MJA
The Medical Journal of Australia

Supporting Information

Supplementary methods

Appendix to: Gidding HF, Machalek DA, Hendry AJ, et al. Seroprevalence of SARS-CoV-2-specific antibodies in Sydney after the first epidemic wave of 2020. *Med J Aust* 2021; doi: 10.5694/mja2.50940.

Bayesian inference for immunofluorescent antibody (IFA)-positive population seroprevalence, accounting for uncertainty in test sensitivity and specificity

The target parameter of interest was the IFA-positive population seroprevalence π, which is related to the observable IFA-positive proportion, p, via the sensitivity δ and specificity γ of the IFA test, as follows:

$$\pi = (p + \gamma - 1)/(\delta + \gamma - 1)$$

(by solving the equation $p = (1 - \gamma)(1 - \pi) + \delta\pi$).

Bayesian inference was undertaken to account for multiple sources of uncertainty including sampling error in the outcome and uncertainty in the test sensitivity and specificity. The outcome model for y, the observed number of IFA positive tests from a sample size of n total tests, was

$$y \sim \text{Binomial}(n, p),$$

with prior distributions required for the parameters π, δ and γ, all constrained to lie between 0 and 1.

Priors for sensitivity (δ) and specificity (γ) were derived from validation study results superimposed on uniform prior distributions (true positives (TP) = 107, false negatives (FN) = 11, true negatives (TN) = 2605, false positives (FP) = 18, adapted from Hueston *et al.*,[1] see main text for details) as follows:

$$\delta \sim \text{Beta}(TP + 1, FN + 1)$$

$$\gamma \sim \text{Beta}(TN + 1, FP + 1)$$

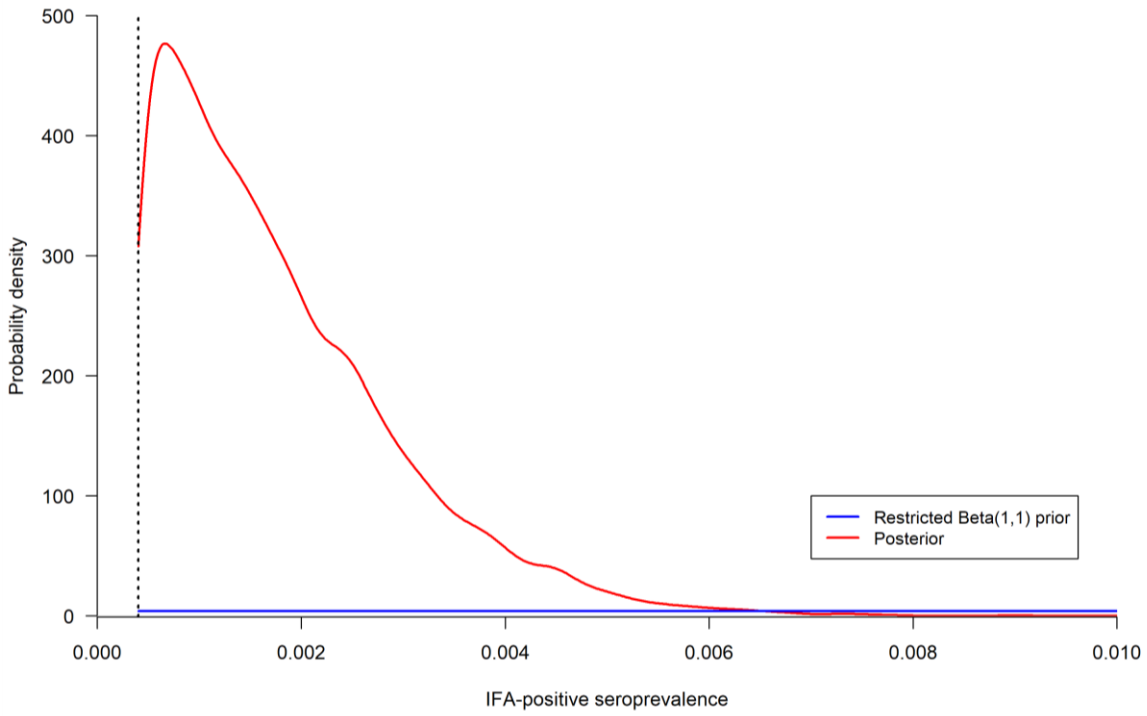
For the primary analysis, the prior distribution for π, the IFA-positive seroprevalence, was assumed to be uniform, i.e. Beta(1,1), restricted to values greater than 0.0004 (0.04%; Figure 1A). The lower bound was established from the ratio of notified cases to the ABS Sydney Estimated Resident Population, ABS 2019.[2] Although substantively unrealistic, the assumption of a uniform prior is a conservative approach that reflects the common preference for 'non-informative' prior assumptions. A sensitivity analysis was performed assuming an alternative prior distribution, Beta(0.2,10), also restricted to values greater than 0.0004 (0.04%; Figure 1B). This prior was more consistent with expert judgement that the true seroprevalence was unlikely to be more than 10-20 times greater than the cumulative incidence of notified cases; thus it assigned much greater weight to low values.

Bayesian models were fitted using the probabilistic programming language Stan (see below for code).[3,4] Point seroprevalence estimates were median values from posterior distributions and 95% credible intervals were calculated as 95% highest posterior density intervals, since posterior distributions were not symmetrical.

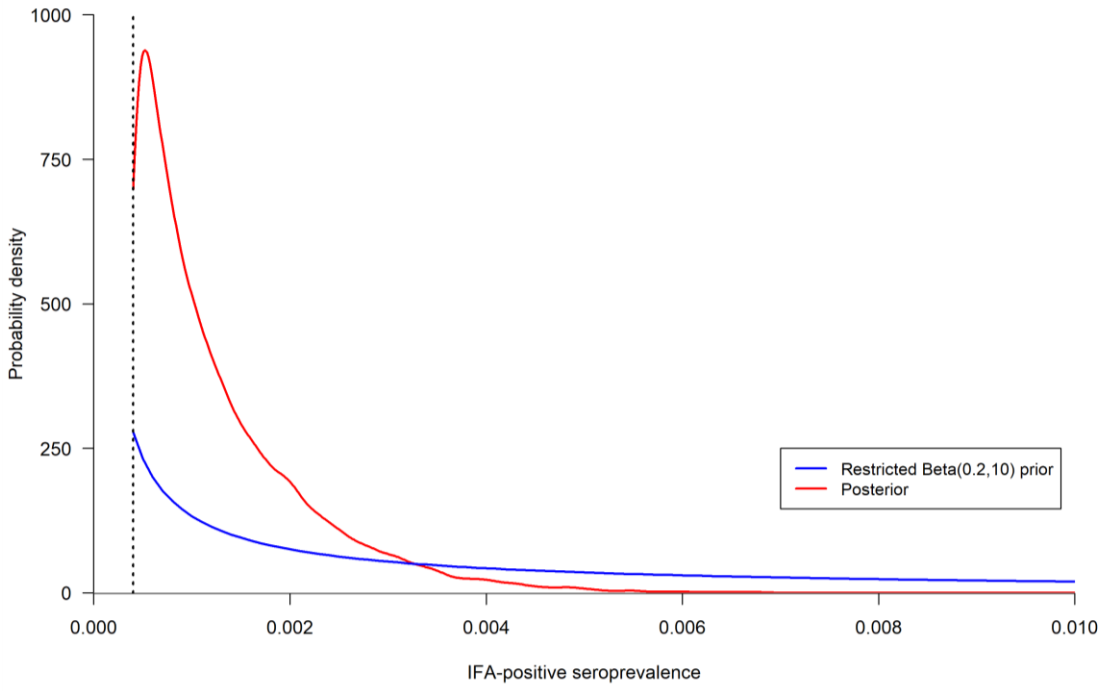
Prior and posterior probability density functions obtained from the primary analysis and sensitivity analysis of the general pathology subpopulation for estimation of IFA-positive seroprevalence are shown in Figures 1A and B.

Figure 1: Prior and posterior probability density functions for IFA-positive seroprevalence based on the general pathology subpopulation using: A) uniform Beta(1,1) prior restricted to π>0.04%; and B) informative Beta(0.2,10) prior restricted to π>0.04%.

A)



B)



Stan code

```
// Bayesian estimation of IFA-positive seroprevalence pooled across age
      groups and sex
// Three collections considered separately (blood donors, general pathology,
      antenatal screening)
// Accounting for uncertainty in test sensitivity and specificity

data{
  int<lower=0> N_pos;                       // observed number of positive tests
  int<lower=0> N_tests;                     // number of total tests
  int<lower=0> TP;                          // number of true positives from
                                                validation study (for test sens)
  int<lower=0> FN;                          // number of false negatives from
                                                validation study (for test sens)
  int<lower=0> FP;                          // number of false positives from
                                                validation study (for test spec)
  int<lower=0> TN;                          // number of true negatives from
                                                validation study (for test spec)
}
parameters{
  real<lower=0,upper=1> sens;               // estimated test sensitivity
  real<lower=0,upper=1> spec;               / estimated test specificity
  real<lower=0.00037,upper=1> theta_true;  // estimated true seroprevalence,
                        lower bound reflects ratio of case numbers to ABS 2019
                         Sydney population size (1,840/4,968,147)
}
transformed parameters{
  real theta_obs = theta_true * sens + (1 - theta_true) * ( 1 - spec);
}
model{
  sens ~ beta(TP+1, FN+1);
  spec ~ beta(TN+1, FP+1);
  theta_true ~ beta(1,1);
// prior distribution for true seroprevalence, must change to beta(0.2,10) for
      sensitivity analysis
  N_pos ~ binomial(N_tests, theta_obs);
}
```

References

1. Hueston L, Kok J, Guibone A, et al. The antibody response to SARS-CoV-2 infection. Open Forum Infectious Diseases 2020; 7: ofaa387.

2. Australian Bureau of Statistics. Estimated resident population (ERP) and components by SA2 and above (ASGS 2016), 2017 onwards. http://stat.data.abs.gov.au/Index.aspx?DataSetCode=ABS_ERP_COMP_SA (viewed Oct 2020).

3. Carpenter B, Gelman A, Hoffman MD, et al. Stan: a probabilistic programming language. J Statist Software 2017; 76: 1–32.

4. Stan Development Team. Stan user's guide, version 2.23. https://mc-stan.org/docs/2 23/stan-users-guide/index.html (viewed Oct 2020).