STROBE checklist for EARTH study

Cohort Study
Checklist

	Item No	Recommendation	Page Number	Section	Additional Information
Title and abstract					
	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	2	Abstract	
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2	Abstract	
Introduction					
Background/ rationale	2	Explain the scientific background and rationale for the investigation being reported	3	Introduction	
Objectives	3	State specific objectives, including any prespecified hypotheses	3	Objectives	
Methods					
Study design	4	Present key elements of study design early in the paper	3	Methods- Study Design	
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	3	Methods- Data Source	
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	4	Methods- Population & Follow Up	
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	NA		
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	4	Methods- Population & Analysis	
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	3	Methods- Data Source	
Bias	9	Describe any efforts to address potential sources of bias	9	Limitations	
Study size	10	Explain how the study size was arrived at	NA		This is a descriptive study, and no comparative analysis is being carried out, and therefore a sample size calculation is not appropriate. Our cohort of over 32,000 patients is very large, and allows precise

					estimates of population variables, as shown in the paper by the narrow 99% confidence intervals.
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	4	Methods-Analysis	
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	4	Methods-Analysis	
		(*b*) Describe any methods used to examine subgroups and interactions	4 & 5	Methods-Follow Up	
		(*c*) Explain how missing data were addressed	4	Methods-Analysis	
		(*d*) If applicable, explain how loss to follow-up was addressed	4	Methods-Analysis	This is a GPRD study so the only type of missing data is values which are not recorded for every patients, such as body mass index. This is addressed in the Methods section. Further missing data is unlikely due to the nature of a GP database, All patients are followed up until death or until they transferred out of practice.
		(*e*) Describe any sensitivity analyses	NA		
Results					
Participants	13*	(*a*) Report numbers of individuals at each stage of study? eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	5	Baseline characteristics	
		(*b*) Give reasons for non-participation at each stage	5	Baseline characteristics	
		(*c*) Consider use of a flow diagram	-	-	
Descriptive data	14*	(*a*)Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	5	Results- Table 1	
		(*b*) Indicate number of participants with missing data for each variable of interest	5	Results- Table 1	

		(*c*) Summarise follow-up time (eg average and total amount)	5	Results- Baseline characteristics
Outcome data	15*	Report numbers of outcome events or summary measures over time	6	Results- Stroke mortality and Recurrent cardiovascular events
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	5 & 6	Results
		(*b*) Report category boundaries when continuous variables were categorized	NA	NA
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	NA	NA
Other analyses	17	Report other analyses done? eg analyses of subgroups and interactions, and sensitivity analyses	6 & 7	Results- Atrial fibrillation
Discussion				
Key results	18	Summarise key results with reference to study objectives	7	Discussion
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	8	Discussion- limitations
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	7 & 8 & 9	Discussion & Implications
Generalisability	21	Discuss the generalisability (external validity) of the study results	7	Discussion
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	10	Funding