

TABLE 1. STREGA reporting recommendations, extended from STROBE Statement

Item	Item number	STROBE Guideline	Extension for Genetic Association Studies (STREGA)	Relevant text from manuscript
Title and Abstract	1	(a) Indicate the study's design with a commonly used term in the title or the abstract.		Title and Abstract: "genome-wide association study"
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found.		Abstract: Methods (including sample size, significance threshold used) and Findings
Introduction				
<i>Background rationale</i>	2	Explain the scientific background and rationale for the investigation being reported.		Abstract: Background Introduction
<i>Objectives</i>	3	State specific objectives, including any pre-specified hypotheses.	State if the study is the first report of a genetic association, a replication effort, or both.	Introduction: Paragraph 4 – first report of genetic associations for chronic sputum production.
Methods				
<i>Study design</i>	4	Present key elements of study design early in the paper.		Abstract: Methods and Findings. Hypothesis and overall approach outlined at end of Introduction.

Item	Item number	STROBE Guideline	Extension for Genetic Association Studies (STREGA)	Relevant text from manuscript
				Methods
<i>Setting</i>	5	Describe the setting, locations and relevant dates, including periods of recruitment, exposure, follow-up, and data collection.		Methods: Subjects Data Availability
<i>Participants</i>	6	<p>(a) Cohort study – Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up.</p> <p>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.</p> <p>Cross-sectional study – Give the eligibility criteria, and the sources and methods of selection of participants.</p>	<i>Give information on the criteria and methods for selection of subsets of participants from a larger study, when relevant.</i>	Case-control study Methods: Study population
		<p>(b) Cohort study – For matched studies, give matching criteria and number of exposed and unexposed.</p> <p>Case-control study – For matched studies, give matching criteria and the number of controls per case.</p>		Case-control study Methods: Study population

Item	Item number	STROBE Guideline	Extension for Genetic Association Studies (STREGA)	Relevant text from manuscript
<i>Variables</i>	7	(a) Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable.	(b) Clearly define genetic exposures (genetic variants) using a widely-used nomenclature system. Identify variables likely to be associated with population stratification (confounding by ethnic origin).	Methods: Genome-wide association study of chronic sputum production Rsids used to describe associated variants.
<i>Data sources measurement</i>	8*	(a) 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.	(b) Describe laboratory methods, including source and storage of DNA, genotyping methods and platforms (including the allele calling algorithm used, and its version), error rates and call rates. State the laboratory/centre where genotyping was done. Describe comparability of laboratory methods if there is more than one group. Specify whether genotypes were assigned using all of the data from the study simultaneously or in smaller batches.	Methods: Genome-wide association study of chronic sputum production
<i>Bias</i>	9	(a) Describe any efforts to address potential sources of bias.	(b) For quantitative outcome variables, specify if any investigation of potential bias resulting from pharmacotherapy was undertaken. If relevant, describe the nature and magnitude of the potential bias,	Sensitivity analyses undertaken to investigate contribution of pre-existing respiratory disease or smoking status, time of year for completion of

Item	Item number	STROBE Guideline	Extension for Genetic Association Studies (STREGA)	Relevant text from manuscript
			<i>and explain what approach was used to deal with this.</i>	survey, sex, current smoking vs ever smoking.
<i>Study size</i>	10	Explain how the study size was arrived at.		Methods: Study population
<i>Quantitative variables</i>	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why.	<i>If applicable, describe how effects of treatment were dealt with.</i>	Methods: Genome-wide association study of chronic sputum production
<i>Statistical methods</i>	12	(a) Describe all statistical methods, including those used to control for confounding.	<i>State software version used and options (or settings) chosen.</i>	Methods: Genome-wide association study of chronic sputum production
		(b) Describe any methods used to examine subgroups and interactions.		Methods: Associations with other phenotypes
		(c) Explain how missing data were addressed.		Methods: Study population
		(d) Cohort study – If applicable, explain how loss to follow-up was addressed.		Methods: Study population

Item	Item number	STROBE Guideline	Extension for Genetic Association Studies (STREGA)	Relevant text from manuscript
		<p>Case-control study – If applicable, explain how matching of cases and controls was addressed.</p> <p>Cross-sectional study – If applicable, describe analytical methods taking account of sampling strategy.</p>		
		(e) Describe any sensitivity analyses.		Methods: Associations with other phenotypes
			<i>(f) State whether Hardy-Weinberg equilibrium was considered and, if so, how.</i>	Methods: Genome-wide association study of chronic sputum production (Paragraph 1)
			<i>(g) Describe any methods used for inferring genotypes or haplotypes.</i>	Methods: Genome-wide association study of chronic sputum production (Paragraph 1)
			<i>(h) Describe any methods used to assess or address population stratification.</i>	Methods: Genome-wide association study of chronic sputum production
			<i>(i) Describe any methods used to address multiple comparisons or to control risk of false positive findings.</i>	Methods: Genome-wide association study of chronic sputum production

Item	Item number	STROBE Guideline	Extension for Genetic Association Studies (STREGA)	Relevant text from manuscript
			<i>(j) Describe any methods used to address and correct for relatedness among subjects</i>	Methods: Study population
Results				
<i>Participants</i>	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.	<i>Report numbers of individuals in whom genotyping was attempted and numbers of individuals in whom genotyping was successful.</i>	Results: Paragraph 1, Table 1, figure 1
		(b) Give reasons for non-participation at each stage.		Results: Paragraph 1, figure 1
		(c) Consider use of a flow diagram.		Described in: Methods: Study population, figure 1
<i>Descriptive data</i>	14*	(a) Give characteristics of study participants (e.g., demographic, clinical, social) and information on exposures and potential confounders.	<i>Consider giving information by genotype.</i>	Results: Table 1

Item	Item number	STROBE Guideline	Extension for Genetic Association Studies (STREGA)	Relevant text from manuscript
		(b) Indicate the number of participants with missing data for each variable of interest.		Results: Table 1
		(c) Cohort study – Summarize follow-up time, e.g. average and total amount.		Not Applicable
<i>Outcome data</i>	15 *	Cohort study -Report numbers of outcome events or summary measures over time.	<i>Report outcomes (phenotypes) for each genotype category over time</i>	
		Case-control study – Report numbers in each exposure category, or summary measures of exposure.	<i>Report numbers in each genotype category</i>	Not applicable for genome-wide study. Coded allele frequencies for variants of interest reported throughout.
		Cross-sectional study – Report numbers of outcome events or summary measures.	<i>Report outcomes (phenotypes) for each genotype category</i>	
<i>Main results</i>	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence intervals). Make clear which confounders		Results: Table 2

Item	Item number	STROBE Guideline	Extension for Genetic Association Studies (STREGA)	Relevant text from manuscript
		were adjusted for and why they were included.		
		(b) Report category boundaries when continuous variables were categorized.		Not applicable
		(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period.		Not applicable
			(d) Report results of any adjustments for multiple comparisons.	Results: Table 2, Figure 4.
<i>Other analyses</i>	17	(a) Report other analyses done – e.g., analyses of subgroups and interactions, and sensitivity analyses.		Results: page 7, paragraph 1, Supplementary Figures 7 and 8, Supplementary Table 5.
			(b) If numerous genetic exposures (genetic variants) were examined, summarize results from all analyses undertaken.	Results: Figure 2

Item	Item number	STROBE Guideline	Extension for Genetic Association Studies (STREGA)	Relevant text from manuscript
			<i>(c) If detailed results are available elsewhere, state how they can be accessed.</i>	Data Availability
Discussion				
<i>Key results</i>	18	Summarize key results with reference to study objectives.		Discussion: Paragraph 1
<i>Limitations</i>	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias.		Discussion: Paragraph 7
<i>Interpretation</i>	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence.		Discussion: Paragraph 8
<i>Generalizability</i>	21	Discuss the generalizability (external validity) of the study results.		Discussion: Paragraph 6
Other Information				
<i>Funding</i>	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.		Acknowledgements

STREGA = STrengthening the REporting of Genetic Association studies; STROBE = STrengthening the Reporting of Observational Studies in Epidemiology.

* Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.