STROBE Statement—Checklist of items that should be included in reports of *case-control studies*

<table>
<thead>
<tr>
<th>Item No</th>
<th>Recommendation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Title and abstract</strong></td>
<td>1</td>
</tr>
<tr>
<td><strong>Introduction</strong></td>
<td>2</td>
</tr>
<tr>
<td><strong>Objectives</strong></td>
<td>3</td>
</tr>
<tr>
<td><strong>Methods</strong></td>
<td>4</td>
</tr>
<tr>
<td><strong>Setting</strong></td>
<td>5</td>
</tr>
</tbody>
</table>
| **Participants** | 6 | *(a)* 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. Eligibility criteria is outlined from lines 110-112 for patients, and in lines 119-120 for...
controls. Ascertainment is described from lines 111-113. Cases were collected in leprosy reference centres, and controls were healthy volunteers or blood donors from the same centres or nearby blood banks, that may share a greater proportion of patient’s environmental factors, including exposure to the parasite (it is estimated that around 70% of exposed individuals to M. leprae do not develop leprosy). This rationale is given in lines 115-117.

(b) For matched studies, give matching criteria and the number of controls per case

This study was unmatched; differences in demographic factors between both groups were corrected through logistic regression.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
</table>
| Variables                                   | 7      | Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable
Diagnostic criteria is given from line 113-115. Potential confounders (geographic origin, sex, age, ethnic group) are listed in Table 1. |
| 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
Genotyping results for all samples were obtained with multiplex PCR-SSP, as described in the topic “CR1 genotyping”, Table 2 and S1 Figure. mRNA measurements were done with a subset of Sinop samples, using TaqMan RT-PCR (described in the next section “mRNA Quantification”), and soluble CR1 levels were measured with ELISA in a group of samples from Curitiba and Sinop, selected in order to achieve a balanced distribution of ethnic groups and genotypes (in the topic “sCR1 Quantification”). |
| Bias                                        | 9      | Describe any efforts to address potential sources of bias
Differences in data distribution resulting from confounding factors, as geographic origin, age, sex and ethnicity, were statistically corrected using logistic regression. This is outlined in line 210-212. |
| Study size                                  | 10     | Explain how the study size was arrived at
We calculated the sample size needed for detecting associations with allele frequencies of at least 10% with 95% confidence level and a confidence interval of 5.0 and arrived at minimal 384 chromosomes (at least 192 individuals). Our sample sizes achieved this minimum number, with the exception of the group of paucibacillary patients, which nevertheless was close to it (182). These explanatory sentences can be found at lines 208-212. |
| Quantitative variables                      | 11     | Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why
Quantitative variables (CR1 mRNA and soluble CR1 levels) were not normally distributed and thus compared using nonparametric methods (see below). They were grouped into patients and controls, multi- and paucibacillary patients, to check if levels alter according to disease status and severity, and according to genotypes, to see if they influence gene and protein expression. Data were transformed into log10 for better graphical visualization. These information can be found in lines 221-226. |
| Statistical methods                         | 12     | (a) Describe all statistical methods, including those used to control for confounding
Statistical methods were described in lines 212-221, 226-228. Logistic regression was used for correcting association results for confounding factors as sex, age, geographic origin and ethnic group.
(b) Describe any methods used to examine subgroups and interactions |
The same methods outlined above were used for subgroups (multi and paucibacillary patients).

(c) Explain how missing data were addressed
We only included full-haplotyped samples in the logistic regression.

(d) If applicable, explain how matching of cases and controls was addressed
Not applicable.

(e) Describe any sensitivity analyses
There were none necessary for genotyping results. Inclusion of the outlier (excluded for better graphical visualization) did not change substantially the results of Mann-Whitney comparisons (they remained significant).

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
Available numbers were given in the “Subjects and samples” section.

(b) Give reasons for non-participation at each stage
Not applicable.

(c) Consider use of a flow diagram
Not necessary for this study.

Descriptive data 14*
(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders
Given in Table 1.

(b) Indicate number of participants with missing data for each variable of interest
Subgroup numbers investigated for mRNA and sCR1 levels are included in supplementary tables 1 and 2.

Outcome data 15*
Report numbers in each exposure category, or summary measures of exposure
Not applicable.

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
Results with these estimates are given in lines 258-266 and Table 3.

(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.
### Other analyses
17 Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses
Outlined in lines 260-263, Tables S1 and S2.

### Discussion

<table>
<thead>
<tr>
<th><strong>Key results</strong></th>
<th>18 Summarise key results with reference to study objectives</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Limitations</strong></td>
<td>Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias. Limitations as sample sizes, genetic background, etc. were discussed in lines 302-304, 320-321, 327.</td>
</tr>
<tr>
<td><strong>Interpretation</strong></td>
<td>Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence. Interpretation of results are given in lines 236-327.</td>
</tr>
<tr>
<td><strong>Generalisability</strong></td>
<td>Discuss the generalisability (external validity) of the study results. The final discussion can be found in lines 328-331.</td>
</tr>
</tbody>
</table>

### 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. Funding information can be found in the funding statement. |

*Give information separately for cases and controls.*