Agreement Between Prospective and Retrospective Measures of Childhood Maltreatment: A Systematic Review and Meta-analysis

Jessie R. Baldwin, PhD; Aaron Reuben, MEM; Joanne B. Newbury, PhD; Andrea Danese, MD, PhD

**IMPORTANCE** Childhood maltreatment is associated with mental illness. Researchers, clinicians, and public health professionals use prospective or retrospective measures interchangeably to assess childhood maltreatment, assuming that the 2 measures identify the same individuals. However, this assumption has not been comprehensively tested.

**OBJECTIVE** To meta-analyze the agreement between prospective and retrospective measures of childhood maltreatment.

**DATA SOURCES** MEDLINE, PsycINFO, Embase, and Sociological Abstracts were searched for peer-reviewed, English-language articles from inception through January 1, 2018. Search terms included child* maltreatment, child* abuse, child* neglect, child bull*, child* trauma, child* advers*, and early life stress combined with prospective* and cohort.

**STUDY SELECTION** Studies with prospective measures of childhood maltreatment were first selected. Among the selected studies, those with corresponding retrospective measures of maltreatment were identified. Of 450 studies with prospective measures of childhood maltreatment, 16 had paired retrospective data to compute the Cohen κ coefficient.

**DATA EXTRACTION AND SYNTHESIS** Multiple investigators independently extracted data according to PRISMA and MOOSE guidelines. Random-effects meta-analyses were used to pool the results and test predictors of heterogeneity.

**MAIN OUTCOMES AND MEASURES** The primary outcome was the agreement between prospective and retrospective measures of childhood maltreatment, expressed as a κ coefficient. Moderators of agreement were selected a priori and included the measure used for prospective or retrospective assessment of childhood maltreatment, age at retrospective report, sample size, sex distribution, and study quality.

**RESULTS** Sixteen unique studies including 25,471 unique participants (52.4% female [SD, 10.6%]; mean [SD] age, 30.6 [11.6] years) contained data on the agreement between prospective and retrospective measures of childhood maltreatment. The agreement between prospective and retrospective measures of childhood maltreatment was poor, with κ = 0.19 (95% CI, 0.14-0.24; P < .001). Agreement was higher when retrospective measures of childhood maltreatment were based on interviews rather than questionnaires (Q = 4.1521; df = 1; P = .04) and in studies with smaller samples (Q = 4.2251; df = 1; P = .04). Agreement was not affected by the type of prospective measure used, age at retrospective report, sex distribution of the sample, or study quality.

**CONCLUSIONS AND RELEVANCE** Prospective and retrospective measures of childhood maltreatment identify different groups of individuals. Therefore, children identified prospectively as having experienced maltreatment may have different risk pathways to mental illness than adults retrospectively reporting childhood maltreatment. Researchers, clinicians, and public health care professionals should recognize these critical measurement differences when conducting research into childhood maltreatment and developing interventions.
Do prospective and retrospective measures of childhood maltreatment identify the same individuals? This question has captivated psychiatrists and psychologists since the inception of our discipline and still permeates many aspects of our professions. Researchers use retrospective reports as a shortcut to better understand the consequences of childhood maltreatment without the significant time or financial investment needed to undertake cohort studies. Clinicians use retrospective reports to swiftly identify individuals who are at heightened risk of mental illness by virtue of their exposure to childhood maltreatment. Public health professionals use retrospective reports to pragmatically estimate the health burden associated with exposure to childhood maltreatment. All these practices rely on the assumption that retrospective reports and prospective measures identify the same, or at least similar, groups of individuals. However, qualitative reviews have raised concerns about the validity of this assumption. Herein we present, to our knowledge, the first quantitative assessment of the agreement between retrospective reports and prospective measures of childhood maltreatment.

Methods

Data Sources
We performed a systematic review and meta-analysis in line with the PRISMA and MOOSE guidelines, following an a priori-defined protocol (eMethods and eTables 1 and 2 in the Supplement). We searched MEDLINE, PsycINFO, Embase, and Sociological Abstracts for peer-reviewed articles written in English and published from database inception to January 1, 2018, that included prospective assessments of childhood maltreatment. We used the following search terms: childhood maltreatment, child abuse, child neglect, child maltreatment, child trauma, child adversity, and early life stress combined with prospective and cohort.

Study Selection
Two authors (J.R.B. and A.R.) independently screened titles and abstracts of all articles retrieved from the search before reviewing the full text of potentially eligible studies. We included original, peer-reviewed articles with prospectively collected information on childhood maltreatment (age <18 years). Measures of maltreatment (sexual abuse, physical abuse, emotional abuse, and neglect), domestic violence, bullying, institutionalization, and broader measures of adverse childhood experiences that included maltreatment were used to define overall childhood maltreatment. From the articles with prospective assessment of childhood maltreatment, we selected studies with data on corresponding retrospective measures (defined as subsequent assessment of the same individuals’ exposure undertaken at any age).

Data Extraction
Three authors (J.R.B., A.R., and J.B.N.) independently extracted data from all studies with prospective assessment of childhood maltreatment on sample characteristics (cohort name, sample size, location, age at latest assessment, and sex distribution), childhood maltreatment type(s) assessed, prospective measure type(s) (official records, interview, and questionnaire), source (child protection services, hospital records, parent, child, teacher, or multiple informants), and availability of retrospective measures. If retrospective measures of childhood maltreatment were available, 2 authors (J.R.B. and A.D.) subsequently extracted data on the retrospective measurement type (interview or questionnaire) and source, agreement between prospective and retrospective measures, and study quality. Inconsistencies were resolved in consensus meetings and confirmed with the authors of the primary studies when necessary. Relevant missing information was requested from authors.

Statistical Analysis
The extracted data were converted to contingency tables comparing prospectively identified childhood maltreatment (yes or no) with retrospectively reported childhood maltreatment (yes or no). From the contingency tables, we derived estimates of prevalence, raw percentage of agreement between measures, and Cohen 𝜅 coefficient. Some studies only reported a 𝜅. Prevalence and raw percentage of agreement estimates were used for descriptive purposes, and our primary outcome was the 𝜅. The other extracted variables were used to explain the heterogeneity in the 𝜅 across studies.

We first described the prevalence of childhood maltreatment based on prospective and retrospective measures of childhood maltreatment. We then examined the (1) prevalence of retrospective reports of childhood maltreatment among those with prospective observations, (2) prevalence of prospective observations among those with retrospective reports, and (3) raw percentage of agreement between the 2 measures through meta-analyses of proportions for different childhood maltreatment types with the metafor R package.

Data from contingency tables were first converted using the Freeman-Tukey double arc sine transformation to normalize and stabilize the variance of the sampling distribution, then aggregated using random-effects model meta-analyses, and finally back-transformed using the inverse of the Freeman-Tukey double arc sine.
Agreement Between Prospective and Retrospective Measures of Childhood Maltreatment

Results

Search Results
The study selection procedure is summarized in Figure 1, and further details are provided in the eResults in the Supplement. We identified 450 independent studies with prospective measures of childhood maltreatment (eTable 5 in the Supplement). Of these studies, we identified 20 studies (26365 participants) with at least partial data on the agreement between prospective and retrospective measures of childhood maltreatment and 16 unique studies13-28 with 25471 unique participants (52.4% female [SD, 10.6%]; mean [SD] age, 30.6 [11.6] years) with direct measures or paired data sufficient to compute measures of \( \kappa \). Details of these studies are reported in the Table.

Overlap Between Individuals Identified by Prospective or Retrospective Measures of Childhood Maltreatment

eFigure 1 in the Supplement displays the range of prevalence estimates for childhood maltreatment based on 32 prospective and retrospective measures extracted from 15 studies.14-20,22-28,31 Capitalizing on the paired nature of the data, we next analyzed (I) the prevalence of retrospective reports...
<table>
<thead>
<tr>
<th>Source (Study Name)</th>
<th>No. of Participants</th>
<th>Male, %</th>
<th>Sample Description</th>
<th>Type of Maltreatment Assessed</th>
<th>Prospective Measure</th>
<th>Retrospective Measure</th>
<th>Age at Report, y</th>
<th>Agreement Data</th>
<th>Paired Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Robins,29 1966 (Deviant Children Grown Up)</td>
<td>411</td>
<td>NA</td>
<td>Children with antisocial behavior problems and matched controls from St Louis, Missouri</td>
<td>Maltreatment</td>
<td>Mixed CPS, self</td>
<td>Interview</td>
<td>Self</td>
<td>45</td>
<td>No No</td>
</tr>
<tr>
<td>Williams,30 1994</td>
<td>129</td>
<td>100</td>
<td>Sexually abused females from a Northeastern US city examined at a hospital, 1973-1975</td>
<td>Sexual abuse</td>
<td>Mixed Medical records, self, parent</td>
<td>Interview</td>
<td>Self</td>
<td>24.5</td>
<td>No No</td>
</tr>
<tr>
<td>Widom and Shepard,31 1996; Widom and Morris,32 1997; Raphael et al,33 2001</td>
<td>1181-1196</td>
<td>48.7</td>
<td>Maltreated children and matched controls from a metropolitan area in the Midwestern United States</td>
<td>Maltreatment, physical abuse, and sexual abuse</td>
<td>Records CPS</td>
<td>Interview</td>
<td>Self</td>
<td>28.7-29.2</td>
<td>Yes Yes</td>
</tr>
<tr>
<td>Johnson et al,34 1999 (Children in the Community Study)</td>
<td>639</td>
<td>47.7</td>
<td>Children from 2 counties in northern New York State</td>
<td>Maltreatment</td>
<td>Records CPS</td>
<td>Interview</td>
<td>Self</td>
<td>22.3</td>
<td>Yes Yes</td>
</tr>
<tr>
<td>Goodman et al,35 2003</td>
<td>175</td>
<td>80.6</td>
<td>Sexually abused children referred from DAs’ offices in Denver, Colorado, 1985-1987</td>
<td>Sexual abuse</td>
<td>Records CPS</td>
<td>Interview</td>
<td>Self</td>
<td>23</td>
<td>No No</td>
</tr>
<tr>
<td>Tajima et al,36 2004 (Lehigh Longitudinal Study)</td>
<td>409</td>
<td>45.7</td>
<td>Children from 2 counties in eastern Pennsylvania with overrepresentation from low-income families</td>
<td>Physical abuse</td>
<td>Interview Parent</td>
<td>Interview</td>
<td>Self</td>
<td>18</td>
<td>Yes Yes</td>
</tr>
<tr>
<td>White et al,37 2007 (Rutgers Health and Human Development Project)</td>
<td>359</td>
<td>50.7</td>
<td>Adolescents from New Jersey recruited through random telephone sampling</td>
<td>Physical abuse</td>
<td>Questionnaire Self</td>
<td>Questionnaire</td>
<td>Self</td>
<td>30.5</td>
<td>Yes Yes</td>
</tr>
<tr>
<td>Everson et al,38 2008 (Longitudinal Studies of Child Abuse and Neglect)</td>
<td>348-350</td>
<td>51</td>
<td>Children from Eastern and Southeastern United States identified to be high risk for poor medical or developmental outcomes</td>
<td>Physical, sexual, and emotional abuse</td>
<td>Records CPS</td>
<td>Interview</td>
<td>Self</td>
<td>12</td>
<td>Yes Yes</td>
</tr>
<tr>
<td>Shaffer et al,39 2008 (Minnesota Longitudinal Study of Parents and Children)</td>
<td>125-170</td>
<td>47.1</td>
<td>Children born to pregnant women with low socioeconomic status who attended prenatal care in Minnesota</td>
<td>Maltreatment, physical and sexual abuse, and neglect</td>
<td>Mixed CPS, self, parent, teacher</td>
<td>Interview</td>
<td>Self</td>
<td>19</td>
<td>Yes Yes</td>
</tr>
<tr>
<td>Barnes et al,40 2009</td>
<td>179</td>
<td>100</td>
<td>Sexually abused females and matched controls in Washington, DC</td>
<td>Sexual abuse</td>
<td>Records CPS</td>
<td>Interview</td>
<td>Self</td>
<td>24</td>
<td>No No</td>
</tr>
<tr>
<td>Scott et al,41 2010 (The New Zealand Mental Health Survey)</td>
<td>2144</td>
<td>55.1</td>
<td>New Zealand, Maori, and Pacific adults in New Zealand</td>
<td>Maltreatment</td>
<td>Records CPS</td>
<td>Interview</td>
<td>Self</td>
<td>21.9</td>
<td>Yes Yes</td>
</tr>
<tr>
<td>Denholm et al,42 2013 (National Child Development Study)43</td>
<td>8461</td>
<td>NA</td>
<td>Children from Great Britain born in a single week in 1958</td>
<td>Neglect</td>
<td>Multiple Parent, teacher</td>
<td>Questionnaire</td>
<td>Self</td>
<td>45</td>
<td>Yes Yes</td>
</tr>
<tr>
<td>Elwyn and Smith,44 2013 (Rochester Youth Development Study)</td>
<td>846</td>
<td>27.1</td>
<td>Children from Rochester, New York, with overrepresentation of boys and students from high-crime census tracts</td>
<td>Maltreatment</td>
<td>Records CPS</td>
<td>Interview</td>
<td>Self</td>
<td>22.7</td>
<td>Yes Yes</td>
</tr>
</tbody>
</table>

(continued)
Table. Description of Studies With Prospective and Retrospective Measures of Childhood Maltreatment (continued)

<table>
<thead>
<tr>
<th>Source (Study Name)</th>
<th>No. of Participants</th>
<th>Female, %</th>
<th>Sample Description</th>
<th>Type of Maltreatment Assessed</th>
<th>Prospective Measure</th>
<th>Retrospective Measure</th>
<th>Age at Report, y</th>
<th>Agreement Data</th>
<th>Paired Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patten et al.21 2015 (National Longitudinal Study Survey of Children and Youth/National Population Health Survey)</td>
<td>1977</td>
<td>48.3</td>
<td>Children from Canada assessed as part of 2 linked surveys in childhood and adulthood</td>
<td>ACEs</td>
<td>Interview Parent</td>
<td>Interview Self</td>
<td>NA</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Plant et al.,27 2015 (South London Child Development Study)</td>
<td>97</td>
<td>52.4</td>
<td>Children born to pregnant women who attended antenatal clinics in South London, UK</td>
<td>Physical and sexual abuse</td>
<td>Interview Parent, Self</td>
<td>Questionnaire Self</td>
<td>25</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Mills et al.,23 2016 (Mater-University of Queensland Study of Pregnancy)</td>
<td>3739</td>
<td>57.3</td>
<td>Children born to pregnant women who attended maternity services at a hospital in Brisbane, Australia</td>
<td>Sexual abuse</td>
<td>Records CPS Questionnaire Self</td>
<td>21</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Reuben et al.,23 2016 (Dunedin Multidisciplinary Health and Development Study)</td>
<td>950</td>
<td>49.9</td>
<td>Children born in Dunedin, New Zealand, in 1972-1973</td>
<td>Physical, sexual, and emotional abuse, neglect, and ACEs</td>
<td>Mixed CPS, parent, research workers, teacher</td>
<td>Interview Self</td>
<td>38</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Shenk et al.,24 2016 (Female Adolescent Development Study)</td>
<td>514</td>
<td>100</td>
<td>Maltreated females and matched controls from an area in the US Midwest</td>
<td>Maltreatment</td>
<td>Records CPS Interview Self</td>
<td>15.7</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Newbery et al.,25 2018 (E-Risk Longitudinal Twin Study)</td>
<td>2055</td>
<td>51</td>
<td>Twin children from England and Wales born in 1994-1995</td>
<td>Physical, sexual, and emotional abuse, neglect, and maltreatment</td>
<td>Interview Parent, research worker</td>
<td>Interview Self</td>
<td>18</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Naicker et al.,28 2017 (Birth to Twenty Plus Cohort)</td>
<td>1506-1565</td>
<td>51.8</td>
<td>Children born in Soweto-Johannesburg, South Africa, during 7 weeks in 1990</td>
<td>Sexual, physical, and emotional abuse</td>
<td>Questionnaire Self</td>
<td>Questionnaire Self</td>
<td>23</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Abbreviations: ACEs, adverse childhood experiences; CPS, child protection services; DA, district attorney; NA, not available.

* Mixed* refers to studies assessing childhood maltreatment using a combination of records and interviews or questionnaires; *multiple* refers to studies assessing childhood maltreatment using interviews and questionnaires (excluding records).

Data on agreement between prospective and retrospective measures were obtained from study authors rather than from the published article.
of childhood maltreatment among those with prospective observations, (2) the prevalence of prospective observations among those with retrospective reports, and (3) the raw percentage of agreement between prospective and retrospective measures. A random-effects meta-analysis of 7 studies\(^\text{14,17,18,20,24,25,32}\) focusing on a broad measure of child maltreatment revealed that the prevalence of retrospective reports among those with prospective observations was 48% (95% CI, 34%-62%; \(I^2 = 96\%\)); the prevalence of prospective observations among those who retrospectively reported childhood maltreatment was 44% (95% CI, 24%-65%; \(I^2 = 99\%\)); and the percentage of agreement between prospective and retrospective measures of childhood maltreatment was 76% (95% CI, 67%-84%; \(I^2 = 99\%\)). Therefore, on average, 52% of individuals with prospective observations of maltreatment did not retrospectively report it, and 56% of individuals retrospectively reporting maltreatment did not have concordant prospective observations (Figure 2A).

We next undertook sensitivity analyses to test whether the overlap between individuals identified as maltreated through prospective or retrospective measures varied as a function of the type of maltreatment (Figure 2B-E and eTable 3 in the Supplement). First, the prevalence of retrospective reports among those with prospective observations in 8 studies\(^\text{13,16,17,22,23,25,27,28}\) that included childhood sexual abuse was 45% (95% CI, 18%-75%; \(I^2 = 97\%\)); the prevalence of prospective observations among those who retrospectively reported childhood sexual abuse was 25% (95% CI, 12%-41%; \(I^2 = 96\%\)); and the percentage of agreement between prospective and retrospective measures of childhood sexual abuse was 86% (95% CI, 75%-94%; \(I^2 = 99\%\)). Second, the prevalence of retrospective reports among those with prospective observations in the 9 studies\(^\text{15-17,19,23,25-28,31}\) that included childhood physical abuse was 38% (95% CI, 18%-60%; \(I^2 = 98\%\)); the prevalence of prospective observations among those who retrospectively reported childhood physical abuse was 42% (95% CI, 19%-66%; \(I^2 = 98\%\)); and the percentage of agreement between prospective and retrospective measures of childhood physical abuse was 75% (95% CI, 62%-86%; \(I^2 = 99\%\)). Third, the prevalence of retrospective reports among those with prospective observations in the 4 studies\(^\text{16,23,25,28}\) that included childhood emotional abuse was 37% (95% CI, 23%-52%; \(I^2 = 84\%\)); the prevalence of prospective observations among those who retrospectively reported childhood emotional abuse was 15% (95% CI, 4%-33%; \(I^2 = 97\%\)); and the percentage of agreement between prospective and retrospective measures of childhood emotional abuse was 76% (95% CI, 57%-91%; \(I^2 = 99\%\)). Finally, the prevalence of retrospective reports among those with prospective observations in the 4 studies\(^\text{17,19,23,25}\) that included childhood neglect was 23% (95% CI, 14%-34%; \(I^2 = 81\%\)); the prevalence of prospective observations among those who recalled childhood neglect was 18% (95% CI, 13%-25%; \(I^2 = 61\%\)); and the percentage of agreement between prospective and retrospective measures of childhood neglect was 84% (95% CI, 70%-94%; \(I^2 = 99\%\)).

**Agreement Between Prospective and Retrospective Measures of Childhood Maltreatment**

Because the raw percentage of agreement can be inflated by chance, we next examined the agreement between prospective and retrospective measures based on the \(\kappa\), which accounts for chance findings and provides an estimate of variation in agreement in the population. A random-effects model meta-analysis of the 16 studies that included any measure of maltreatment revealed that the agreement between prospective and retrospective measures of childhood maltreatment was poor, with \(\kappa = 0.19\) (95% CI, 0.14-0.24; \(P < .001\); \(I^2 = 93\%\)). A forest plot displaying the meta-analytic findings is reported in Figure 3.

We found some evidence of publication bias, as suggested by slight asymmetry of the funnel plot (eFigure 2A in...
and-fill procedure. The trim-and-fill results with 17 studies plot asymmetry arising from publication bias, we used a trim-and-fill procedure. The trim-and-fill results with 17 studies were similar and had overlap-

In Figure 3, the Cohen $\kappa$ agreement between prospective and retrospective measures of childhood maltreatment is depicted. The forest plot shows the Cohen $\kappa$ values for each study, with the overall Cohen $\kappa$ being 0.19 (95% CI: 0.14-0.24). The diamond marker indicates the overall effect size and its variation; different sizes of markers are a function of the standard error for $\kappa$ in individual studies in the random-effects model.

### Predictors of Heterogeneity in Agreement Between Prospective and Retrospective Measures of Childhood Maltreatment

Finally, we tested predictors of heterogeneity across studies with subgroup and metaregression analyses. First, we considered whether the measure used for prospective assessment of maltreatment could explain heterogeneity in effect sizes. Agreement with retrospective reports was similar regardless of whether prospective assessment was based on records (eg, child protection records or medical records; $k = 0.16; 95\% CI, 0.09-0.24), reports (eg, questionnaires or interviews by parents or young people; $k = 0.22; 95\% CI, 0.14-0.31), or mixed measures (records and reports; $k = 0.23; 95\% CI, 0.01-0.48). An overall test of moderation showed that prospective measure type did not explain the heterogeneity in agreement ($Q = 1.1755; df = 2; P = .56$). Second, we considered whether the measure used for retrospective assessment of maltreatment could explain heterogeneity in effect sizes. As shown in Figure 4, retrospective recall during interviews (eg, verbal assessment, including reading a questionnaire aloud) showed higher $\kappa$ agreement compared to retrospective measures. In meta-regression analyses, we included interview vs questionnaire ($Q = 4.1521; df = 1; P = .04$). Diamond markers indicate overall effect size and its variation; different sizes of markers are a function of the standard error for $\kappa$ in individual studies in the random-effects model.

### Meta-analysis Stratified by the Type of Retrospective Measure Used

A formal test of moderation across type of childhood maltreatment was not possible because the subgroups were not independent (ie, different types of childhood maltreatment were measured in the same individuals). Fourth, we tested whether the type of childhood maltreatment could explain heterogeneity in effect sizes. As shown in Figure 4, recall with interview ($k = 0.34; 95\% CI, 0.23-0.45$) compared to recall with questionnaire ($k = 0.09; 95\% CI, 0.04-0.14$) showed the strongest agreement.

The results are reported in Table 4, which shows the Cohen $\kappa$ values for each study, with the overall Cohen $\kappa$ being 0.19 (95% CI: 0.14-0.24). The diamond marker indicates the overall effect size and its variation; different sizes of markers are a function of the standard error for $\kappa$ in individual studies in the random-effects model.

<table>
<thead>
<tr>
<th>Source</th>
<th>Cohen $\kappa$ (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>0.22 (0.16-0.27)</td>
</tr>
<tr>
<td>All</td>
<td>0.11 (0.06-0.16)</td>
</tr>
<tr>
<td>Overall</td>
<td>0.19 (0.14-0.24)</td>
</tr>
</tbody>
</table>

Table 4. Forest Plot Depicting the Results of a Random-Effects Meta-analysis
between prospective and retrospective measures, such as the agreement may be owing to other systematic differences between prospective and retrospective reports might detect more true cases. Third, low capture only the most severe cases of maltreatment, whereas prospective measures (eFigure 1 in the Supplement); for example, childhood maltreatment identified by prospective vs retrospective measures) (Q = 1.1653; df = 1; P = .28) or age at retrospective report (Q = 1.0561; df = 1; P = .30). Variation in study quality also did not explain heterogeneity in effect sizes (Q = 0.1632; df = 1; P = .69) (eTable 4 in the Supplement). Finally, in sensitivity analyses where we selected the highest effect size for the 7 studies reporting multiple effect sizes for different abuse types (instead of calculating the mean as above), we found similar results (eResults in the Supplement).

Discussion

This meta-analysis is the first, to our knowledge, to examine the agreement between prospective and retrospective measures of childhood maltreatment. Across 16 studies that included 25 471 individuals, we found that prospective and retrospective measures of childhood maltreatment showed poor agreement. Notably, more than half of individuals with prospective observations of childhood maltreatment did not report it retrospectively, and likewise more than half of individuals retrospectively reporting childhood maltreatment did not have concordant prospective observations (Figure 2). This finding suggests that prospective and retrospective measures of childhood maltreatment identify largely different groups of individuals and, thus, cannot be used interchangeably.

Low agreement between prospective and retrospective measures of childhood maltreatment could be explained by multiple factors, such as motivation of reporters, measurement features, and memory biases. Motivation can reduce agreement if prospective or retrospective reporters may gain something by intentionally withholding information about childhood maltreatment (ie, nondisclosure, for example owing to embarrassment, feeling uncomfortable with the interviewer, not wanting to discuss upsetting events, or fear of referral to the authorities) or by fabricating information (ie, false disclosure, for example in the context of harassment, revenge, or family disputes).

Measurement features can also reduce agreement in several ways. First, all childhood maltreatment measures have imperfect test-retest reliability, and constraints on reliability add error variance, ultimately reducing agreement between prospective and retrospective measures. Second, low agreement may be due to systematic differences in the sensitivity of the measures (as reflected by the lower prevalence of childhood maltreatment identified by prospective vs retrospective measures) (eFigure 1 in the Supplement); for example, prospective measures through official records might capture only the most severe cases of maltreatment, whereas retrospective reports might detect more true cases. Third, low agreement may be owing to other systematic differences between prospective and retrospective measures, such as the reporting period (eg, official records vs later self-reports), the reporting period (eg, prospective observation until 12 years of age vs retrospective recall of experiences from 0-18 years of age, or official records capturing maltreatment limited to early childhood owing to the focus of child protection services), or the definition of the maltreatment experience between prospective and retrospective measures (eg, neglect measured prospectively as lack of parental affection and retrospectively as lack of input or stimulation).

Finally, memory biases can reduce agreement by promoting underreporting and overreporting of actual experiences. On the one hand, underreporting may occur because of (1) deficits in encoding the maltreatment experience in early life owing to immature, delayed, or impaired brain development; (2) deficits in consolidating the maltreatment memory owing to low emotional valence (ie, not experiencing the event as distressing) or disrupted sleep patterns; (3) deficits in reconsolidating the maltreatment memory owing to memory updating during subsequent reactivation if false feedback is given (eg, because the experience was not abusive), if the memory is no longer associated with distressing emotions (eg, after successful psychotherapy), or if reappraisal is positively biased by personality features (eg, high agreeableness); (4) deficits in memory storage owing to brain injury or aging; or (5) deficits in retrieving the maltreatment memory owing to infantile amnesia, forgetting (eg, because of low contextual reinforcement or interference by competing memories), or cognitive avoidance strategies to regulate affect. On the other hand, overreporting may occur because of (6) bias in memory encoding or reconsolidation owing to individual suggestibility (as shown in experimental paradigms of imagination inflation, false feedback, or memory implantation) or a source-monitoring error (eg, misinterpretation of internal images or dreams as lived experiences) or (7) inaccurate retrieval linked to negative bias in autobiographical memory (eg, in depression).

Our findings support some of these factors. First, we found that the agreement between prospective and retrospective measures of childhood maltreatment was higher in studies that used interviews rather than questionnaires to elicit retrospective recall (Figure 4). This finding is consistent with broader observations regarding the assessment of life stress and may occur because interviews enable provision of a more detailed definition of maltreatment, contextual or anchoring methods, and greater engagement of participants. Second, agreement was also higher in studies with smaller samples (eFigure 5 in the Supplement), which might reflect the presence of more detailed retrospective assessments. Finally, the agreement for any of the childhood maltreatment measures included was substantially lower than the agreement for more clear-cut forms of adversity, such as parental loss (k = 0.83 in the study by Reuben et al; Figure 2F), suggesting that subjective interpretation of the childhood maltreatment measures may contribute to the observed heterogeneity. More research is clearly needed to disentangle factors contributing to the low agreement between prospective and retrospective measures of childhood maltreatment.

Limitations

Our findings should be interpreted in the context of some limitations. First, because of the high levels of heterogeneity, the
average meta-analytical effect sizes for agreement should be interpreted with caution. However, we used random-effects models to minimize bias linked to high heterogeneity and note that the meta-analytical CIs are narrow and consistent with the interpretation given.

Second, the results describe the agreement between prospective and retrospective measures of childhood maltreatment commonly used in the context of research studies. Therefore, the results cannot be extrapolated to infer agreement or validity of measures of childhood maltreatment used in other contexts (eg, retrospective allegations brought to the attention of the criminal justice system).

Third, although prospective measures are generally considered to be more valid (specific) indicators of the occurrence of maltreatment,\textsuperscript{52} the low agreement between prospective and retrospective measures cannot be interpreted to directly indicate poor validity of retrospective measures. For example, prospective measures may have lower sensitivity (ie, may identify a lower proportion of individuals who were maltreated), and the higher prevalence of retrospective measures could, thus, indicate greater ability to identify true cases of childhood maltreatment. If that was the case, predictions from retrospective measures should converge on the same outcomes as those of more specific prospective measures (convergent validity), and retrospective measures should not only be associated with outcomes assessed with the same method (ie, self-reports) but should also be associated with outcomes assessed with other methods, such as objective measures (eg, medical examinations or biomarkers [discriminant validity]).\textsuperscript{53} A few studies\textsuperscript{13,15,23,25} have tested these questions and have observed that prospective and retrospective measures assessed in the same individuals are associated with similar outcomes. However, retrospective measures showed stronger associations with self-reported outcomes than objectively assessed outcomes,\textsuperscript{23,31} raising concerns about potential common method bias.\textsuperscript{54} Therefore, further research in other samples is needed to comprehensively evaluate the construct validity of retrospective measures. Regardless of any concerns regarding their validity, retrospective reports may still be pragmatically used in the clinic as risk indicators associated with incidence of psychopathology, its course of illness, or treatment response.\textsuperscript{1,55}

Conclusions

Our findings have implications for researchers and health care professionals. Although retrospective reports and prospective measures identify at-risk individuals, the groups of individuals identified with either measure are not the same (Figure 2). Therefore, assuming that the underlying risk mechanisms are the same in both groups may be inaccurate. That is, the mechanisms underlying disease risk in children identified as being maltreated through prospective assessments may be different from the mechanisms underlying disease risk in adults retrospectively reporting childhood maltreatment. If risk mechanisms are different, then the 2 groups will need different interventions to effectively prevent and treat disease. As such, our findings provide a new framework for etiologic research on childhood maltreatment and intervention development.

ARTICLE INFORMATION

Accepted for Publication: December 13, 2018.
Published Online: March 20, 2019.

Author Contributions: Drs Baldwin and Danese had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Concept and design: Baldwin, Danese.

Acquisition, analysis, or interpretation of data: All authors.

Drafting of the manuscript: Baldwin, Danese.

Critical revision of the manuscript for important intellectual content: All authors.

Statistical analysis: Baldwin, Danese.

Obtained funding: Danese

Administrative, technical, or material support: Newbury, Danese.

Supervision: Danese.

Conflict of Interest Disclosures: None reported.

Funding/Sponsor: This study was supported by award U24AG047867 from the US National Institute on Aging (NIA) of the National Institutes of Health and award ES/M00919X/1 from the UK Economic and Social Research Council and the Biotechnology and Biological Sciences Research Council. Drs Baldwin and Newbury reported support from the UK Economic and Social Research Council. Mr Reuben reported support from grant AG032282 from the NIA; grant MR/P005918 from the UK Medical Research Council; and grant F31ES029358 from the US National Institute of Environmental Health Sciences. Dr Danese reported funding from the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King’s College London.

Role of the Funder/Sponsor: The funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Disclaimer: The content is solely the responsibility of the authors and does not necessarily represent the official views of the NHS, the NIHR, or the Department of Health and Social Care.

REFERENCES


Agreement Between Prospective and Retrospective Measures of Childhood Maltreatment


http://www.jama-psychiatry.com/content/early/2019/02/10/jamapsychiatry.2019.02.10.E10

Agreement Between Prospective and Retrospective Measures of Childhood Maltreatment