The RECORD statement – checklist of items, extended from the STROBE statement, that should be reported in observational studies using routinely collected health data.

	Item No.	STROBE items	Location in manuscript where items are reported	RECORD items	Location in manuscript where items are reported
Title and abstract					
	1	(a) Indicate the study's design with a commonly used term in the title or the abstract (b) Provide in the abstract an informative and balanced summary of what was done and what was found	Study design is in the abstract	RECORD 1.1: The type of data used should be specified in the title or abstract. When possible, the name of the databases used should be included. RECORD 1.2: If applicable, the geographic region and timeframe within which the study took place should be reported in the title or abstract. RECORD 1.3: If linkage between databases was conducted for the study, this should be clearly stated in the title or abstract.	Type of data (EMR) and database is specified in the title and abstract. P. 2 line 7-8 Geographic region is mentioned in methods of abstract. P. 2 line 4
Introduction					
Background rationale	2	Explain the scientific background and rationale for the investigation being reported	Included in introduction		P. 3, line 21-23
Objectives	3	State specific objectives, including any prespecified hypotheses	Included in introduction		P. 4 line 2-4
Methods					

Study Design	4	Present key elements of study design early in the paper	Methods section, first paragraph		p. 4 line 8
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	All included in methods		p. 4 lines 8-11, 13-24
Participants	6	*(a) Cohort study* - Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study* - Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study* - Give the eligibility criteria, and the sources and methods of selection of participants *(b) Cohort study* - For matched studies, give matching criteria and number of exposed and unexposed *Case-control study* - For matched studies, give matching criteria and the number of controls per case	Described in methods section on study design and population. As well in figure 1. Study inclusion criteria are described on page 4 lines 13-24.	RECORD 6.1: The methods of study population selection (such as codes or algorithms used to identify subjects) should be listed in detail. If this is not possible, an explanation should be provided. RECORD 6.2: Any validation studies of the codes or algorithms used to select the population should be referenced. If validation was conducted for this study and not published elsewhere, detailed methods and results should be provided. RECORD 6.3: If the study involved linkage of databases, consider use of a flow diagram or other graphical display to demonstrate the data linkage process, including the number of individuals with linked data at each stage.	No algorithms were used to select the population of interest. None were used to selection population. A flow chart of study selection was provided (Figure 1 on page 12)

Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable.		RECORD 7.1: A complete list of codes and algorithms used to classify exposures, outcomes, confounders, and effect modifiers should be provided. If these cannot be reported, an explanation should be provided.	A reference to the algorithm to determine child ethnicity is provided in the methods. Page. 5, lines 12-13
Data sources/ measurement	8	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	These are described in the section "Measurement" in the methods		Page 5 lines 4-22
Bias	9	Describe any efforts to address potential sources of bias			
Study size	10	Explain how the study size was arrived at	All eligible children in the EMRALD cohort and described in Figure 1.		Page 12
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why	Described in the methods section.		Page 5 lines 4-22
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding (b) Describe any methods used to examine subgroups and interactions	Described in methods section. No subgroup analysis performed.		Page 6 lines 14-23 Continues on page 7 lines 1-6

		(c) Explain how missing data were addressed (d) *Cohort study* - If applicable, explain how loss to follow-up was addressed *Case-control study* - If applicable, explain how matching of cases and controls was addressed *Cross-sectional study* - If applicable, describe analytical methods taking account of sampling strategy (e) Describe any sensitivity analyses	Complete case analysis.		
Data access and cleaning methods		..		RECORD 12.1: Authors should describe the extent to which the investigators had access to the database population used to create the study population. RECORD 12.2: Authors should provide information on the data cleaning methods used in the study.	EMRALD database access described in Methods: study population. P. 4 lines 8-14 Data cleaning methods are described in the definition of exposure and reference previously published papers. Page 5 lines 6-9
Linkage		..		RECORD 12.3: State whether the	Data linkage is

				study included person-level, institutional-level, or other data linkage across two or more databases. The methods of linkage and methods of linkage quality evaluation should be provided.	described in the section on study population in methods. Page 4 lines 8-14
Results					
Participants	13	(a) Report the numbers of individuals at each stage of the study (*e.g.*, numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed) (b) Give reasons for non-participation at each stage. (c) Consider use of a flow diagram		RECORD 13.1: Describe in detail the selection of the persons included in the study (*i.e.,* study population selection) including filtering based on data quality, data availability and linkage. The selection of included persons can be described in the text and/or by means of the study flow diagram.	There is a description of the study population in the first paragraph of the methods and a follow chart to describe exclusions (figure 1). Page 12
Descriptive data	14	(a) Give characteristics of study participants (*e.g.*, demographic, clinical, social) and information on exposures and potential confounders (b) Indicate the number of participants with missing data for each variable of interest (c) *Cohort study* - summarise follow-up time (*e.g.*, average and total amount)	Table 1 includes this information. All missing data on covariates was below 5%.		Page 7 and 13
Outcome data	15	*Cohort study* - Report numbers of outcome events or summary measures over time	Table 2 presents prevalence estimates by age		Page 7 and 14

		Case-control study - Report numbers in each exposure category, or summary measures of exposure *Cross-sectional study* - Report numbers of outcome events or summary measures	and sex.		
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence interval). Make clear which confounders were adjusted for and why they were included (b) Report category boundaries when continuous variables were categorized (c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	Table 3 shows the adjusted models with 95% confidence intervals and p-values. Temporal trends are presented in Figure 2. As well by sex in supplemental figures.		Page 8 and table on page 14-15
Other analyses	17	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses	N/A		
Discussion					
Key results	18	Summarise key results with reference to study objectives	First paragraph of discussion		Page 8-9 lines 19-24, 1-4
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both	A summary and discussion of the limitations are presented in the 3rd	RECORD 19.1: Discuss the implications of using data that were not created or collected to answer the specific research question(s).	We discuss our attempts to mitigate specific limitations and

		direction and magnitude of any potential bias	paragraph.	Include discussion of misclassification bias, unmeasured confounding, missing data, and changing eligibility over time, as they pertain to the study being reported.	are presented in the paragraph on limitations. Page 10 lines13-23
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	Interpretation in given in the conclusion of the paper.		Page 11, lines 4-16
Generalisability	21	Discuss the generalisability (external validity) of the study results	Discussed in limitations paragraph.		Page 10 lines13-23
Other Information					
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	Provided on title page of the manuscript		Page 1
Accessibility of protocol, raw data, and programming code		..		RECORD 22.1: Authors should provide information on how to access any supplemental information such as the study protocol, raw data, or programming code.	All relevant information is kept at the Institute of Clinical Evaluative Sciences.

*Reference: Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, Sørensen HT, von Elm E, Langan SM, the RECORD Working Committee. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. *PLoS Medicine* 2015; in press.