person-time in the state for the population and would likely lead to attenuated results. 3) The multi-agency outcome measure was limited due to incomplete law enforcement and CAC data.

Conclusion

Child maltreatment is a substantial public health problem; however, etiologic analyses are needed to inform public health prevention efforts. Comprehensive population-representative data linkage studies are essential to detangling the multifaceted etiologies and interplay of factors that contribute to child maltreatment. Further, our confidence in assessing the impact of public health prevention efforts and policy over time relies on reliable, consistent estimates. PRAMS provide a rich set of measures for prospective cohort studies and when linked with administrative sources (such as Medicaid claims, hospital visits, and follow-up surveys) can efficiently increase the breadth of information available for longitudinal analysis. Other PRAMS states should consider the utility of the ALCANLink methodology for studying reported child maltreatment longitudinally. This study underscores the importance of manual review of data linkages to monitor linkage quality and suggests the need for increased transparency and standardizations in linkage studies. We also highlight the importance of adjustment for out-of-state emigration, especially for states like Alaska that may have large population movements among population subsets. Data linkage did not substantially improve the detection of reported

maltreatment in this study; additional research is needed to develop methods to improve the identification and classification of maltreatment.

Additional file

Additional file 1: Detailed information regarding cohort follow-up and censoring, maltreatment classification, and data linkage methods. (DOCX 161 kb)

Abbreviations

ALCANLink: Alaska longitudinal child abuse and neglect linkage project; CAC: Child advocacy center; CDC: Centers for disease control and prevention; CDR: Child death review; CPS: Child protective services; HR: Hazard ratio; PFD: Permanent fund dividend; PRAMS: Pregnancy risk assessment monitoring system; SCAN: Surveillance of child abuse and neglect; US: United States

Acknowledgements

This work was made possible by the contributions of the Alaska PRAMS team, Office of Children's Services staff, Child Advocacy Center Directors, and other statewide partners. We recognize the efforts of the multiple agencies that provided data making the development of the ALCANLink project possible. Specifically we recognize Kathy Perham-Hester, the Alaska PRAMS coordinator for conducting manuscript review and methodology discussion; Margaret Young and Abigail Newby-Kew with the Alaska MCH-Epidemiology Unit for manuscript review; Christy Lawton and Travis Erickson with the Office of Children's Services for manuscript review and comment.

Funding

The research was supported in part by the "Travis Fund". This fund, established by a private donor in Alaska, supports the scholarly research efforts of a graduate student at the University of North Carolina to benefit the children of Alaska. Its primary mission is to support efforts to quantify and elucidate the complex causal mechanisms leading to maltreatment. The primary author was supported by this research assistantship and in part by a grant from the Maternal and Child Health Bureau (Title V, Social Security Act, B04/MC-29332-01). S.W.M. and the University of North Carolina Injury Prevention Research Center are partially supported by an award from the National Center for Injury Prevention and Control, Centers for Disease Control and Prevention (R49/CE002479). The Alaska PRAMS was supported in part by a grant from the Centers for Disease Control and Prevention (U50/CCU01343 and UR6/DP000475).

Availability of data and materials

The dataset supporting the conclusions of this article are not publicly available and must be requested in writing from the Alaska Department of Health and Social Services, MCH-Epidemiology Unit (data can be obtained by completing the form at http://dhss.alaska.gov/dph/wcfh/Documents/mchepi/Fillable_Data_Request_Form_2-2016.pdf and submitting it to mchepi@alaska.gov). The dataset from which these conclusions are derived will be retained by the MCH-Epidemiology Unit for 10 years. A de-identified research relational database is available for academic researchers for defined research projects and are reviewed and approved on a case-by-case bases by the MCH-Epidemiology Unit senior epidemiologist.

Authors' contributions

JWP: Conceived and organized the study, analyzed these data and contributed to the manuscript. MES: Contributed to the development of the study methodology, analysis plan and writing. PGS: Contributed to the development of the methodology, analysis plan, and writing. PL: Contributed to the development of the study methodology, analysis plan, analysis activities and writing. JLD: Contributed to the development of the study methodology, analysis plan, analysis activities and writing. SWM: Contributed to the development of the study methodology, analysis plan, analysis activities and writing. He supervised the study. All authors read and approved the final manuscript.

Authors' information

Dr. Jared Parrish is currently the senior epidemiologist with the Maternal and Child Health Epidemiology Unit in the Alaska Division of Public Health and has spent over 10 years working on improving surveillance methodologies for child maltreatment, with an emphasis on the integration of data to quantify child maltreatment and improving maltreatment fatality classification consistency. He is skilled at utilizing epidemiologic methods and leveraging these methods for applied research that can be used to inform policy and practice. Dr. Parrish serves on multiple committees focused on improving national estimates of maltreatment related mortality, and has provided states and counties across the U.S. technical assistance on implementing maltreatment surveillance. Dr. Parrish focus his research on quantifying the influence of systematic error on effect estimates, data integration, and incorporating novel methods for applied surveillance with an emphasis on improving timeliness, efficacy, and utility of data that lead to prevention.

Ethics approval and consent to participate

The current study was approved by the University of North Carolina at Chapel Hill Office of Human Subjects Research. By responding to the PRAMS survey, respondents provide consent to have their responses combined with other information the health department has about them. The data linkages were conducted by the Alaska Surveillance of Child Abuse and Neglect (SCAN) program under the authority of the Alaska Division of Public Health. Identifiers were used for record linkages only, and conducted solely by state staff and not affixed to the survey or administrative responses. De-identified data were shared via a Data Use Agreement between the Alaska Division of Public Health and the University of North Carolina at Chapel Hill. The Alaska PRAMS project is reviewed by Institutional Review Boards at the University of Alaska Anchorage and the Centers for Disease Control and Prevention.

Consent for publication

Not applicable.

Competing interests

During the time of this study the corresponding author is both a doctoral student at the University of North Carolina and an employee of the Alaska Division of Public Health (ADPH). The ADPH requires all manuscripts to be reviewed by an executive leadership committee. No other potential conflicts of interest are noted.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Alaska Division of Public Health, Section of Women's, Children's, and Family Health, 3601 C St., Suite 358, Anchorage, AK 99503, USA. ²The University of North Carolina at Chapel Hill, School of Public Health, 2101 McGavran-Greenberg Hall, CB# 7435, Chapel Hill, NC 27599, USA. ³The National Center for the Review & Prevention of Child Deaths, c/o Michigan Public Health Institute, 1115 Massachusetts Ave. NW, Washington, DC 20005, USA. ⁴The University of North Carolina at Chapel Hill, School of Social Work, 325 Pittsboro St. #3550, Chapel Hill, NC 27516, USA. ⁵The University of North Carolina at Chapel Hill, Injury Prevention Research Center, 137 East Franklin St, CB# 7505, Chapel Hill, NC 27599, USA.

Received: 14 April 2017 Accepted: 28 June 2017

Published online: 07 August 2017

References

- Alaska Department of Revenue: Permanent Fund Dividend Division. Summary of Dividend Applications & Payments. 2016; Available at: <http://pfd.alaska.gov/Division-Info/Summary-of-Applications-and-Payments>. Accessed 12 Apr 2016.
- Bertolli J, Morgenstern H, Sorenson SB. Estimating the occurrence of child maltreatment and risk-factor effects: benefits of a mixed-design strategy in epidemiologic research. *Child Abuse Negl.* 1995;19(8):1007–16.
- Bohensky M. Bias in data linkage studies. *Methodological Developments in Data Linkage.* 2016:63–82.
- Brownell MD, Jutte DP. Administrative data linkage as a tool for child maltreatment research. *Child Abuse Negl.* 2013;37(2):120–4.

- Butchart A, Harvey AP, Mian M, Furniss T. Preventing child maltreatment: a guide to taking action and generating evidence. 2006.
- Calderwood L, Lessof C. Enhancing longitudinal surveys by linking to administrative data. *Methodology of longitudinal surveys*. Chichester: John Wiley & Sons; 2009. p. 55–72.
- Cicchetti D. Advances and challenges in the study of the sequelae of child maltreatment. *Dev Psychopathol*. 1994;6(01):1–3.
- Cicchetti D, Carlson V. *Child maltreatment: Theory and research on the causes and consequences of child abuse and neglect*. Cambridge University Press; 1989.
- Delaronde S, King G, Bendel R, Reece R. Opinions among mandated reporters toward child maltreatment reporting policies. *Child Abuse Negl*. 2000;24(7):901–10.
- Drake B. Unraveling “unsubstantiated”. *Child Maltreat*. 1996;1(3):261–71.
- Drake B, Zuravin S. Bias in child maltreatment reporting: revisiting the myth of classlessness. *Am J Orthop*. 1998;68(2):295.
- Dubowitz H, Pitts SC, Litrownik AJ, Cox CE, Runyan D, Black MM. Defining child neglect based on child protective services data. *Child Abuse Negl*. 2005;29(5):493–511.
- Ewigman B, Kivlahan C, Land G. The Missouri child fatality study: underreporting of maltreatment fatalities among children younger than five years of age, 1983 through 1986. *Pediatrics*. 1993;91(2):330–7.
- Fallon B, Trocmé N, Fluke J, MacLaurin B, Tonmyr L, Yuan Y. Methodological challenges in measuring child maltreatment. *Child Abuse Negl*. 2010;34(1):70–9.
- Finkelhor D, Turner H, Ormrod R, Hamby SL. Violence, abuse, and crime exposure in a national sample of children and youth. *Pediatrics*. 2009;124(5):1411–23.
- Fluke JD, Shusterman GR, Hollinshead DM, Yuan YT. Longitudinal analysis of repeated child abuse reporting and victimization: multistate analysis of associated factors. *Child Maltreat*. 2008;13(1):76–88.
- Gessner BD, Moore M, Hamilton B, Muth PT. The incidence of infant physical abuse in Alaska. *Child Abuse Negl*. 2004;28(1):9–23.
- Gilbert R, Kemp A, Thoburn J, Sidebotham P, Radford L, Glaser D, et al. Recognizing and responding to child maltreatment. *Lancet*. 2009;373(9658):167–80.
- Goldsmith S. *The Alaska Permanent Fund Dividend: an experiment in wealth distribution*. 9th International Congress, BIEN, Geneva; 2002.
- Greene N, Greenland S, Olsen J, Nohr EA. Estimating bias from loss to follow-up in the Danish National Birth Cohort. *Epidemiology*. 2011;22(6):815–22.
- Gunn VL, Hickson GB, Cooper WO. Factors affecting pediatricians’ reporting of suspected child maltreatment. *Ambul Pediatr*. 2005;5(2):96–101.
- Holman CJ, Bass AJ, Rouse IL, Hobbs MS. Population-based linkage of health records in Western Australia: development of a health services research linked database. *Aust N Z J Public Health*. 1999;23(5):453–9.
- Hussey JM, Marshall JM, English DJ, Knight ED, Lau AS, Dubowitz H, et al. Defining maltreatment according to substantiation: distinction without a difference? *Child Abuse Negl*. 2005;29(5):479–92.
- Hussey JM, Chang JJ, Kotch JB. Child maltreatment in the United States: prevalence, risk factors, and adolescent health consequences. *Pediatrics*. 2006;118(3):933–42.
- Jack SM. The role of public health in addressing child maltreatment in Canada. *Chronic Dis Can*. 2010;31(1):39–44.
- Jonson-Reid M, Drake B. Multisector longitudinal administrative databases: an indispensable tool for evidence-based policy for maltreated children and their families. *Child Maltreat*. 2008;13(4):392–9.
- Jutte DP, Roos LL, Brownell MD. Administrative record linkage as a tool for public health research. *Annu Rev Public Health*. 2011;32:91–108.
- Klein JP, Moeschberger ML. *Survival analysis: techniques for censored and truncated data*. New York: Springer Science & Business Media; 2005.
- Kohl PL, Jonson-Reid M, Drake B. Time to leave substantiation behind: findings from a national probability study. *Child Maltreat*. 2009;14(1):17–26.
- Leeb RT, Paulozzi L, Melanson C, Simon T, Arias I. *Child maltreatment surveillance: uniform definitions for public health and recommended data elements*; 2008. p. 1.0.
- Leiter J, Myers KA, Zingraff MT. Substantiated and unsubstantiated cases of child maltreatment: do their consequences differ? *Soc Work Res*. 1994;18(2):67–82.
- Link CL. Confidence intervals for the survival function using Cox’s proportional-hazard model with covariates. *Biometrics*. 1984;601–9.
- Lumley T. *Survey: analysis of complex survey samples*. 2012;R package version 3.28.2.
- McGee RA, Wolfe DA, Yuen SA, Wilson SK, Carnochan J. The measurement of maltreatment: a comparison of approaches. *Child Abuse Negl*. 1995;19(2):233–49.
- Parrish JW, Young MB, Perham-Hester KA, Gessner BD. Identifying risk factors for child maltreatment in Alaska: a population-based approach. *Am J Prev Med*. 2011 Jun;40(6):666–73.
- Putnam-Hornstein E. Report of maltreatment as a risk factor for injury death: a prospective birth cohort study. *Child Maltreat*. 2011;16(3):163–74.
- Putnam-Hornstein E, Needell B. Predictors of child protective service contact between birth and age five: an examination of California’s 2002 birth cohort. *Child Youth Serv Rev*. 2011;33(8):1337–44.
- Putnam-Hornstein E, Wood JN, Fluke J, Yoshioka-Maxwell A, Berger RP. Preventing severe and fatal child maltreatment: making the case for the expanded use and integration of data. *Child Welfare*. 2013a;92(2):59–75.
- Putnam-Hornstein E, Needell B, King B, Johnson-Motoyama M. Racial and ethnic disparities: a population-based examination of risk factors for involvement with child protective services. *Child Abuse Negl*. 2013b;37(1):33–46.
- Putnam-Hornstein E, Mitchell M, Hammond I. A birth cohort study of involvement with child protective services before age 5. *Children’s Data Network*. 2014;2:1–11.
- Qayad MG, Zhang H. Accuracy of public health data linkages. *Matern Child Health J*. 2009;13(4):531–8.
- R Core Team. R: a language and environment for statistical computing. 2014;3.1.0.
- Rentz ED, Martin SL, Gibbs DA, Clinton-Sherrod M, Hardison J, Marshall SW. Family violence in the military: a review of the literature. *Trauma Violence Abuse*. 2006;7(2):93–108.
- Robinson JR, Young TK, Roos LL, Gelskey DE. Estimating the burden of disease. Comparing administrative data and self-reports. *Med Care*. 1997;35(9):932–47.
- Rothman KJ. *Epidemiology: an introduction*. New York: Oxford University Press; 2012.
- Rothman KJ, Greenland S, Lash TL. *Modern epidemiology*. Philadelphia: Lippincott Williams & Wilkins; 2008.
- Runyan DK, Curtis PA, Hunter WM, Black MM, Kotch JB, Bangdiwala S, et al. LONGSCAN: a consortium for longitudinal studies of maltreatment and the life course of children. *Aggress Violent Behav*. 1998;3(3):275–85.
- Runyan DK, Cox CE, Dubowitz H, Newton RR, Upadhyaya M, Kotch JB, et al. Describing maltreatment: do child protective service reports and research definitions agree? *Child Abuse Negl*. 2005;29(5):461–77.
- Sariyar M, Borg A. The RecordLinkage package: detecting errors in data. *The R Journal*. 2010;2(2):61–7.
- Schnitzer PG, Slusher P, Van Tuinen M. Child maltreatment in Missouri: combining data for public health surveillance. *Am J Prev Med*. 2004;27(5):379–84.
- Schnitzer PG, Slusher PL, Kruse RL, Tarleton MM. Identification of ICD codes suggestive of child maltreatment. *Child Abuse Negl*. 2011;35(1):3–17.
- Sedlak AJ, Mettenberg J, Basena M, Peta I, McPherson K, Greene A. Fourth national incidence study of child abuse and neglect (NIS-4). Washington, DC: US Department of Health and Human Services. Retrieved on July 2010;9; 2010a.
- Sedlak AJ, Mettenberg J, Basena M, Peta I, McPherson K, Greene A. Fourth national incidence study of child abuse and neglect (NIS-4). Washington, DC: US Department of Health and Human Services. Retrieved on July 2010;9; 2010b.
- Shulman HB, Gilbert BC, Lansky A. The pregnancy risk assessment monitoring system (PRAMS): current methods and evaluation of 2001 response rates. *Public Health Rep*. 2006;121(1):74–83.
- Socolar RR, Runyan DK, Amaya-Jackson L. Methodological and ethical issues related to studying child maltreatment. *J Fam Issues*. 1995;16(5):565–86.
- Stanley FJ, Croft ML, Gibbins J, Read AW. A population database for maternal and child health research in Western Australia using record linkage. *Paediatr Perinat Epidemiol*. 1994;8(4):433–47.
- Tonmyr L, Hovdestad WE, Draca J. Commentary on Canadian child maltreatment data. *J Interpers Violence*. 2014;29(1):186–97.
- Turner HA, Finkelhor D, Ormrod R, Hamby SL. Infant victimization in a nationally representative sample. *Pediatrics*. 2010;126(1):44–52.
- US Department of Health and Human Services, Administration for Children and Families, Administration for Children, Youth and Families, Children’s Bureau. *Child maltreatment 2014*. 2016.
- Van Horne BS, Moffitt KB, Canfield MA, Case AP, Greeley CS, Morgan R, et al. Maltreatment of children under age 2 with specific birth defects: a population-based study. *Pediatrics*. 2015;136(6):e1504–12.
- Wu SS, Ma C, Carter RL, Ariet M, Feaver EA, Resnick MB, et al. Risk factors for infant maltreatment: a population-based study. *Child Abuse Negl*. 2004;28(12):1253–64.