Identifying adverse childhood experiences with electronic health records of linked mothers and children in England: a multistage development and validation study

Shabeer Syed, Arturo Gonzalez-Izquierdo, Janice Allister, Gene Feder, Leah Li, Ruth Gilbert

Summary

Background Electronic health records (EHRs) of mothers and children provide an opportunity to identify adverse childhood experiences (ACEs) during crucial periods of childhood development, yet well developed indicators of ACEs remain scarce. We aimed to develop clinically relevant indicators of ACEs for linked EHRs of mothers and children using a multistage prediction model of child maltreatment and maternal intimate partner violence (IPV).

Methods In this multistage development and validation study, we developed a representative population-based birth cohort of mothers and children in England, followed from up to 2 years before birth to up to 5 years after birth across the Clinical Practice Research Datalink (CPRD) GOLD (primary care), Hospital Episode Statistics (secondary care), and the Office for National Statistics mortality register. We included livebirths in England between July 1, 2004, and June 30, 2016, to mothers aged 16–55 years, who had registered with a general practitioner (GP) that met CPRD quality standards before 21 weeks of gestation. The primary outcome (reference standard) was any child maltreatment or maternal IPV in either the mother’s or child’s record from 2 years before birth (maternal IPV only) to 5 years after birth. We used seven prediction models, combined with expert ratings, to systematically develop indicators. We validated the final indicators by integrating results from machine learning models, survival analyses, and clustering analyses in the validation cohort.

Findings We included data collected between July 1, 2002, and June 27, 2018. Of 376006 eligible births, we included 211393 mother-child pairs (422786 patients) from 400 practices, of whom 126837 mother-child pairs (60.0%; 240 practices) were randomly assigned to a derivation cohort and 84556 pairs (40.0%; 160 practices) to a validation cohort. We included 63 indicators in six ACE domains: maternal mental health problems, maternal substance misuse, adverse family environments, child maltreatment, maternal IPV, and high-risk presentations of child maltreatment. Excluding the seven indicators in the reference standard, 56 indicators showed high discriminative validity for the reference standard of any child maltreatment or maternal IPV between 2 years before and 5 years after birth (validation cohort, area under the receiver operating characteristic curve 0.85 [95% CI 0.84–0.86]). During the 2 years before birth and 5 years after birth, the overall period prevalence of maternal IPV and child maltreatment (reference standard) was 2.3% (2876 of 126837 pairs) in the derivation cohort and 2.3% (1916 of 84556 pairs) in the validation cohort. During the 2 years before and after birth, the period prevalence was 39.1% (95% CI 38.7–39.5; 34773 pairs) for any of the 63 ACE indicators, 22.2% (21.8–22.5%; 20122 pairs) for maternal mental health problems, 15.7% (15.4–16.0%; 14549 pairs) for adverse family environments, 8.1% (7.8–8.3%; 6808 pairs) for high-risk presentations of child maltreatment, 6.9% (6.7–7.2%; 7856 pairs) for maternal substance misuse, and 3.0% (2.9–3.2%; 2540 pairs) for any child maltreatment (2.4% [2.3–5.6%; 2051 pairs]) and maternal IPV (1.0% [0.8–1.0%; 875 pairs]). 62.6% (21785 of 34773 pairs) of ACEs were recorded in primary care only, and 72.3% (25140 cases) were recorded in the maternal record only.

Interpretation We developed clinically relevant indicators for identifying ACEs using the EHRs of mothers and children presenting to general practices and hospital admissions. Over 70% of ACEs were identified via maternal records and were recorded in primary care by GPs within 2 years of birth, reinforcing the importance of reviewing parental and carer records to inform clinical responses to children. ACE indicators can contribute to longitudinal surveillance informing public health policy and resource allocation. Further evaluation is required to determine how ACE indicators can be used in clinical practice.

Funding None.

Copyright © 2022 The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY 4.0 license.

Introduction Adverse childhood experiences (ACEs) are potentially traumatic or neglectful experiences that can profoundly affect the health and development of children. Adverse childhood experiences often co-occur in the family, ranging from child maltreatment and neglect, to growing up in a household
with parental mental health problems, and maternal intimate partner violence (IPV). Failure to identify and measure ACEs can substantially undermine opportunities to act, and can place children at risk of considerable harm in the longer term. Children with ACEs face increased risks of hospital admissions, chronic conditions, teenage pregnancies, suicide, and inter-generational violence.

In the UK and the USA, national prevention strategies expanded their aims in 2020 to reduce ACEs as a precursor to long-term health problems. However, to achieve this goal, data systems for identifying ACEs and vulnerable families must first be carefully developed and quantified. The absence of validated indicators for identifying ACEs in routine electronic health records (EHRs) represents an important obstacle. Policy makers and services often rely on small self-report studies, local samples, and non-validated methods for identifying ACEs and vulnerable families. The few longitudinal studies in the UK that have measured ACEs in EHRs were based on individual data sources in specific populations, or on family members in isolation (eg, maternal mental health problems; appendix pp 45–46). Studies to date have also differed in their approach to defining ACEs or have combined multiple risk groups that are difficult to disentangle.

In the UK, first-time mothers are offered 13 antenatal appointments and between one and three postnatal primary care appointments. Children routinely attend primary care for surveillance and vaccinations from 6–8 weeks to 4–6 years after birth. The ability to link EHRs of mothers and children provides an opportunity to identify ACEs during crucial periods of prevention of long-term harm. This study aimed to develop clinically relevant indicators of ACEs in linked EHRs, with the potential to aid longitudinal public health monitoring and prompt early assessment of support needs for vulnerable children and mothers before and after birth. We used a multistage prediction framework of child maltreatment and maternal IPV to develop indicators. First, we identified candidate ACE indicators using a systematic review. Then, we assessed the relevance of indicators against a reference standard of child maltreatment or maternal IPV using an integrative predictive and explanatory approach, combining risk estimates from multiple variable selection models, cluster analyses, and expert ratings. We used a large representative birth cohort of mothers and children in England followed from 2 years before birth up to 5 years after birth across primary and secondary care. Indicator selections and predictions of child maltreatment or maternal IPV were theoretically informed (appendix p 3).

Methods

Study design and participants

This study followed the guidelines for accurate and transparent health estimates reporting, and the reporting guidelines for prediction model and development (appendix pp 47–54). We derived a population-based birth cohort using the mother–baby link in the Clinical Practice Research Datalink (CPRD) GOLD database via the CALIBER platform (appendix p 4). The CPRD GOLD database contains anonymised primary care data from...
approximately 6.9% of UK general practices, with recorded symptoms, diagnoses, prescriptions, and referrals to secondary care. CPRD GOLD is broadly representative of the general population. Mothers and children are linked in the CPRD via unique household identifiers and maternity records, matched with high validity. In 2004, the mother–baby link contained data from 676 practices, with 400 practices (59%) consenting for linkage to other data sources. Most practices (329 [82%] of 400 general practices) contributed data for at least 10 years or until 2014.

Mothers and children were linked to the Hospital Episodes Statistics Admitted Patient Care database (HES-APC), the mortality register from the Office for National Statistics, and the Index of Multiple Deprivation 2015. The HES-APC provides data on hospital admissions, including birth or delivery records, from all hospitals in England funded by the National Health Service (NHS). The Office for National Statistics includes data on cause-specific mortality from registered death certificates. The Index of Multiple Deprivation 2015 is the official national specific mortality from registered death certificates. The Office for National Statistics includes data on cause-specific mortality from registered death certificates. The Index of Multiple Deprivation 2015 is the official national metric for relative deprivation in England and comprises a composite score across seven domains based on a patient’s postcode (eg, average income, employment status, or crime levels). This index can be classified into five quantiles, from the least deprived to the most deprived. Detailed study procedures and information on data sources are provided in the appendix (pp 4–5).

We included livebirths in England between July 1, 2004, and June 30, 2016, to mothers aged 16–55 years, who had registered with a general practitioner (GP) that met CPRD GOLD quality standards before 21 weeks of gestation. We selected a random child per mother to avoid clustering effects due to multiple children, given that several variable selection models could not computationally handle clustering on the secure analytical server (random-effects forest >168 h; appendix pp 21–22). Children had to be registered with the practice within 6 months after birth, with follow-up data collected until the child’s first birthday (figure 1). We calculated conception dates using validated algorithms. The 20-week minimum cutoff before birth avoided exclusion of extreme preterm births and captured pre-birth data for the full cohort. This study was approved by the UK Medicines and Healthcare products Regulatory Agency Independent Scientific Advisory Committee (19_162R), under Section 251 (NHS Social Care Act 2006), for the use of anonymised records in research without individual participant written consent.

### Data selection

We included data collected between July 1, 2002, and June 27, 2018, with births starting from July 1, 2004, to June 30, 2016. This period allowed for 2 years of follow-up before and after birth. Follow-up for the primary analyses ranged from a minimum of 20 weeks of gestation and 1 year following birth, to a maximum of 2 years before birth to 5 years after birth, and included practice deregistration, the last data collection date of the practice, and death or the study end date, whichever came first (appendix p 4). We restricted the period for ACEs to the first 5 years after birth to minimise attrition (eg, practice deregistration) and to avoid ACEs related to events outside the immediate family (eg, teenage peer relationships). This period is consistent with prioritised care interventions for mothers and children.

### Indicators and domains of ACEs

We developed two measures of ACEs for EHRs: indicators (ie, grouped codes or measures) and domains (ie, grouped indicators; table 1). Definitions, selection procedures, and excluded candidate indicators are provided in the appendix (pp 6–31). Briefly, we identified 408 candidate indicators for ACEs using systematic reviews of code lists and previous studies. We used predefined criteria from Public Health England, WHO, and previous reviews to classify indicators of ACEs (appendix pp 6–9). We manually grouped candidate indicators into six broader ACE domains consistent with the original ACE study. We added the domain, high-risk presentations of child maltreatment. This domain encompassed indicators based on guidance from the National Institute for Health and Care Excellence and the Royal College of General Practitioners, which addressed presentations that should raise concerns for child maltreatment.

We defined candidate indicators by combining information from the EHRs of both mothers and children across all sources (eg, Read codes, International Classification of Diseases 9th or 10th edition, prescriptions, or self-report measures). We treated mothers and children with no indicator as unexposed. We used multiple rule-based algorithms to prevent misclassification of specific indicators as due to other causes (eg, accidents; appendix pp 16–17). The complete code lists are available on the ACEs in EHRs website.

We assessed the relevance of candidate indicators on the basis of their association with a clinically defined reference standard in a multistage prediction model. We expected the selected indicators to reflect a continuum of clinical relevance, ranging from high relevance to low relevance.

To add covariates to the prediction models, we included variables on demographics, deprivation quantiles, and clinical characteristics recorded during pregnancy or up to 2 years before birth. We ascertained birthweight, gestational age, maternal age, and congenital anomalies (European Registry of Congenital Anomalies and Twins guidelines) from both the mothers’ and children’s records in the CPRD and the Hospital Episodes Statistics database.

### Outcomes

The primary outcome (ie, reference standard) was any recorded child maltreatment or maternal IPV in the
Validation of ACE indicators based on a birth cohort approach.

**Phase 1: Identifying Indicators**

- Screening of ACE indicators and initial classification of codes and measures
- Indicators excluded if rated irrelevant or if of lower predictive value for child maltreatment
- Internal expert review
- 408 candidate ACE indicators with mapped codes and measures
- Reference standard selected: child maltreatment and maternal IPV (seven indicators)

**Phase 2: Birth Cohort and Indicator Selections**

- Birth cohort selection: 576,000 births from 2004 to 2016 recorded in the mother-baby link in the CPRD-GOLD, eligible for linkage across the Hospital Episodes Statistics, index of multiple deprivation, and the Office for National Statistics
- Linked mother-child data sources: Primary care (CPRD), hospital admissions (Hospital Episode Statistics for admitted patient care), death registry (Office for National Statistics), index of multiple deprivation
- 164,651 mother-child pairs included; 71,226 randomly excluded siblings; 55,963 registered <20 weeks before birth; 37,444 children registered late or <1 year of follow-up
- 211,393 mother-child pairs included (400 general practices)

**Validation Cohort**

- Random assignment of general practices
- 84,556 (40.0%) mother-child pairs (160 general practices) included in the validation cohort
- 126,837 (60.0%) mother-child pairs (240 general practices) included in the derivation cohort
- 166 indicators and 12 covariates

**Selection via Ranking Index**

- Indicators and covariates entered into seven variable selection models to predict child maltreatment and maternal IPV
- Variables: Child maltreatment and maternal IPV variable selection using random forests (thresholding step), Boruta algorithm, recursive feature selection, random survival forest, Cox and logistic regression with LASSO, and backward selection in the Cox model

**Selection of ACE Indicators**

- 408 candidate ACE indicators with mapped codes and measures
- 31 tentative indicators excluded following expert ratings and thresholding against minimum positive predictive value
- 41 indicators and seven covariates excluded
- 71 indicators retained on the basis of model consistency over decisions or rankings
- 54 tentative indicators reviewed by experts
- 31 tentative indicators excluded following expert ratings and thresholding against minimum positive predictive value
- 36 indicators retained and merged into 56 final indicators

**Phase 3: Validation and Sensitivity Analyses**

- Final selected ACEs
  - 56 indicators (63 with reference standard)
- Analyses in the validation cohort
- Aggregate hierarchal clustering analyses
- Ten-fold repeated cross-validation balanced random forest
- Adjusted weighted Cox proportional hazard models
- Adjusted weighted logistic models with marginal predictions

**Validation Metrics of ACEs**

- Data-driven clustering of indicators based on distance metrics
- Area under the curve of receiver operator characteristics, sensitivity, and specificity (positive and negative predictive value)
- Balanced random forest variable importance of any child maltreatment or maternal IPV
- Probability of any child maltreatment or maternal IPV over time
- Weighted adjusted period prevalences of ACE domains

Figure 1: Overview of the development and validation of ACE indicators.

ACE = adverse childhood experience. CPRD = Clinical Practice Research Datalink. GP = general practitioner. IPV = intimate partner violence. LASSO = Least absolute shrinkage and selection operator. UCLA = University of California at Los Angeles.
child’s or the mother’s record from 2 years before birth (maternal IPV only) to 5 years after birth. Child maltreatment included neglect,* harm caused during pregnancy (eg, neonatal abstinence syndrome),19–25 deaths related to child maltreatment, and social service referrals for children (appendix pp 8–9, 24–25).19–25 In the absence of a consensus reference standard for ACEs, child maltreatment and maternal IPV provide a clinically important outcome measured in childhood,26–27 and represent a probable cumulation of underlying adversity (appendix pp 1).26–27 Because GPs do not record ACEs at every presentation that raises concern,28–30 indicators were not required to occur before the outcome (excluded in sensitivity analyses; appendix p 42). The seven indicators of child maltreatment or maternal IPV in the reference standard were excluded from the risk prediction models, but were included for estimating prevalences and for clustering analyses.

**Statistical analysis**

Detailed statistical methods are provided in the appendix (pp 21–22, 26–27). Briefly, we randomly assigned 60% of general practices to a derivation cohort for model development (appendix p 4). The remaining 40% of general practices were assigned to a validation cohort. Some indicators were used in their existing form, whereas indicators with less than 100 unique records were reclassified into neighbouring indicators for statistical power. Continuous variables (eg, alcohol units) were dichotomised on the basis of validated higher risk cutoff scores (appendix pp 16–17). We used pairwise correlation matrices, network plots, and adjusted Cox proportional hazards models with inverse probability weighting to reclassify indicators (appendix pp 21–22).31–35

Having preprocessed 166 candidate indicators, we established the relevance of each indicator by combining results from multiple variable selection models of child maltreatment and maternal IPV into a ranking index (appendix pp 26–27). We followed the procedures of Haq and colleagues36 to rank indicators in descending order of risk association for each model (ie, greatest risk association was ranked first for highest relevance).36 The median (IQR) ranking of each indicator acted as a summary measure of relevance. Given the aims of this study to establish indicators that reflected a continuum of clinically important outcome measured in childhood,26–27 and emotional or psychological abuse,*

| ACE domains and final indicators | Number of codes |
|---------------------------------|-----------------|
| CM (reference standard)*        |                 |
| CM1*   | Child protection or safeguarding* | 50 |
| CM2*   | CM not otherwise specified, including physical or sexual abuse (merged)* | 154 |
| CM3*   | Neglect (including neonatal abstinence syndrome or fetal alcohol spectrum disorders), and emotional or psychological abuse* | 76 |
| CM4*   | Social service involved (including parental imprisonment or criminal activity)* | 80 |
| CM5*   | Child in care* | 107 |
| Suspected CM† | | |
| CM6     | Suspected CM, not otherwise specified (including neglect and social service involvements) | 244 |
| CM7     | Child assaulted, not otherwise specified (including physical or sexual abuse [≥10], rib fractures [≥3]) | 545 |

| Maternal IPV (reference standard)* | |
|---------------------------------|---|
| Maternal IPV1* Maternal IPV, not otherwise specified (including physical or sexual abuse)* | 67 |
| Maternal IPV2* Mother assaulted plus child protection recording or incident during pregnancy* | 554 |
| Suspected maternal IPV† | |
| Maternal IPV3 Suspected maternal IPV, not otherwise specified | 33 |
| Maternal IPV4 Suspected maternal IPV, physical or sexual abuse | 45 |
| Maternal IPV5 Mother assaulted, not otherwise specified (hospital admission only) | 119 |
| Maternal IPV6 Mother assaulted plus high-risk presentations (algorithm)* | 236 |

| HRP-CM | |
|--------|---|
| Child injuries | |
| HRP-CM1 | Bruising and contusions (≥3)† | 114 |
| HRP-CM2 | Superficial injuries of head, neck, or multiple body parts (≥3)‡ | 37 |
| HRP-CM3 | Thermal injuries: head, face, or neck (≥3)‡ | 161 |
| HRP-CM4 | Thermal injuries: trunk, back, or trachea (≥3)‡ | 53 |
| HRP-CM5 | Skull fractures or intracranial crush injury (≥3)‡ | 16 |

| Harm by undetermined intent | |
|-----------------------------|---|
| HRP-CM6 Child harm by undetermined intent: rare injuries and life-threatening events (eg, retinal haemorrhages, drownings, sudden unexpected death in infancy, or firearm injuries [≤10])† | 239 |
| HRP-CM7 Child harm by undetermined intent: exposure to unspecified factor (≥10)† | 4 |

| Potential failure to provide | |
|-----------------------------|---|
| HRP-CM8 Failure to thrive (eg, excessive thirst or suspected malnutrition [≥10])‡ | 48 |
| HRP-CM9 Non-attendance of child appointments (≥3) appointments within 2 years [≥10])‡ | 16 |

| MSM | |
|-----|---|
| MSM1 Severe drug misuse (dependence) | 564 |
| MSM2 Moderate drug misuse (all other) | 213 |
| MSM3 Maternal drug prescription for opioid dependence (multipurpose usage) | 21 |
| MSM4 Family substance misuse (ie, unspecified family member) | 19 |
| MSM5 Severe alcohol misuse (including self-report measures of ≥35 alcohol units per week)‡ | 273 |

| AFEs | |
|------|---|
| Antenatal care and health visit concerns | |
| AFE1 High-risk antenatal presentation: specific to social risk | 2 |
| AFE2 High-risk antenatal presentation: psychosocial risk, not otherwise specified | 38 |
| AFE3 Unwanted or concealed pregnancy (including attempted abortion of current child) | 46 |
| AFE4 Psychosocial health problem with lower-level intervention | 20 |
| AFE5 Increasing concern of health visitor | 11 |

(Tables 1 continues on next page)
**(Continued from previous page)**

### ACE domains and final indicators

| Parental conflicts, disruptions, and causes for concerns | Number of codes |
|----------------------------------------------------------|-----------------|
| AFE6 Family disruptions and parental conflicts, not otherwise specified | 108 |
| AFE7 Parental separations | 27 |
| AFE8 Mother with legal problems | 32 |
| AFE9 Family is cause for concern† | 182 |
| AFE10 Problems related to negative childhood events | 26 |
| AFE11 Mother assaulted, not otherwise specified (GP record only) | 1 |

### Vulnerable families

| AFE12 Housing problems, effects of deprivation, and refugees (excluding homelessness) | 57 |
| AFE13 Homelessness (child or mother) | 22 |
| AFE14 Vulnerable family, not otherwise specified (including care programme approach) | 31 |
| AFE15 Family or parental support referral | 12 |
| AFE16 Problems related to psychosocial circumstances | 24 |
| AFE17 Maternal learning or intellectual disability | 276 |
| AFE18 Increased concerns of maternal incapacity | 10 |
| AFE19 Maternal problems with daily living or limited capacity to work (including financial concerns) | 41 |

### Maternal MHPs

| Common MHPs | Number |
|-------------|--------|
| Maternal MHP1 Depression (including use of antidepressants)† | 818 |
| Maternal MHP2 Self-harm or suicide attempts | 744 |
| Maternal MHP3 Anxiety disorder, not otherwise specified (including use of anxiolytics)† | 549 |
| Maternal MHP4 Panic disorder (including agoraphobia or health anxiety) | 24 |
| Maternal MHP5 Obsessive-compulsive disorders | 27 |
| Maternal MHP6 Post-traumatic stress disorder (including acute stress disorder) | 72 |
| Maternal MHP7 Sleep-wake disorders | 33 |
| Maternal MHP8 MHPs not otherwise specified | 17 |
| Maternal MHP9 Referred to or seen by a mental health professional (tier 3 service or above) | 180 |
| Maternal MHP10 Puerperal MHPs, not otherwise specified | 5 |

### Eating disorders

| Maternal MHP11 Anorexia nervosa | 13 |
| Maternal MHP12 Eating disorders, not otherwise specified (including bulimia) | 49 |

### Psychosis and personality disorders

| Maternal MHP13 Psychosis (including mental health sections not otherwise specified) | 339 |
| Maternal MHP14 Use of antipsychotics | 324 |
| Maternal MHP15 Bipolar disorders | 66 |
| Maternal MHP16 Personality disorders (eg, borderline personality disorder) | 177 |

### Neurodevelopmental disorders

| Maternal MHP17 Neurodevelopmental conditions and conduct disorders | 245 |

Details regarding ascertainment of indicators are provided in the appendix (pp 10–31). ACE=adverse childhood experience. CM=child maltreatment. IPV=intimate partner violence. HRP=high-risk presentation. MSM=maternal substance misuse. AFEs=adverse family environments. MHP=mental health problem. *Indicators were combined into the primary outcome (reference standard) and excluded from the development and validation phase. †Suspected CM and suspected maternal IPV were subdomains containing less specific maltreatment-related indicators used in the development process to expand the final CM and maternal IPV domains, respectively. ‡Indicators are defined by multiple rule-based algorithms, including age restrictions in years (upper age cutoff denoted in brackets), exclusions of accidental injuries, genetic predispositions (eg, bone diseases), traumatic birth injuries, transmissions of diseases from mother to child during birth, or need to meet higher cutoff score on a validated self-report instrument. Medications, interventions, and psychiatric symptoms were combined into appropriate disorder clusters using validated algorithms. §Neurodevelopmental disorders are included as a diagnostic cluster in the Diagnostic and Statistical Manual of Mental Disorders (5th edn) and International Classification of Diseases (10th or 11th edn).

**Table 1**: Final six ACE domains and 63 indicators included in validation analyses

---

Algorithm, logistic and Cox regression models with least absolute shrinkage and selection operator (Harrell C index), and a Cox model with stepwise backward variable selection (Akaike’s information criterion). We included indicators consistently retained across models and excluded consistently omitted indicators (Fleiss’ kappa agreement statistic ≥0.6). We used a ten-fold repeated cross-validation for variable selection models (appendix pp 21–22, 26–27). All machine learning models were balanced by randomly downsampling the derivation cohort to minimise overfitting by the imbalanced number of reference standard cases compared with non-cases.

This process left tentative indicators with a median ranking in between the two thresholds. Four experts on family violence (RG, LL, GF, and JA) independently rated tentative indicators (plus 50% already included or excluded) on the basis of clinical credibility and relevance criteria on a scale of 1–10 (appendix pp 33–37). The expert panel were masked to the decisions of the variable selection model. We retained indicators that were consistently rated by experts to be at least five on the relevance and credibility scale (Fleiss’ kappa statistic ≥0.6). As we retained a large amount of indicators in the final stage, we combined indicators with less than 150 observations into neighbouring indicators by repeating the initial reclassification step, before applying them to the validation cohort (original indicators shown in the appendix pp 11–16).

We validated indicators in four ways in the validation cohort. First, we confirmed the manually grouped indicators of ACE domains by entering them into an agglomerative hierarchical clustering model based on Jaccard’s similarity index. This bottom-up approach systematically clustered indicator pairs into larger domains without pre-specification. Second, we assessed the accuracy and predictive value of ACEs for identifying the reference standard without covariates. We used a weighted-balanced random-forest model, trained in the derivation cohort with ten-fold repeated cross-validation to model predictions in the validation cohort. The random-forest model was built with a minimum of 1000 trees and 30 observations to attempt splits. To gauge predictive performance, we computed receiver operating characteristic (ROC) curves and the area under the ROC curve (AUC). We used Delong’s method to compute 95% CIs for the AUC. We examined agreement between predicted and observed probabilities over risk deciles of any child maltreatment or maternal IPV by plotting calibration curves and calculating the Brier score (appendix p 39).

Third, as expected by the cumulative stress model, we established if there was a dose–response relationship between the number of ACE domains and the reference standard. We used inverse probability-weighted and adjusted Cox proportional hazard models and Kaplan-Meier curves to estimate differences in overall
probability of child maltreatment or maternal IPV by ACE domain and by number of ACEs over time,\(^4\) compared with non-exposed mother–child pairs. We checked model assumptions of Kaplan-Meier estimates using log–log plots and the link test.\(^6\) We also provided the cross-validated random-forest model permutation importance values. These values are defined as the mean decrease in the model’s overall predictive ability when data for an indicator is randomly shuffled—i.e., how much the model depends on the indicator. The scaled permutation importance values scores range from 0% (not important) to 100% (important). Finally, to aid with the comparison of external estimates,\(^20\) we used inverse probability-weighted logistic regression models (adjusted for birth year) to compute the period prevalence of ACEs between 2 years before and after birth. This time interval is consistent with period prevalences from previous studies,\(^19\) and ensured that most mother–child pairs at baseline could be included in the denominator.

We did sensitivity analyses in the validation cohort to test the robustness of the final indicators, including extending the exposure period, cohort inclusion criteria (e.g., siblings or birth years), excluding patients with outcomes recorded after the indicator, comparing birth years (2004 vs 2014), and comparing general practices’ years of data contribution (ending before 2014 vs ending after 2014).

There were no missing data for ACE indicators because we assumed that children with no event were unexposed. Data were missing for birth characteristics of covariates on parity, gestational age, birthweight, and social deprivation obtained during pregnancy or up to 2 years before birth (table 1). We imputed missing values separately for each cohort under the missing-at-random assumption. Predictors in the model included all maternal and birth characteristics listed in table 1, any suspected child maltreatment or maternal IPV, maternal ACEs (maternal substance misuse and maternal mental health problems) based on validated indicators, and the reference standard of child maltreatment or maternal IPV (appendix pp 21–22). We used the multivariate imputation by chained equations package in R to create 25 imputed datasets (25 iterations for each imputation) and the sjmisc::merge_imputations function to combine estimates from imputed datasets.

We did all analyses on University College London’s secure analytical server (the Data Safe Haven; certified to ISO27001 information security standards), using Stata (version 16) and R (version 4.1) with the caret,\(^85\) ranger, hmisc, rms, mice, glmnet, pROC, recipes, meta, rmda, and tidyverse packages (complete list available online).

Role of the funding source
The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results
Of 376 006 eligible births between July 1, 2004, and June 30, 2016, we included 211 393 mother–child pairs (422 786 patients), with mothers followed up from up to 2 years before birth (median follow-up 2.0 years [IQR 2.0–2.0]) and mother–child pairs followed up to 5 years after birth (median 4.5 years [2.5–5.0]; figure I). We

| Maternal characteristics | Overall cohort (n=211 393) | Derivation cohort (n=126 837) | Validation cohort (n=84 556) |
|--------------------------|--------------------------|-----------------------------|-----------------------------|
| Follow-up time before birth, years | 2.0 (2.0–2.0) | 2.0 (2.0–2.0) | 2.0 (2.0–2.0) |
| Available follow-up time after birth, years | 4.5 (2.5–5.0) | 4.5 (2.5–5.0) | 4.5 (2.5–5.0) |
| Age at birth, years* | <19 | 7054 (3.3%) | 4243 (3.3%) | 2821 (3.3%) |
| | 20–39 | 147 234 (69.6%) | 88 489 (69.8%) | 58 746 (69.5%) |
| | ≥40 | 57 105 (27.0%) | 34 105 (26.9%) | 23 000 (27.2%) |
| Maternal parity* | 0 | 65 214 (30.8%) | 39 167 (30.9%) | 26 047 (30.8%) |
| | 1–3 | 83 17 (39.3%) | 49 891 (39.3%) | 32 284 (34.9%) |
| | ≥4* | 8259 (3.9%) | 4935 (3.9%) | 3234 (3.9%) |
| Missing data | 54 745 (25.9%) | 32 844 (25.9%) | 21 901 (25.9%) |
| Socioeconomic status quintile* | 1 (least deprived) | 46 559 (22.0%) | 27 873 (22.0%) | 18 666 (22.1%) |
| | 2 | 42 695 (20.2%) | 25 415 (20.0%) | 17 280 (20.4%) |
| | 3 | 41 306 (19.5%) | 24 873 (19.6%) | 16 433 (19.4%) |
| | 4 | 41 634 (19.7%) | 24 986 (19.7%) | 16 648 (19.7%) |
| | 5 (most deprived) | 39 066 (18.5%) | 23 593 (18.6%) | 15 473 (18.3%) |
| Missing data | 153 (0.1%) | 97 (0.1%) | 56 (0.1%) |
| Ethnicity | White | 17 836 (84.4%) | 10 706 (84.4%) | 7 138 (84.4%) |
| | Asian | 14 070 (7.0%) | 8 791 (6.9%) | 5 913 (7.0%) |
| | Black | 7 656 (3.6%) | 4 631 (3.7%) | 3 025 (3.6%) |
| | Other | 2 662 (1.3%) | 1 591 (1.3%) | 1 071 (1.3%) |
| | Mixed | 3 248 (1.5%) | 2 003 (1.6%) | 1 245 (1.5%) |
| | Missing data | 4 723 (2.2%) | 2 813 (2.2%) | 1 920 (2.3%) |
| Location of general practice (region of England, UK) | London | 37 999 (18.0%) | 22 780 (18.0%) | 15 219 (18.0%) |
| | Northeast, northwest, and Yorkshire | 39 387 (18.6%) | 23 681 (18.7%) | 15 706 (18.6%) |
| | East and west Midlands | 27 418 (13.0%) | 16 473 (13.0%) | 10 945 (12.9%) |
| | East | 24 713 (11.4%) | 14 522 (11.4%) | 9 651 (11.4%) |
| | Southeast, southwest, and south central | 82 416 (39.0%) | 49 381 (38.9%) | 33 035 (39.1%) |

Health comorbidities (Global Burden of Disease classification scheme)*

| Cardiovascular and circulatory diseases | 25 874 (12.2%) | 15 612 (12.3%) | 10 262 (12.1%) |
| Chronic respiratory diseases | 24 101 (11.4%) | 14 551 (11.5%) | 9 550 (11.3%) |
| Diabetes and endocrine diseases | 22 352 (10.6%) | 13 499 (10.6%) | 8 883 (10.5%) |
| Musculoskeletal disorders | 10 514 (5.0%) | 6 331 (5.0%) | 4 138 (4.9%) |
| Neurological disorders* | 1 377 (0.7%) | 855 (0.7%) | 522 (0.6%) |
| Lower respiratory infections and other infections* | 22 261 (10.5%) | 13 481 (10.6%) | 8 780 (10.4%) |
| Maternal disorders$ | 24 101 (11.4%) | 14 551 (11.5%) | 9 550 (11.3%) |

(continued)
### Articles

**Cohort characteristics**

- **Other infections included diarrhoeal diseases, malaria, meningitis, and sexually transmitted infections.**
- **Maternal**
- **Data** are median (IQR) or n (%). GP = general practitioner. *Denotes covariates included across risk prediction models for sex of child.

#### Delivery characteristics

**Time to GP registration after birth, months**

| Time          | Overall cohort (n=211,393) | Derivation cohort (n=126,837) | Validation cohort (n=84,556) |
|---------------|----------------------------|-------------------------------|-----------------------------|
| <2 months     | 197,555 (93.5%)            | 118,549 (93.5%)              | 79,066 (93.4%)             |
| 2–3 months    | 9405 (4.4%)                | 5622 (4.4%)                  | 3783 (4.5%)                |
| 3–6 months    | 4433 (2.1%)                | 2666 (2.1%)                  | 1767 (2.1%)                |

**Sex of child**

|            | Male                      | Female                     |
|------------|---------------------------|----------------------------|
| Male       | 108,221 (51.2%)           | 64,962 (51.2%)             | 43,259 (51.2%)             |
| Female     | 103,172 (48.8%)           | 61,875 (48.8%)             | 41,297 (48.8%)             |

**Multiple pregnancy**

|            | Singleton                 | Multiple (eg, twins)       |
|------------|---------------------------|----------------------------|
| Singleton  | 201,261 (95.2%)           | 120,766 (95.2%)            | 80,495 (95.2%)             |
| Multiple   | 3244 (1.5%)               | 1960 (1.5%)                | 1284 (1.5%)                |

**Missing data**

|            | Missing                   |
|------------|---------------------------|
| Missing    | 6289 (2.9%)               | 4071 (2.9%)                | 2758 (2.9%)                |

**Gestational age at birth, weeks**

|            | ≥37                       | <37                        |
|------------|---------------------------|----------------------------|
| ≥37        | 165,535 (78.3%)           | 99,338 (78.3%)             | 66,197 (78.3%)             |
| <37        | 13,232 (6.3%)             | 7,900 (6.2%)               | 5,332 (6.3%)               |

**Birthweight, g**

|            | ≥2500                     | <2500                      |
|------------|---------------------------|---------------------------|
| ≥2500      | 80,992 (38.3%)            | 48,604 (38.3%)            | 32,388 (38.3%)             |
| <2500      | 95,749 (45.3%)            | 57,477 (45.3%)            | 38,272 (45.3%)             |

**Missing data**

|            | Missing                   |
|------------|---------------------------|
| Missing    | 32,626 (15.4%)            | 19,599 (15.5%)             | 13,027 (15.4%)             |

**Birth year**

|            | 2004–10                   | 2011–18                    |
|------------|---------------------------|---------------------------|
| 2004–10    | 237,216 (64.9%)           | 74,177 (35.1%)             | 44,629 (35.2%)             | 29,548 (34.9%) |
| 2011–18    | 82,208 (64.8%)            | 45,079 (35.2%)             | 29,548 (34.9%)             |

---

Data are median (IQR) or n (%). GP = general practitioner. *Denotes covariates included across risk prediction models for the selection of indicators. †Neurological disorders included epilepsy, multiple sclerosis, and motor neuron disease. §Other infections included diarrhoeal diseases, malaria, meningitis, and sexually transmitted infections. ¶Maternal mental health problems. ¶¶Child maltreatment, maternal IPV, and high-risk presentations of child maltreatment. In the dendrogram of hierarchically clustered indicators, including indicators as part of the reference standard, a few individual indicators clustered outside their originally grouped clusters (appendix p 33). We made no changes to the original domains, given that underlying coding descriptions of outlying indicators matched originally grouped domains better conceptually. For instance, compared with other maternal mental health problems, post-traumatic stress disorder clustered closer to violence-related indicators but was retained under maternal mental health problems.

Table 3 shows validation estimates of the grouped final indicators (domains) in the validation cohort (calibration curves and ROC curves are shown in the appendix [pp 39–40]). The full model involving all 56 ACE indicators showed a good balance between sensitivity (72%) and specificity (84%) for identifying the reference standard of any child maltreatment or maternal IPV (AUC 0.85 [95% CI 0.84–0.86]; table 3). The model also showed good agreement between predicted and observed probabilities of any child maltreatment or maternal IPV over the range of risk deciles (calibration intercept 0, slope 1, Brier score 0.137).

Individual ACE domains showed a consistent dose–response relationship with any child maltreatment or maternal IPV (figure 3A). More closely related indicators of the reference standard (eg, suspected child maltreatment or maternal IPV) showed high specificity (range 99%–100%), positive predictive values (41–43%), and good calibration (Brier score 0.19), but lower overall ROC AUC (range 0.55–0.64). By contrast, broader domains (eg, maternal mental health problems, adverse family environments, and maternal substance misuse) showed higher overall AUC values (range 0.67–0.72), but included 400 general practices (median 500 pairs per practice [IQR 247–720]). We randomly assigned 240 general practices (126,837 [60%-0%] mother–child pairs) to the derivation cohort and 160 practices (84,556 [40%-0%] pairs) to the validation cohort (figure 1). The derivation and validation cohorts were similar across key child and maternal characteristics (table 2). For the main follow-up period (from 2 years before birth to 5 years after birth), the overall period prevalence of the reference standard of maternal IPV and child maltreatment was 2.3% (2,876 of 126,837 pairs) in the derivation cohort and 2.3% in the validation cohort (1,916 of 84,556 pairs).

We initially identified 408 ACE indicators, which were condensed into 166 indicators (178 with covariates) following redistribution of rare indicators (figure 1). We entered all 166 indicators and 12 covariates of birth characteristics and maternal comorbidities into seven variable selection models of child maltreatment and maternal IPV (table 2; figure 2). There was large variation in selected indicators between different models (appendix p 31). Overall, we identified 71 consistently retained and 41 consistently excluded indicators across at least five models, leaving 34 tentative indicators for expert relevance ratings (appendix pp 10–15). The largest proportion (28 [68%]) of exclusions were indicators relating to high-risk presentations of child maltreatment (eg, fractures, intracranial injuries, contusions, or anogenital symptoms). Models consistently retained any suspected child maltreatment or maternal IPV and covariates related to ACEs (eg, younger maternal age or social deprivation). After the expert panel resolved disagreements over inclusions (Fleiss kappa interrater agreement >0.64), 94 indicators were combined into 56 final indicators (63 including the seven indicators making up the reference standard; table 1). This step ensured a minimum of 100 observations per indicator when applied to the smaller validation cohort. Covariates were excluded from the final selection.

The final 63 indicators clustered into six ACE domains, broadly confirming the manual groupings based on clinical relevance and existing ACE concepts. These domains were maternal mental health problems, maternal substance misuse, adverse family environments, child maltreatment, maternal IPV, and high-risk presentations of child maltreatment. In the dendrogram of hierarchically clustered indicators, including indicators as part of the reference standard, a few individual indicators clustered outside their originally grouped clusters (appendix p 33). We made no changes to the original domains, given that underlying coding descriptions of outlying indicators matched originally grouped domains better conceptually. For instance, compared with other maternal mental health problems, post-traumatic stress disorder clustered closer to violence-related indicators but was retained under maternal mental health problems.
lower specificity (78–90%) and underpredictions of any child maltreatment or maternal IPV (table 3; appendix p 39). This pattern remained consistent in survival analyses of the probability of child maltreatment or maternal IPV within 5 years after birth. Compared with no ACEs, the adjusted and weighted hazard ratio (HR) ranged from 10.0 (95% CI 8.1–12.2) for any ACE and 32.97 (26.63–40.83) for any suspected child maltreatment or maternal IPV, to 7.05 (5.74–8.67) for maternal substance misuse and 3.42 (2.86–4.10) for high-risk presentations of child maltreatment (figure 3A). The probability of any child maltreatment or maternal IPV also increased for each increase in the number of ACEs (adjusted weighted HR range 3.1–100.0; figure 3B). Random forest importance values for individual ACE indicators are shown in figure 4A.

**Table 3:** Performance metrics of each ACE domain’s predictive ability of any child maltreatment or maternal IPV using repeated cross-validated weighted-balanced random forests in the validation cohort from up to 2 years before birth to 5 years after birth (n=84 556)

| ACE Domain                                | AUC (95% CI) | F1 score* | True positives | True negatives | False positives | False negatives | Sensitivity (%) | Specificity (%) | Positive predictive value (%) | Negative predictive value (%) |
|-------------------------------------------|--------------|-----------|----------------|----------------|----------------|----------------|----------------|----------------|-----------------------------|-----------------------------|
| Any ACE                                   | 0.85 (0.84–0.86) | 0.17     | 1371           | 69 388         | 13 252         | 545            | 72%            | 84%            | 9%                          | 99%                         |
| Maternal mental health problems           | 0.72 (0.70–0.73) | 0.11     | 1149           | 64 736         | 17 904         | 767            | 60%            | 78%            | 6%                          | 99%                         |
| Maternal substance misuse                 | 0.68 (0.66–0.69) | 0.14     | 870            | 72 626         | 10 014         | 1046           | 45%            | 88%            | 8%                          | 99%                         |
| Adverse family environment                 | 0.67 (0.66–0.69) | 0.14     | 809            | 74 166         | 8474           | 1107           | 42%            | 90%            | 9%                          | 99%                         |
| Suspected child maltreatment               | 0.64 (0.63–0.65) | 0.24     | 551            | 81 906         | 734            | 1355           | 29%            | 99%            | 43%                         | 98%                         |
| High-risk presentations of child maltreatment | 0.59 (0.58–0.60) | 0.11  | 487            | 76 049         | 6591           | 1429           | 25%            | 92%            | 7%                          | 98%                         |
| Suspected maternal IPV                    | 0.55 (0.55–0.56) | 0.17     | 207            | 82 344         | 296            | 1709           | 11%            | 100%           | 41%                         | 98%                         |

Arranged in descending order of AUC estimates. ACE=adverse childhood experience. AUC=area under the curve. IPV=intimate partner violence. F1 scores portray a measure of precision and recall.

**Figure 2:** Median (IQR) relevance rankings for 166 ACE indicators and 12 covariates from seven cross-validated variable selection models of child maltreatment and maternal IPV in the derivation cohort

Rankings provided in descending order of risk association with any child maltreatment or maternal IPV from 2 years before birth to 5 years after birth (ie, highest value ranked first and lowest value ranked 12/817). Dark blue error bars indicate IQR and light blue error bars depict the complete range for all cross-validated models. Vertical purple lines represent the model-specific cutoffs for inclusion of indicators. Indicator-specific rankings are available in the appendix (pp 10-15). ACE=adverse childhood experience. FGM=female genital mutilation. IPV=intimate partner violence. LASSO=Least absolute shrinkage and selection operator. PPV=positive predictive value.
When applying all six ACE domains (63 indicators) to the validation cohort (84556 mother–child pairs), the adjusted and weighted period prevalence during the 2 years before and after birth was 39·1% (95% CI 38·7–39·5; 34773 pairs) for any ACE, 22·2% (21·8–22·5%; 20122 pairs) for maternal mental health problems, 15·7% (15·4–16·0%; 14549 pairs) for adverse family environments, 6·9% (6·7–7·2%; 7856 pairs) for

Figure 3: Kaplan-Meier curves for the probability of any child maltreatment or maternal IPV by ACE domain and by number of ACEs in the validation cohort (n=84556)

Kaplan-Meier curves present the probability of any child maltreatment or maternal IPV (reference standard) over time by ACE domain (A) and by the number of ACEs (B), from 2 years before birth to 5 years after birth, relative to mother–child pairs with no ACEs. ACE=adverse childhood experience. HR=hazard ratio. IPV=intimate partner violence. *At baseline, the number at risk represents the total number of exposed mother–child pairs by ACE domain who had not yet experienced the reference standard for the entire 7-year period.
maternal substance misuse, 8·1% (7·8–8·3%; 6808 pairs) for high-risk presentations of child maltreatment, 2·4% (2·3–5·6%; 2051 pairs) for child maltreatment, 1·0% (0·8–1·0%; 875 pairs) for maternal IPV, and 3·0% (2·9–3·2%; 2540 pairs) for any child maltreatment or maternal IPV (appendix p 4; figure 4B). For the same

Figure 4: Random forest variable importance values and PPVs of ACE indicators measured 2 years before to 5 years after birth, and Venn diagrams of ACE overlap measured 2 years before and after birth in the validation cohort

(A) PIMP values (circles) and PPVs (dashed line) of ACE indicators for predicting any child maltreatment or maternal IPV (reference standard) from 2 years before birth to 5 years after birth in the validation cohort from a split sample, cross-validated, random forest model. Indicators refer to both children and mothers, unless specified. PIMP values refer to the average decrease in the model’s prediction performance after randomly shuffling indicators (ie, breaking the association with the outcome). A higher PIMP value meant that the model relied more on the specific indicator for prediction performance. PPVs calculated by dividing reference standard cases by indicator positive cases. Overlap of recorded ACEs measured 2 years before birth to 5 years after birth by five different ACE domains collapsed into three domains (B), which included 31 836 exposed pairs and excluded the high-risk presentation of child maltreatment domain; by individual (C; ie, maternal vs child record), and by data source (D) for the final individual indicators. All estimates are from the validation cohort (n=84 556 mother–child pairs). ACE=adverse childhood experience. CPRD=Clinical Practice Research Datalink. GP=general practitioner. HES-APC=Hospital Episodes Statistics database for admitted patient care. IPV=intimate partner violence. PIMP=permutation variable importance. PPV=positive predictive value. PTSD=post-traumatic stress disorder.

www.thelancet.com/digital-health Vol 4 July 2022 e492
4-year period, 62·6% of ACEs were recorded in primary care only (21785 of 34773 pairs), and 16·8% were from hospital admissions or death records only (5853 of 34773 pairs; figure 4C, D). For any of the ACEs, 72·3% (25140 cases) were recorded in the maternal record only. For any of the child maltreatment or maternal IPV, 90·6% (2302 of 2540 cases) were identified in primary care only (overall crude prevalence 2·7%).

The results of the sensitivity analyses are provided in the appendix (pp 40–44). The results remained relatively robust across all sensitivity analyses (any ACE, range AUC 0·73–0·84). However, stratification by data source showed that the ACEs recorded in the HES-APC and the Office for National Statistics mortality register only reduced the overall AUC to 0·73, relative to CPRD only, or all sources.

**Discussion**

We developed and evaluated 63 indicators of ACEs, which represent six distinct and clinically meaningful domains, using data from a large representative English birth cohort of 21393 mothers and children. ACE indicators were derived by combining evidence from national guidance, systematic reviews, multistage risk prediction models of child maltreatment or maternal IPV, and clinical review. Validation estimates remained robust in different subgroups analyses, and manually grouped indicator domains were broadly confirmed in clustering analyses. Overall, we found that 39·1% of linked mothers and children had ACEs recorded in primary and secondary care 2 years before and after birth. The findings underscore the potential utility of linked data systems and secondary care 2 years before and after birth. The findings forward for external validation. Third, CPRD GOLD contains EHRs from the Vision data system, one of the three main primary care data systems in the UK (ie, Egton Medical Information Systems, SystmOne, and Vision); however, it has the least data coverage. Ever-changing policies and coding practices mean that the recording of ACEs might be influenced by different EHR systems, time-specific trends, and changes in NHS reporting demands. However, CPRD GOLD is the most widely used primary care data source for epidemiological research in the UK, and the only primary care data source with validated mother–baby linkage. Further research to validate mother–baby linkages and ACEs in larger GP databases are needed to generalise findings (eg, the CPRD Aurum). Finally, we could not link children with their fathers, a long-standing issue of anonymised secondary and primary care data. Therefore, indicators would have provided an underestimate of ACEs in the family. Nevertheless, a Swedish registry study showed that maternal mental health problems (ie, the largest form of ACE) were associated with a small increased risk of child injuries, but had a significantly larger effect relative to paternal mental health problems.

Our study is unique in developing ACE indicators based on the EHRs of mothers and children in primary care only (overall crude prevalence 2·7%). The longitudinal mechanisms underlying these associations are unclear and produce modest effect sizes compared with other contextual factors. A multisite birth cohort study of 3269 children found that ACEs reported in childhood showed relatively low accuracy in predicting poor health outcomes in adulthood (aged 18–45 years) when adjusting for socioeconomic factors. We used a reference standard that separated the adverse experience from the adverse stress response of the child, overcoming previous measurement limitations of recall bias and influences by other life factors as children grow up.

This study has several limitations. First, the absence of an independent reference standard meant that indicators could have overestimated risk for families, in whom ACEs are more likely to be noticed and recorded (eg, families with more complex needs). Second, structured data on child maltreatment and maternal IPV reflect only a small proportion of affected women and children presenting to services, and many children might have concerns recorded only in free-text data that is not captured in coded data. However, linkage to self-report data of children is practically challenging and susceptible to self-report biases. Future linkage of EHRs to children’s social care data offers an alternative way forward for external validation. Third, CPRD GOLD contains EHRs from the Vision data system, one of the three main primary care data systems in the UK (ie, Egton Medical Information Systems, SystmOne, and Vision); however, it has the least data coverage. Ever-changing policies and coding practices mean that the recording of ACEs might be influenced by different EHR systems, time-specific trends, and changes in NHS reporting demands. However, CPRD GOLD is the most widely used primary care data source for epidemiological research in the UK, and the only primary care data source with validated mother–baby linkage. Further research to validate mother–baby linkages and ACEs in larger GP databases are needed to generalise findings (eg, CPRD Aurum). Finally, we could not link children with their fathers, a long-standing issue of anonymised secondary and primary care data. Therefore, indicators would have provided an underestimation of ACEs in the family. Nevertheless, a Swedish registry study showed that maternal mental health problems (ie, the largest form of ACE) were associated with a small increased risk of child injuries, but had a significantly larger effect relative to paternal mental health problems.

Our study is unique in developing ACE indicators based on the EHRs of mothers and children in primary care only (overall crude prevalence 2·7%). The longitudinal mechanisms underlying these associations are unclear and produce modest effect sizes compared with other contextual factors. A multisite birth cohort study of 3269 children found that ACEs reported in childhood showed relatively low accuracy in predicting poor health outcomes in adulthood (aged 18–45 years) when adjusting for socioeconomic factors. We used a reference standard that separated the adverse experience from the adverse stress response of the child, overcoming previous measurement limitations of recall bias and influences by other life factors as children grow up.

This study has several limitations. First, the absence of an independent reference standard meant that indicators could have overestimated risk for families, in whom ACEs are more likely to be noticed and recorded (eg, families with more complex needs). Second, structured data on child maltreatment and maternal IPV reflect only a small proportion of affected women and children presenting to services, and many children might have concerns recorded only in free-text data that is not captured in coded data. However, linkage to self-report data of children is practically challenging and susceptible to self-report biases. Future linkage of EHRs to children’s social care data offers an alternative way forward for external validation. Third, CPRD GOLD contains EHRs from the Vision data system, one of the three main primary care data systems in the UK (ie, Egton Medical Information Systems, SystmOne, and Vision); however, it has the least data coverage. Ever-changing policies and coding practices mean that the recording of ACEs might be influenced by different EHR systems, time-specific trends, and changes in NHS reporting demands. However, CPRD GOLD is the most widely used primary care data source for epidemiological research in the UK, and the only primary care data source with validated mother–baby linkage. Further research to validate mother–baby linkages and ACEs in larger GP databases are needed to generalise findings (eg, CPRD Aurum). Finally, we could not link children with their fathers, a long-standing issue of anonymised secondary and primary care data. Therefore, indicators would have provided an underestimation of ACEs in the family. Nevertheless, a Swedish registry study showed that maternal mental health problems (ie, the largest form of ACE) were associated with a small increased risk of child injuries, but had a significantly larger effect relative to paternal mental health problems.
Articles

and secondary care. Over 70% of ACEs were identified from maternal records, with most ACEs recorded in primary care in the first years of the child’s life course. ACEs are preventable and are a clinically important target for early primary care responses, including prompting additional questions, arranging home visits and referrals, and increasing monitoring opportunities.9,10  Therefore, our findings emphasise the importance of using a think-family approach and linked EHR systems (eg, GP family tab) to scrutinise the primary care records of family members for ACEs. Additionally, our findings represent an important first step towards future evaluations of integrating ACE indicators into workflows for prioritising resources.9 Care staff work under immense pressure, with recurrent staffing shortages, waiting lists, and time limits (eg, 10 mins per GP consultation).98 In many cases, staff do not have the time to process both the child’s and mother’s records to examine the need for potential support to inform early responses.9,98 At a broader level, ACE domains have the potential to help organisations better understand trends and relationships between ACEs, other risk factors (eg, obesity or smoking), and chronic conditions to commission trauma-informed care aimed at reducing risk factors.9,11,12

Nevertheless, the potential use of digitally curated ACE indicators in the future requires additional research before service implementation. Current ACE indicators draw on maternal records to trigger actions for the child and the mother. This method potentially exposes the mother’s confidential information to the individuals involved in the child’s care without her consent, including to a perpetrator of abuse (eg, potential child maltreatment triggered by maternal IPV perpetrated by the child’s father). Hence, routine use of ACE indicators needs careful piloting for potential harms and benefits, potential stigma, ethics, and public acceptability, followed by randomised trials to test their efficacy in improving outcomes for children exposed to ACEs.91 We openly provide all excluded and included ACE indicators, cross-mapped across different systems, to aid further evaluation.

Contributors
SS and RG conceived the study, SS, RG, and LL designed the study. SS and LL completed the statistical analyses. RG and LL provided study supervision. AG-I had full access to the CPRD database to extract the study population. SS, LL, AG-I, and RG accessed and verified all the data in the study. All authors contributed to the interpretation of the data, drafting of the manuscript revisions, and had final responsibility for the decision to submit for publication.

Declaration of interests
All authors declare no competing interests.

Data sharing
This study uses data from the CPRD, a research service that provides primary care and linked data for public health research. CPRD data governance does not allow distribution or access to data to other parties outside of the approved study protocol. Researchers can apply for data access with a study protocol at the CPRD website and would need approval by the Research Data Governance Secretariat. We provide all relevant code lists and coding scripts on the ACEs in EHRs library of indicators. All code is shared without investigator support.

Acknowledgments
This work uses data provided by patients and collected by the NHS as part of their care and support. We are incredibly grateful to the generosity of the patients and their families, along with the participating general practitioner practices and NHS staff, for their ongoing contribution to research into mental health and family violence. We also acknowledge Rachel Ashwick (University of Oxford, Oxford, UK), who supported the independent screening in developing the ACE indicators partly published elsewhere. In addition, we acknowledge Linda Willaars (University College London, London, UK) for supporting data management. This study was carried out as part of the CALIBER resource. CALIBER, led by the UCL Institute of Health Informatics, is a research resource providing validated electronic health record phenotyping algorithms and tools for national structured data sources.9 This study is based on data from the CPRD obtained under licence from the UK Medicines and Healthcare Products Regulatory Agency. The interpretation and conclusions contained in this study are those of the authors alone. Hospital Episodes Statistics and the Office for National Statistics are under copyright (2020), re-used with the permission of the Health and Social Care Information Centre. All rights reserved. The research was supported in part by the National Institute for Health Research (NIHR) Great Ormond Street Hospital Biomedical Research Centre. This research benefits from and contributes to the NIHR Children and Families Policy Research Unit, but was not commissioned by the NIHR Policy Research Programme. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, the Department of Health and Social Care or its arm’s length bodies, and other governmental departments. RG was (in part) supported by the Health Data Research UK (grant LOND1).

References
1 Syed S, Ashwick R, Schlוסser M, Gonzalez-Izquierdo A, Li I, Gilbert R. Predictive value of indicators for identifying child maltreatment and intimate partner violence in coded electronic health records: a systematic review and meta-analysis. Arch Dis Child 2021; 106: 44–53.
2 Hughes K, Bellis MA, Hardcastle KA, et al. The effect of multiple adverse childhood experiences on health: a systematic review and meta-analysis. Lancet Public Health 2017; 2: e356–66.
3 Bellis MA, Hughes K, Ford K, Ramos Rodriguez G, Sethi D, Passmore J. Life course health consequences and associated annual costs of adverse childhood experiences across Europe and North America: a systematic review and meta-analysis. Lancet Public Health 2019; 4: e517–28.
4 Gilbert R, Widom CS, Browne K, Fergusson D, Webb E, Janson S. Burden and consequences of child maltreatment in high-income countries. Lancet 2009; 373: 68–81.
5 Merrick MT, Ford DC, Ports KA, Guinan AS. Prevalence of adverse childhood experiences from the 2011–2014 behavioral risk factor surveillance system in 23 states. JAMA Pediatr 2018; 172: 1038–44.
6 Gilbert R, Kemp A, Thoburn J, et al. Recognising and responding to child maltreatment. Lancet 2009; 373: 167–80.
7 Sidebotham P, Brandon M, Bailey S, et al. Pathways to harm, pathways to protection: a triennial analysis of serious case reviews 2011 to 2014. London: Department for Education, 2016.
8 Paranjothy S, Evans A, Bandypadhyay A, et al. Risk of emergency hospital admission in children associated with mental disorders and alcohol misuse in the household: an electronic birth cohort study. Lancet Public Health 2018; 3: e279–88.
9 Hillis SD, Anda RF, Duhe SR, Felitti VJ, Marchbanks PA, Marks JS. The association between adverse childhood experiences and adolescent pregnancy, long-term psychosocial consequences, and fetal death. Pediatrics 2004; 113: 120–27.
10 Kessler RC, McLaughlin KA, Green JG, et al. Childhood adversities and adult psychopathology in the WHO World Mental Health Surveys. Br J Psychiatry 2010; 197: 378–85.
11 Narayan AJ, Lieberman AF, Masten AS. Intergenerational transmission and prevention of adverse childhood experiences (ACEs). Clin Psychol Rev 2021; 85: 101997.
12 Centers for Disease Control and Prevention. Preventing adverse childhood experiences: data to action (PACE/DJA). Aug 19, 2021. https://www.cdc.gov/violenceprevention/aces/preventingace-dataaction.html (accessed Feb 14, 2022).

For the CPRD website see https://www.cprd.com/
For the ACEs in EHRs library see www.ACEsinEHRs.com
13 Public Health England. No child left behind: a public health informed approach to improving outcomes for vulnerable children. Sept 2, 2020. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/933764/Public_health_approach_to_vulnerability_in_childhood.pdf (accessed Feb 14, 2022).

14 The Scottish Government. Scotland’s public health priorities. June 14, 2018. https://www.gov.scot/publications/scottlands-public-health-priorities/documents/ (accessed Feb 14, 2022).

15 Miktón CR, Tanaka M, Tomlinson M, et al. Global research priorities for interpersonal violence prevention: a modified Delphi study. Bull World Health Organ 2017; 95: 36–48.

16 WHO. Global plan of action to strengthen the role of the health system within a national multisectoral response to address interpersonal violence, in particular against women and girls, and against children. Geneva: World Health Organization, 2016.

17 Oram S, Trevillion K, Feder G, Howard LM. Prevalence of experiences of domestic violence among psychiatric patients: systematic review. Br J Psychiatry 2013; 202: 94–99.

18 Ramsay J, Rutterford C, Gregory A, et al. Domestic violence: knowledge, attitudes, and clinical practice of selected UK primary healthcare clinicians. Br J Gen Pract 2012; 62: e647–55.

19 González-Izquierdo A, Cortina-Borja M, Woodman J, et al. Maltreatment or violence-related injury in children and adolescents admitted to the NHS: comparison of trends in England and Scotland between 2005 and 2011. BMJ Open 2014; 4: e004474.

20 Abel KM, Hope H, Swift E, et al. Prevalence of maternal mental illness among children and adolescents in the UK between 2005 and 2017: a national retrospective cohort analysis. Lancet Public Health 2019; 4: e291–300.

21 Rod NH, Bengtsson J, Budtz-Jørgensen E, et al. Trajectories of childhood adversity and mortality in early adulthood: a population-based cohort study. Lancet 2020; 396: 489–97.

22 NHS England. Your pregnancy care. March 20, 2020. https://www.nhs.uk/pregnancy/your-pregnancy-care (accessed Feb 14, 2022).

23 Leite A, Thomas SL, Andrews NJ. Implementing near-real-time vaccine safety surveillance using the Clinical Practice Research Datalink (CPRD). Vaccine 2017; 35: 6885–92.

24 HM Government. The best start for life: a vision for the 1,001 critical days. March 25, 2021. https://www.gov.uk/government/publications/the-best-start-for-life-a-vision-for-the-1001-critical-days. March 25, 2021. https://www.gov.uk/government/publications/the-best-start-for-life-a-vision-for-the-1001-critical-days. March 25, 2021. https://www.gov.uk/government/publications/the-best-start-for-life-a-vision-for-the-1001-critical-days.

25 Hoffman JM, Watts DJ, Athey S, et al. Integrating explanation and prediction in computational social science. Nature 2021; 595: 181–88.

26 Fitch K, Bernstein SJ, Aguilar MD, Burnand B, LaCalle JR. The RAND/UCLA appropriateness method user’s manual. Santa Monica, CA: RAND, 2001.

27 Felitti VJ, Anda RF, Nordenberg D, et al. Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults. The Adverse Childhood Experiences (ACE) study. Am J Prev Med 1998; 14: 245–58.

28 Shonkoff JP, Corso PS, editor. Trajectories of childhood adversity and the development coming of age: science through the life course. New York, NY: United Nations Children’s Fund, 2018.

29 Leite A, Thomas SL, Andrews NJ. Implementing near-real-time vaccine safety surveillance using the Clinical Practice Research Datalink (CPRD). Vaccine 2017; 35: 6885–92.

30 Fitch K, Bernstein SJ, Aguilar MD, Burnand B, LaCalle JR. The RAND/UCLA appropriateness method user’s manual. Santa Monica, CA: RAND, 2001.

31 Shonkoff JP, Corso PS, editor. Trajectories of childhood adversity and the development coming of age: science through the life course. New York, NY: United Nations Children’s Fund, 2018.

32 Belsky J. Child maltreatment: an ecological integration. Am Psychol 1980; 35: 120–35.

33 Stevens GA, Alkema L, Black RE, et al. Guidelines for accurate and transparent health estimates reporting: the GATHER statement. PLoS Med 2016; 13:e002656.

34 Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. Br J Surg 2015; 102: 148–58.

35 Denexas S, González-Izquierdo A, Direk K, et al. UK phenomics platform for developing and validating electronic health record phenotypes: CALIBER. J Am Med Inform Assoc 2019; 26: 1545–59.

36 Herret E, Gallagher AM, Bhaskaran K, et al. Data resource profile: clinical practice research datalink (CPRD). Int J Epidemiol 2019; 48: 827–38.

37 Minassian C, Williams R, Meears WH, Smeth L, Campbell OMR, Thomas SL. Methods to generate and validate a pregnancy registry in the UK Clinical Practice Research Datalink primary care database. Pharmacoepidemiol Drug Saf 2019; 28: 923–31.

38 Herbert A, Wijalas I, Zylbersztejn A, Cromwell D, Hardell P. Data resource profile: hospital episode statistics admitted patient care (HES APC). Int J Epidemiol 2017; 46: 1093–1093i.

39 Ministry of Housing, Communities & Local Government. English indices of deprivation 2015. Sept 30, 2015. https://www.gov.uk/government/statistics/english-indices-of-deprivation-2015 (accessed Feb 14, 2022).

40 Harren K, Gilbert R, Fagg J, Gottmann A, van der Meulen J. Associations between pre-pregnancy psychosocial risk factors and infant outcomes: a population-based cohort study in England. Lancet Public Health 2021; 6: e97–105.

41 Public Health England. Supporting public health: children, young people and families. London: Public Health England, July 1, 2014. https://www.gov.uk/government/publications/commissioning-of-public-health-services-for-children (accessed Feb 14, 2022).

42 Marie-Mitchell A, Kostolansky R. A systematic review of trials to improve child outcomes associated with adverse childhood experiences. Am J Prev Med 2019; 56: 756–64.

43 HM Government. Working together to safeguard children: a guide to inter-agency working to safeguard and promote the welfare of children. July 2018. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/942454/Working_together_to_safeguard_children_inter_agency_guidance.pdf (accessed Feb 14, 2022).

44 Department of Health. Alcohol guidelines review: report from the guidelines development group to the UK chief medical officers. January, 2016. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/545739/GDG_report-Jan2016.pdf (accessed April 14, 2022).

45 Woodman J, Allister J, Rafi I, et al. A simple approach to improve recording of concerns about child maltreatment in primary care records: developing a quality improvement intervention. Br J Gen Pract 2012; 62: e78–86.

46 John A, McGregor J, Fone D, et al. Case-finding for common mental disorders of anxiety and depression in primary care: an external validation of routinely collected data. BMC Med Inform Decis Mak 2016; 16: 1–10.

47 Public Health England. No child left behind: understanding and quantifying vulnerability. London: Public Health England, 2020.

48 Krinner LM, Warren-Findlow J, Bowling J, Issel LM, Reeve CL. The dimensionality of adverse childhood experiences: a scoping review of ACE dimensions measurement. Child Abuse Negl 2021; 121: 105270.

49 Kalmakis KA, Chandler GE. Adverse childhood experiences: towards a clear conceptual meaning. J Adv Nurs 2014; 70: 489–501.

50 National Institute for Health and Care Excellence. Child maltreatment: when to suspect maltreatment in under 18s. London: National Institute for Health and Care Excellence, 2021.

51 Royal College of General Practitioners. Child safeguarding toolkit. https://elearning.rcgp.org.uk/mod/book/view.php?id=12531 (accessed Feb 14, 2022).

52 Black MM, Walker SP, Fernald LCH, et al. Early childhood development coming of age: science through the life course. Lancet 2017; 389: 77–90.

53 Zylbersztejn A, Verfürden M, Hardell P, Gilbert R, Wijalas L. Phenotyping congenital anomalies in administrative hospital records. Paediatr Perinat Epidemiol 2020; 34: 21–28.

54 Simkiss DE, Spencer NJ, Stallard N, Thorogood M. Health service use in families where children enter public care: a nested case control study using the General Practice Research Database. BMC Health Serv Res 2012; 12: 65.

55 Gilbert R, Fluke J, O’Donnell M, et al. Child maltreatment: variation in trends and policies in six developed countries. Lancet 2012; 379: 758–72.
