Supplementary Material

Sensitivity analysis: taking the uncertainty of the sensitivity and specificity estimates into account.

In the main text we have considered the sensitivity and specificity of the assay to be known exactly, as is current practice in methods for adjusting prevalence estimates for the assay characteristics[1]. However, both the sensitivity (88.1% [CI 83.7%, 91.8%]) and the specificity (93.2 CI [85.7%, 97.5%]) for the assay are only estimates and it will be statistically more principled to propagate that uncertainty through into the final adjusted seroprevalence estimate.

To do this, we proceed in 4 steps:

1. Find the best-fit beta distribution for each of (i) the seroprevalence estimate, (ii) the assay sensitivity, (iii) the assay specificity.
2. Use parametric bootstrap simulations to draw a large (10^7) sample of values from all 3 distributions.
3. For each of the 10^7 triplets of values thus obtained, compute the adjusted seroprevalence point estimate using the usual adjustment equation ($p_{adj} = \frac{p_{raw} - (1 - specificity)}{sensitivity - (1 - specificity)}$), thereby building up an empirical distribution of adjusted seroprevalence estimates.
4. Use the highest density interval with probability mass 0.95 from this empirical distribution as the final 95% confidence interval.

Applying this to our data, we get the same adjusted seroprevalence estimate (12.3%, since the same equation was used) but with a larger 95% confidence interval: [3.9%, 19.0%].

The conclusions from the main text are not affected by this wider confidence interval.

We have published the R code to do the above adjustment on GitHub:

https://github.com/gitMarcH/bootComb.

References

1. Reiczigel, J., Földi, J. & Ózsvári, L. Exact confidence limits for prevalence of a disease with an imperfect diagnostic test. *Epidemiol. Infect.* **138**, 1674–1678 (2010).