Applying mixture model methods to SARS-CoV-2 serosurvey data from the SEROCoV-POP study

Judith A Bouman[1,*], Sarah Kadelka[1], Silvia Stringhini[2,6,11], Francesco Pennacchio[2], Benjamin Meyer[9], Sabine Yerly[3,4], Laurent Kaiser[4,5,7], Idris Guessous[2,6], Andrew S Azman[6,8,10], Sebastian Bonhoeffer[1], and Roland R Regoes[1,*]

[1]Institute of Integrative Biology, ETH Zurich, Zurich, Switzerland
[2]Division of Primary Care, Geneva University Hospitals, Geneva, Switzerland
[3]Division of Laboratory Medicine, Geneva University Hospitals, Geneva, Switzerland
[4]Geneva Center for Emerging Viral Diseases and Laboratory of Virology, Geneva University Hospitals, Geneva, Switzerland
[5]Division of Infectious Diseases, Geneva University Hospitals, Geneva, Switzerland
[6]Department of Health and Community Medicine, Faculty of Medicine, University of Geneva, Geneva, Switzerland
[7]Department of Medicine, Faculty of Medicine, University of Geneva, Geneva, Switzerland
[8]Institute of Global Health, Faculty of Medicine, University of Geneva, Geneva, Switzerland
[9]Centre for Vaccinology, Department of Pathology and Immunology, University of Geneva, Geneva, Switzerland
[10]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA
[11]University Centre for General Medicine and Public Health, University of Lausanne, Lausanne, Switzerland
[*]Corresponding authors: Judith A. Bouman and Roland R. Regoes, (judith.bouman@env.ethz.ch, roland.regoes@env.ethz.ch).

July 19, 2021

Abstract

Serosurveys are an important tool to estimate the true extent of the current SARS-CoV-2 pandemic. So far, most serosurvey data have been analysed with cut-off based methods, which dichotomize individual measurements into sero-positives or negatives based on a predefined cutoff. However, mixture model methods can gain additional information from the same serosurvey data. Such methods refrain from dichotomizing individual values and instead use the full distribution of the serological measurements from pre-pandemic and COVID-19 controls to estimate the cumulative incidence. This study presents an application of mixture model methods to SARS-CoV-2 serosurvey data from the SEROCoV-POP study from April and May 2020 (2766 individuals). Besides estimating the total cumulative incidence in these data (8.1% (95% CI: 6.8% - 9.8%)), we applied extended mixture model methods to estimate an indirect indicator of disease severity, which is the fraction of cases with a distribution of antibody levels similar to hospitalised COVID-19 patients. This fraction is 51.2% (95% CI: 15.2% - 79.5%) across the full serosurvey, but differs between three age classes: 21.4% (95% CI: 0% - 59.6%) for individuals between 5 and 40 years old, 60.2% (95% CI: 21.5% - 100%) for individuals between 41 and 65 years old and 100% (95% CI: 20.1% - 100%) for individuals between 66 and 90 years old. Additionally, we find a mismatch between the inferred negative distribution of the serosurvey and the validation data of pre-pandemic controls. Overall, this study illustrates that mixture model methods can provide additional insights from serosurvey data.

Introduction

Serological surveys (serosurveys) are an important tool to estimate the cumulative incidence of SARS-CoV-2 infections in various geographic locations or risk groups during the current pandemic [1]. Based on the estimated cumulative incidence, one can even calculate several related parameters such as: the ascertainment rate, i.e. the fraction of cases detected, the relative risk of infection for sub-groups [2], the infection fatality rate [3].In 2020, many serosurveys have been conducted in a wide variety of geographic locations [4]. The vast majority of these serosurvey studies have been analysed with cutoff-based methods, meaning that each individual serological measurement has been dichotomized into sero-negative or positive based on a predefined cutoff value. This cutoff value has been defined based on a receiver operating characteristic (ROC) curve constructed from samples from pre-pandemic controls and known SARS-CoV-2 infections.

The cutoff-based method for analysing serosurveys has two main challenges. Firstly, the cutoff depends on validation data from known SARS-CoV-2 infections, which are often not representative of the full spectrum of possible infections. Instead, cases used for the validation data are, especially at the beginning of an epidemic, biased towards severe infections and early convalescent periods [5]. However, it is known that disease severity influences the antibody level after infection [6] and that antibody levels wane over time [7]. This can lead to overly confident estimates of the sensitivity and specificity of the serological test and therefore bias the estimated cumulative incidence. Secondly, the amount of information obtained from the serosurvey is reduced by dichotomizing the continuous measurements. As a result, the cutoff-based method does not allow to differentiate between several types of SARS-CoV-2 infections (for instance mild and severe infections), nor to detect or correct for a possible mismatch between the cases included in the validation data and those in the serosurvey.

Both of the posed challenges can be circumvented by using mixture model methods. Instead of dichotomizing the individual serological observations, mixture model methods estimate the cumulative incidence directly based on the full distribution of serological measurements for the pre-pandemic controls and known SARS-CoV-2 infections [8]. As a result, this inference framework can also be used to determine whether the cases included as positive COVID-19 controls are a good representation of the cases in the serosurvey data or whether cases with a distinct distribution of serological measurements (such as measurements from individuals with an asymptomatic or mild infection) are missing in the validation data. Moreover, mixture model methods allow to use multiple distinct distributions of cases separately in the analysis. Even though mixture models have been successfully applied to serosurvey data for several pathogens [9]–[12], they are rarely used to analyse serosurvey data from SARS-CoV-2 studies [13].

In this study, we apply mixture model methods to serosurvey data from the SEROCoV-POP study that was performed in Geneva in April and May of 2020 [2]. In addition to corroborating previous estimates of the cumulative incidence for these data (4.6 % in first week (95% CI: 2.4%-8.0%) to 10.9 % in the fifth week (95% CI: 8.2%-13.9%)) – we estimated a cumulative incidence of 8.1% (95% CI: 6.8% - 9.8%) over the whole period of sampling –, our aim is to show how mixture model methods can be used to extract more information from serosurveys. We use an extended mixture model that takes into consideration the distribution of antibody levels of both hospitalized COVID-19 patients and outpatients. This results in an estimate of what we call the indirect indicator of severity, which is defined as the fraction of individuals in the serosurvey that display a distribution of antibody levels similar to that of the hospitalized patients in the control data. This fraction is not a direct estimate of the fraction of cases in the serosurvey that were treated in a hospital, as the validation data does not contain positive control data from asymptomatic and mild cases. Therefore, we rather refer to this quantity as the indirect indicator of disease severity.

Methods

Data

We used the pre-pandemic and COVID-19 control data from Meyer et al [14]. These data contain the age of the individual, the measured IgG OD ratio of the Euroimmun SARS-CoV-2 serological assay and severity of the SARS-CoV-2 infection. We also used data from the SEROCoV-POP study from April and May 2020

from Stringhini et al [2]. Besides the age of the individual and the measured IgG OD ratio of the Euroimmun SARS-CoV-2 serological assay this data also reports the sex of each individual, the household structure between the individuals and the date of the measurement. The Euroimmun SARS-CoV-2 serological assay measures the IgG and IgA antibodies against the S1-domain of the spike protein of SARS-CoV-2 [14]. The IgG ratio is the result of the immunoreactivity of the sample measured at an optical density of 450 nm (OD450) divided by the OD450 of the calibrator [14], [15].

Mixture model methods

We have assembled all observations of the SEROCoV-POP study from April and May and apply the mixture model described by Bouman et al. [8]. The basic mixture model maximizes the likelihood equation 1. Here, U is the vector of observed IgG OD ratios in the serosurvey data, σ is a binary vector of length n with their underlying true serological status (1 for past infection and 0 for no past infection). The probabilities $p(U_i|\sigma_i = 0)$ and $p(U_i|\sigma_i = 1)$ capture the empirical distributions of IgG OD ratios for the pre-pandemic and COVID-19 control measurements, and π is the cumulative incidence.

$$ll(U) = \sum_{i=1}^{i=n} \log \left(p(\sigma_i = 1|\pi)\, p(U_i|\sigma_i = 1) + p(\sigma_i = 0|\pi)\, p(U_i|\sigma_i = 0) \right) \tag{1}$$

The likelihood is extended for the model where the outpatient and hospitalized cases are estimated separately, see equation 2. Here, π_{out} is the cumulative incidence of outpatient cases and π_{hosp} the cumulative incidence of hospitalized cases, σ_i can be 0 (no past infection), 1 (past outpatient infection) or 2 (past hospitalized infection).

$$\begin{aligned} ll(U) = \sum_{i=1}^{i=n} \log \Bigg(& p(\sigma_i = 1|\pi_{out})\, p(U_i|\sigma_i = 1) + \\ & p(\sigma_i = 2|\pi_{hosp})\, p(U_i|\sigma_i = 2) + p(\sigma_i = 0|\pi_{out}, \pi_{hosp})\, p(U_i|\sigma_i = 0) \Bigg) \end{aligned} \tag{2}$$

The 95% confidence intervals are estimated by bootstrapping the control distributions as well as the observations from the serosurvey. The various mixture models are compared with a likelihood ratio test.

We applied the extended model described above to the serosurvey data segregated into three age categories: 5-40 years, 41-65 years and 66-90 years. Even though the ages of the outpatient and hospitalized case distributions are significantly different, we used the whole distribution of both distributions for these analyses.

Testing for a mismatch between serosurvey and validation data To test if there is a mismatch between the observed serosurvey data and the validation data, we extend Equation 2 with an additional class (see equation 3). Thus, σ can now take one of four categorical values where the new one represents an additional, yet unknown, category of cases. The distribution of this additional category ($p(U_i|\sigma_i = 3)$) is modelled to be a normal distribution, where the mean and standard deviation are under optimization.

$$\begin{aligned} ll(U) = \sum_{i=1}^{i=n} \log \Bigg(& p(\sigma_i = 1|\pi_{out})\, p(U_i|\sigma_i = 1) + p(\sigma_i = 2|\pi_{hosp})\, p(U_i|\sigma_i = 2) + \\ & p(\sigma_i = 3|\pi_{add})\, p(U_i|\sigma_i = 3) + p(\sigma_i = 0|\pi_{out}, \pi_{hosp}, \pi_{add})\, p(U_i|\sigma_i = 0) \Bigg) \end{aligned} \tag{3}$$

This model is then compared to the model of Equation 2 to test if the additional distribution has significantly improved the likelihood of observing the serosurvey data.

We have also used an adjusted version of the method described above, where we summarized all observations below 0.34 into a point mass for the empirical distribution. The value of 0.34 is two standard deviations larger than the mean of the inferred mismatch in the distribution of pre-pandemic controls, to make sure that this mismatch is not included in the new distributions. The model is then performed with these distributions instead of the original empirical distributions of the negative and positive controls.

Results

Distributions of IgG OD Ratios significantly differ for hospitalized and outpatient SARS-CoV-2 positive controls

Meyer et al. (2020) validated the diagnostic accuracy of the Euroimmun SARS-CoV-2 IgG and IgA immunoassay for SARS-CoV-2 infection [14]. For this validation, they used a pre-pandemic negative control group (Negative controls, 326 individuals) and two clinically distinguishable positive control groups: individuals who were hospitalized in the University Hospitals of Geneva (COVID-19 hospitalized, 91 individuals), and individuals who were treated in outpatient clinics (COVID-19 outpatients, 90 individuals). All positive controls tested positive for SARS-CoV-2 by PCR. The observed IgG OD ratios of the Euroimmun SARS-CoV-2 immunoassay are shown in Figure 1 for the negative controls and both groups of positive controls. The distribution of the IgG OD ratios for the hospitalized positive controls is significantly different from the outpatient positive controls (two-sample Wilcoxon test, p-value = 1.122e-05).

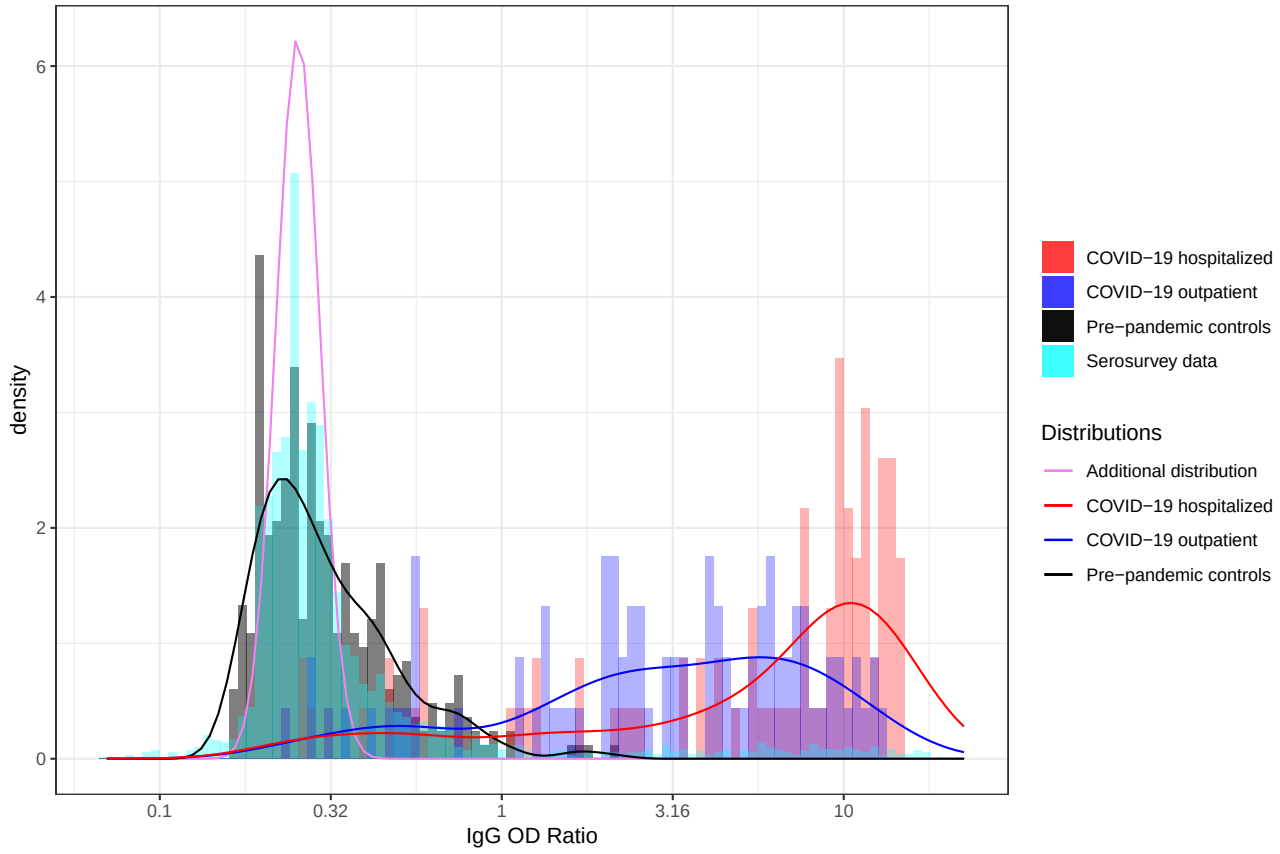

Figure 1: Histograms of IgG OD ratios of the Euroimmun SARS-CoV-2 IgG from the SEROCoV-POP study from April to May [2] and the validation data from Meyer et al [14]. Solid lines indicate the empirical distributions. The purple solid line shows the inferred additional distribution that is an indication of the mismatch between the pre-pandemic controls and the serosurvey data.

Model that separately estimates the cumulative incidence for hospitalized and outpatient control data is significantly better than model based on one type of controls only

The significant difference between the distributions of the IgG OD ratios for the hospitalized and the outpatient controls allows the mixture model method to simultaneously estimate the cumulative incidence of both types of cases in the data from the SEROCoV-POP study from April and May 2020 (see Equation 2). We find a cumulative incidence of 4.0% (95% CI: 0.8% - 7.4%)) for cases with a distribution of antibody levels similar to hospitalized controls and 4.2% (95% CI: 1.4% - 7.4%)) for cases with a distribution of antibody levels similar to outpatient controls. As a result, the fraction of cases in the serosurvey that can be explained with the distribution of the IgG OD ratios from the hospitalized controls, which we refer to as the *indirect indicator of disease severity*, is 51.2% (95% CI: 9.9% – 83.7%). The large 95% CI of this indicator of disease severity is caused by the overlap in the two positive control distributions.

To investigate if the model improves by including a separate estimate for both types of positive controls, we compared the likelihood of the estimates above to the likelihood from a model that is based on either the hospitalized or outpatient control data only (see Equation 1 and Table 1). The p-values in Table 1 indicate that the model is indeed significantly improved by estimating two cumulative incidences separately. Table 1 also shows that the point estimate of the total cumulative incidence estimate is higher if the mixture model is based on the outpatient controls only and lower if it is based on the hospitalized controls only, compared to the model that uses both distributions. This is expected, as the distribution obtained from the COVID-19 hospitalized controls is more distinguishable from the pre-pandemic controls than the COVID-19 outpatient controls.

Type of positive control data	**Cumulative incidence estimate**	**p-value compared with separate outpatient and hospitalized data**
Outpatient data only	8.4% (6.9% - 10.1%)	4.1e-08
Hospitalized data only	7.7% (6.3% - 9.6%)	5.0e-07
Outpatient and hospitalized data treated as separate distributions	8.1% (6.8% - 9.9%)	–

Table 1: Overview of cumulative incidence estimates based on various positive control data. The p-values are the result of a likelihood ratio test.

Indirect indicator of disease severity differs between age groups

It is known that there is a correlation between the age of an infected individual and the severity of a SARS-CoV-2 infection [16]. To validate our methodology, we estimated the indirect indicator of disease severity for three age-classes: 5 to 40 years, 41 to 65 years and 66 to 90 years. These estimates, together with the total cumulative incidence estimates for the age-classes, are shown in Table 2. Indeed, the indirect indicator of disease severity is highest for the oldest age class: we estimated that 100 % of the cases in the serosurvey can be explained by the distribution of the hospitalized COVID-19 controls, for the middle and young class this is 60.2 % and 21.4 % respectively (see Figure 2). Figure 3 shows that the maximal observed IgG ratio as well as the median of all values above the cutoff provided by the manufacturer (red dots) increase with age. However, the overall median of the distribution does not increase with age (black dots). This illustrates that the observed increase in the indirect indicator of disease severity is driven by the upper part of the IgG ratio distributions. The model that separately considers the age classes is significantly better than the model without these age classes after correcting for the increased amount of parameters (likelihood-ratio test, p-value = 0.009).

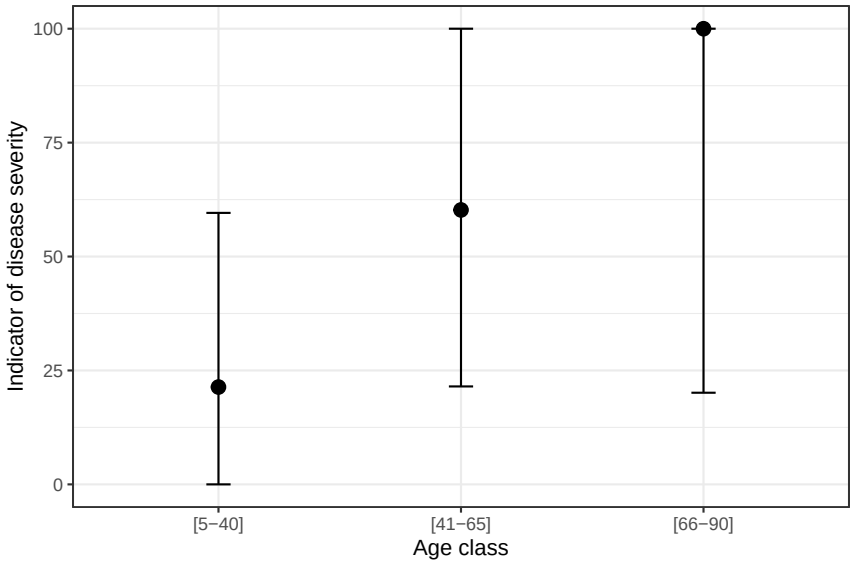

Figure 2: The indirect indicator of disease severity per age class, including the 95% confidence intervals.

Men, compared to women, are more likely to suffer from a severe SARS-CoV-2 infection [17]. Again, this can also be found by applying the mixture model method to the serosurvey data (see table 2). The point estimate of the indirect indicator of disease severity is higher for males compared to females, although this difference is not significant. The p-value of a likelihood-ratio test for the model that separates female and male participants with the original model is 0.046. The age distribution of the males and females are comparable in the serosurvey (two-sample Wilcoxon test, p-value = 0.18).

Sub-population	**Number of observations**	**Total cumulative incidence**	**indirect indicator of disease severity**
age-range [5-40]	1077	10.1% (7.8% - 12.7%)	21.4% (0%-59.6%)
age-range [41-65]	1355	7.6% (5.8% - 9.6%)	60.2% (21.5% - 100%)
age-range [66-90]	334	4.1% (0% - 6.9%)	100% (20.1% - 100%)
Females	1454	7.1% (5.3% - 9.0%)	45.3% (4.1% - 91.9%)
Males	1312	9.3% (7.3% - 11.6%)	50.8% (16.7% - 90.7%)

Table 2: Cumulative incidence and indicator of disease severity for three age-classes.

Mismatch between pre-pandemic controls and individuals without previous SARS-CoV-2 infection in the serosurvey

The mixture model can also be used to test if the validation data is a good representation of the observations in the serosurvey, or, in other words, whether the observed serosurvey data can be constructed by combining the pre-pandemic and COVID-19 control distributions from the validation data. In the SEROCoV-POP serosurvey, we infer a mismatch between the validation and the serosurvey data (p-value likelihood ratio test = $8e-105$). We find a missing distribution of measurements on the lower end of the range of the antibody levels, overlaying part of the distribution of the pre-pandemic control samples (see Figure 1). This indicates that the

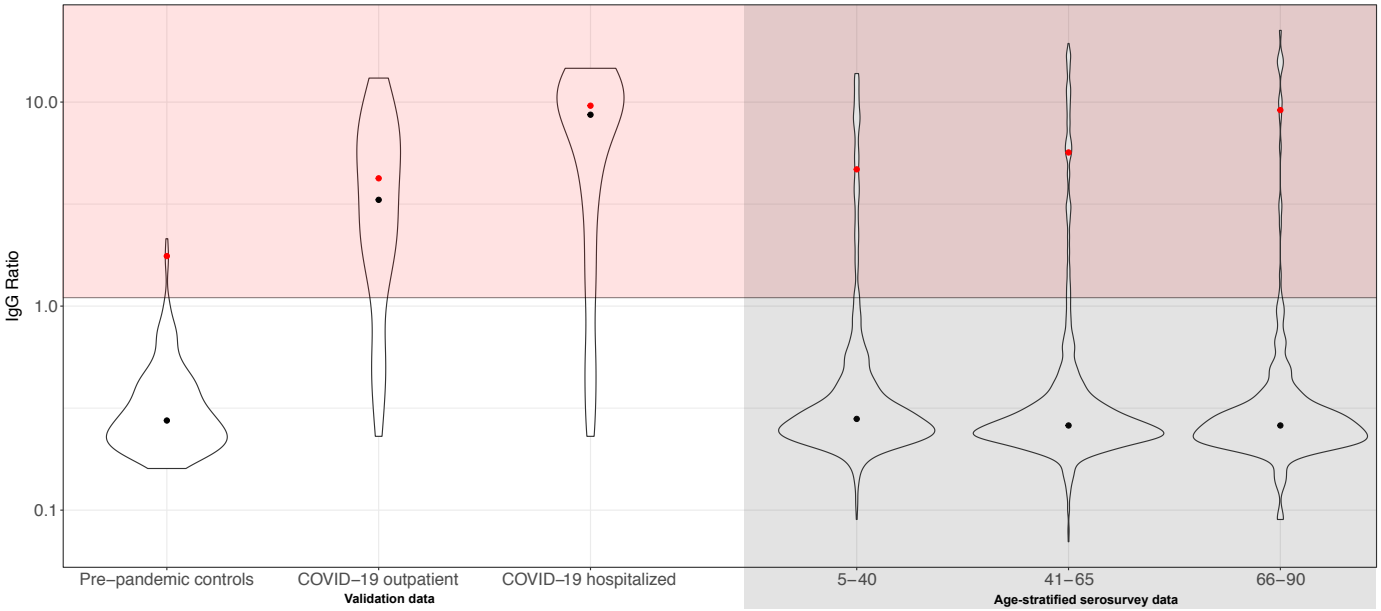

Figure 3: Violin plots of the distributions from the validation data (pre-pandemic controls, COVID-19 outpatient cases and COVID-19 hospitalized cases) and age-stratified serosurvey data. The black dots indicate the median of the full distribution and the red dots the median of all values larger than the cutoff of seropositivity provided by the manufacturer (1.1).

mismatch exists between the measurements from the pre-pandemic control samples and the individuals from the serosurvey study who likely did not have a past SARS-CoV-2 infection. The total cumulative incidence of SARS-CoV-2 infections for the model that includes this missing case distribution is 9.3% (95% CI: 6.7% - 10.6 %) and thus higher than without this distribution. The point estimate of the indirect indicator of disease severity is 37.4% (95% CI: 23.0% - 86.0 %) in this model.

No evidence for a missing positive control distribution with lower mean

The mismatch we identified between the control data and the serosurvey data appeared in the lower range of the IgG OD ratio. Hence, we assembled all values below 0.34 into a single point mass to eliminate this mismatch and test for an additional mismatch on the higher end of the observed IgG OD ratios. However, we did not find any evidence for such an additional mismatch. This suggests that the individuals with high IgG OD ratios in the serosurvey are well represented by the positive control data.

Discussion

In this study, we present an application of mixture model methods to SARS-CoV-2 serosurvey data. Serosurvey data are currently used to determine the proportion of seropositivity and to estimate the cumulative incidence and the relative risk of seropositivity in various sub-groups. This is usually done by introducing a cutoff for seropositivity.

We show that mixture models that use the entire distribution of the antibody levels rather than a cut-off for seropositivity, provide additional insights into aspects of an epidemic that are usually not addressed in serosurveys. Specifically, we have used mixture models to infer the cumulative incidence from distinct serological distributions, in this case those from hospitalized and outpatient COVID-19 positive controls. We found that the indirect indicator of disease severity (the fraction of individuals with antibody distributions similar to hospitalized cases) increases with age mirroring evindence from clinical studies. Additionally, mixture model methods can be used to test for a mismatch between the pre-pandemic and COVID-19 control

data and the serosurvey data, which could indicate that the cases observed in the population are not well represented by those included in the control data. While we provide evidence for such discrepancies, they are not indicative of a large fraction of cases with intermediate antibody levels that would be expected for asymptomatically infected individuals.

Although the mixture model approach naturally allows to implement declining antibody levels and sero-reversion [18], we have not corrected our estimate of the cumulative incidence for the possible effect of sero-reversion. The reason for this is that the serosurvey was conducted within 4 months of the start of the pandemic. Current estimates of anti-body half lives IgG RBD are around 50–106 days [19]. Therefore we expect the effect of sero-reversion to be negligible. Furthermore, we did not correct the estimate of cumulative incidence for age nor household structure because our study was aiming to provide a proof of concept rather than additional estimates for the sero-prevalence in Geneva. As a result, the estimates presented here are only representative for the study population and not for the general population of Geneva. Estimates for the cumulative incidence of the general population of Geneva from these data can be found in Stringhini et al. (2020) [2].

The presented estimates of the indirect indicator of disease severity have wide confidence intervals. This is caused by the fact that while the distributions of the antibody levels for COVID-19 hospitalized and outpatient cases are significantly different, there is quite a lot of overlap (see figure S1). This could potentially be improved if more detailed positive control data would be available to guide the construction of more distinguishable distributions of IgG OD ratios based on characteristics of the infections or infected individuals. Despite the large confidence intervals, we found that the indirect indicator of disease severity increases with age, corroborating previous reports [16]. Similarly, the point estimate of the indirect indicator of disease severity is higher for males compared to females, consistent with reported sex differences in ICU admission and death [17].

Furthermore, we have tested for a mismatch between the data used for the validation of the serological test and those of the serosurvey. We hypothesized to find such a mismatch because the validation data is solely derived from relatively severe cases whereas the majority of the cases in the serosurvey data are expected to have been mild. Mild infections have been shown to result in lower antibody responses compared to severe infections [15], [20], Not including the distribution of antibody levels in mild cases can lead to more false negatives and thus to an underestimation of the cumulative incidence [8]. We could indeed establish a mismatch between the validation and serosurvey data.

Interestingly, however, the mismatch we identify is not characterized by an intermediate level of antibodies in between the level of the pre-pandemic sera and the outpatients as we would expect for a missing distribution of mild or asymptotic cases. Opposite to our expectation we found that the serosurvey data display a narrower distribution at the lower end of the antibody levels than the pre-pandemic, negative controls — as if there were asymptotic or mild SARS-CoV-2 infections among the pre-pandemic controls. A more detailed characterization of the individuals from whom the pre-pandemic control sera were sampled, as well as the determination of antibody levels in asymptomatic and mild cases could shed further light on this mismatch and thus further improve the estimation of the cumulative incidence. An additional improvement could be obtained when a quantitative immuno-assay would be used, instead of the semi-quantitative Euroimmun that was available at the beginning of the pandemic.

Acknowledgements

We would like to thank Peter Ashcroft, Sonja Lehtinen and Jana Huisman for valuable comments on the manuscript. Roland Regoes gratefully acknowledges funding from the Botnar Research Centre for Child Health (grant number 2020-FS-354).

Ethics Statement

The SEROCoV-POP study was approved by the Cantonal Research Ethics Commission of Geneva, Switzerland (CER16-363). The full study protocol is available online (in French).

References

[1] M. Koopmans and B. Haagmans, "Assessing the extent of SARS-CoV-2 circulation through serological studies," *Nature Medicine*, vol. 26, no. 8, Aug. 2020, ISSN: 1078-8956. DOI: 10.1038/s41591-020-1018-x.

[2] S. Stringhini, A. Wisniak, G. Piumatti, A. S. Azman, S. A. Lauer, H. Baysson, D. De Ridder, D. Petrovic, S. Schrempft, K. Marcus, S. Yerly, I. Arm Vernez, O. Keiser, S. Hurst, K. M. Posfay-Barbe, D. Trono, D. Pittet, L. Gétaz, F. Chappuis, I. Eckerle, N. Vuilleumier, B. Meyer, A. Flahault, L. Kaiser, and I. Guessous, "Seroprevalence of anti-SARS-CoV-2 IgG antibodies in Geneva, Switzerland (SEROCoV-POP): a population-based study," *The Lancet*, vol. 396, no. 10247, pp. 313–319, Aug. 2020, ISSN: 1474547X. DOI: 10.1016/S0140-6736(20)31304-0.

[3] A. T. Levin, W. P. Hanage, N. Owusu-Boaitey, K. B. Cochran, S. P. Walsh, and G. Meyerowitz-Katz, "Assessing the age specificity of infection fatality rates for COVID-19: systematic review, meta-analysis, and public policy implications," *European Journal of Epidemiology*, vol. 35, no. 12, Dec. 2020, ISSN: 0393-2990. DOI: 10.1007/s10654-020-00698-1.

[4] X. Chen, Z. Chen, A. S. Azman, X. Deng, R. Sun, Z. Zhao, N. Zheng, X. Chen, W. Lu, T. Zhuang, J. Yang, C. Viboud, M. Ajelli, D. T. Leung, and H. Yu, "Serological evidence of human infection with SARS-CoV-2: a systematic review and meta-analysis," *The Lancet Global Health*, Mar. 2021, ISSN: 2214109X. DOI: 10.1016/S2214-109X(21)00026-7.

[5] S. Føns and K. A. Krogfelt, "How can we interpret SARS-CoV-2 antibody test results?" *Pathogens and Disease*, vol. 79, no. 1, Feb. 2021, ISSN: 2049-632X. DOI: 10.1093/femspd/ftaa069.

[6] C. Serre-Miranda, C. Nobrega, S. Roque, J. Canto-Gomes, C. Silva, N. Vieira, P. Barreira-Silva, P. Alves-Peixoto, J. Cotter, A. Reis, M. Formigo, H. Sarmento, O. Pires, A. Carvalho, D. Petrovykh, L. Diéguez, J. Sousa, N. Sousa, C. Capela, J. Palha, P. Cunha, and M. Correia-Neves, "Performance assessment of 11 commercial serological tests for SARS-CoV-2 on hospitalised COVID-19 patients," *International Journal of Infectious Diseases*, vol. 104, Mar. 2021, ISSN: 12019712. DOI: 10.1016/j.ijid.2021.01.038.

[7] G. den Hartog, E. R. A. Vos, L. L. van den Hoogen, M. van Boven, R. M. Schepp, G. Smits, J. van Vliet, L. Woudstra, A. J. Wijmenga-Monsuur, C. C. E. van Hagen, E. A. M. Sanders, H. E. de Melker, F. R. M. van der Klis, and R. S. van Binnendijk, "Persistence of antibodies to SARS-CoV-2 in relation to symptoms in a nationwide prospective study," *Clinical Infectious Diseases*, Feb. 2021, ISSN: 1058-4838. DOI: 10.1093/cid/ciab172.

[8] J. A. Bouman, J. Riou, S. Bonhoeffer, and R. R. Regoes, "Estimating the cumulative incidence of SARSCoV-2 with imperfect serological tests: Exploiting cutoff-free approaches," *PLoS Computational Biology*, vol. 17, no. 2, Feb. 2021, ISSN: 15537358. DOI: 10.1371/JOURNAL.PCBI.1008728.

[9] A. J. VYSE, N. J. GAY, L. M. HESKETH, R. PEBODY, P. MORGAN-CAPNER, and E. MILLER, "Interpreting serological surveys using mixture models: the seroepidemiology of measles, mumps and rubella in England and Wales at the beginning of the 21st century," *Epidemiology and Infection*, vol. 134, no. 6, Dec. 2006, ISSN: 0950-2688. DOI: 10.1017/S0950268806006340.

[10] M. C. Rota, M. Massari, G. Gabutti, M. Guido, A. De Donno, and M. L. C. d. Atti, "Measles serological survey in the Italian population: Interpretation of results using mixture model," *Vaccine*, vol. 26, no. 34, Aug. 2008, ISSN: 0264410X. DOI: 10.1016/j.vaccine.2008.05.094.

[11] M. A. Vink, J. van de Kassteele, J. Wallinga, P. F. M. Teunis, and J. A. Bogaards, "Estimating Seroprevalence of Human Papillomavirus Type 16 Using a Mixture Model with Smoothed Age-dependent Mixing Proportions," *Epidemiology*, vol. 26, no. 1, Jan. 2015, ISSN: 1044-3983. DOI: 10.1097/EDE.0000000000000196.

[12] M. van Boven, J. van de Kassteele, M. J. Korndewal, C. H. van Dorp, M. Kretzschmar, F. van der Klis, H. E. de Melker, A. C. Vossen, and D. van Baarle, "Infectious reactivation of cytomegalovirus explaining age- and sex-specific patterns of seroprevalence," *PLOS Computational Biology*, vol. 13, no. 9, Sep. 2017, ISSN: 1553-7358. DOI: 10.1371/journal.pcbi.1005719.

[13] C. Bottomley, M. Otiende, S. Uyoga, K. Gallagher, E.W. Kagucia, A.O. Etyang, D. Mugo, J. Gitonga, H. Karanja, J. Nyagwange, I.M.O. Adetifa, A. Agweyu, D.J. Nokes, G.M. Warimwe, and J.A.G. Scott, "Improving SARS-CoV-2 cumulative incidence estimation through mixture modelling of antibody levels," *MedRChiv*, 2021.

[14] B. Meyer, G. Torriani, S. Yerly, L. Mazza, A. Calame, I. Arm-Vernez, G. Zimmer, T. Agoritsas, J. Stirnemann, H. Spechbach, I. Guessous, S. Stringhini, J. Pugin, P. Roux-Lombard, L. Fontao, C. A. Siegrist, I. Eckerle, N. Vuilleumier, and L. Kaiser, "Validation of a commercially available SARS-CoV-2 serological immunoassay," *Clinical Microbiology and Infection*, vol. 26, no. 10, pp. 1386–1394, Oct. 2020, ISSN: 14690691. DOI: 10.1016/j.cmi.2020.06.024.

[15] N. M. Okba, M. A. Müller, W. Li, C. Wang, C. H. GeurtsvanKessel, V. M. Corman, M. M. Lamers, R. S. Sikkema, E. de Bruin, F. D. Chandler, Y. Yazdanpanah, Q. Le Hingrat, D. Descamps, N. Houhou-Fidouh, C. B. Reusken, B.-J. Bosch, C. Drosten, M. P. Koopmans, and B. L. Haagmans, "Severe Acute Respiratory Syndrome Coronavirus 2Specific Antibody Responses in Coronavirus Disease Patients," *Emerging Infectious Diseases*, vol. 26, no. 7, Jul. 2020, ISSN: 1080-6040. DOI: 10.3201/eid2607.200841.

[16] Y. Liu, B. Mao, S. Liang, J.-W. Yang, H.-W. Lu, Y.-H. Chai, L. Wang, L. Zhang, Q.-H. Li, L. Zhao, Y. He, X.-L. Gu, X.-B. Ji, L. Li, Z.-J. Jie, Q. Li, X.-Y. Li, H.-Z. Lu, W.-H. Zhang, Y.-L. Song, J.-M. Qu, and J.-F. Xu, "Association between age and clinical characteristics and outcomes of COVID-19," *European Respiratory Journal*, vol. 55, no. 5, May 2020, ISSN: 0903-1936. DOI: 10.1183/13993003.01112-2020.

[17] H. Peckham, N. M. de Gruijter, C. Raine, A. Radziszewska, C. Ciurtin, L. R. Wedderburn, E. C. Rosser, K. Webb, and C. T. Deakin, "Male sex identified by global COVID-19 meta-analysis as a risk factor for death and ITU admission," *Nature Communications*, vol. 11, no. 1, Dec. 2020, ISSN: 2041-1723. DOI: 10.1038/s41467-020-19741-6.

[18] S. Kadelka, J. A. Bouman, P. Ashcroft, and R. R. Regoes, "Comment on Buss et al., Science 2021: An alternative, empirically-supported adjustment for sero-reversion yields a 10 percentage point lower estimate of the cumulative incidence of SARS-CoV-2 in Manaus by October 2020," *arxiv*, 2021.

[19] J. M. Dan, J. Mateus, Y. Kato, K. M. Hastie, E. D. Yu, C. E. Faliti, A. Grifoni, S. I. Ramirez, S. Haupt, A. Frazier, C. Nakao, V. Rayaprolu, S. A. Rawlings, B. Peters, F. Krammer, V. Simon, E. O. Saphire, D. M. Smith, D. Weiskopf, A. Sette, and S. Crotty, "Immunological memory to SARS-CoV-2 assessed for up to 8 months after infection," *Science*, vol. 371, no. 6529, Feb. 2021, ISSN: 0036-8075. DOI: 10.1126/science.abf4063.

[20] C. H. GeurtsvanKessel, N. M. A. Okba, Z. Igloi, S. Bogers, C. W. E. Embregts, B. M. Laksono, L. Leijten, C. Rokx, B. Rijnders, J. Rahamat-Langendoen, J. P. C. van den Akker, J. J. A. van Kampen, A. A. van der Eijk, R. S. van Binnendijk, B. Haagmans, and M. Koopmans, "An evaluation of COVID-19 serological assays informs future diagnostics and exposure assessment," *Nature Communications*, vol. 11, no. 1, Dec. 2020, ISSN: 2041-1723. DOI: 10.1038/s41467-020-17317-y.

Supplementary Material

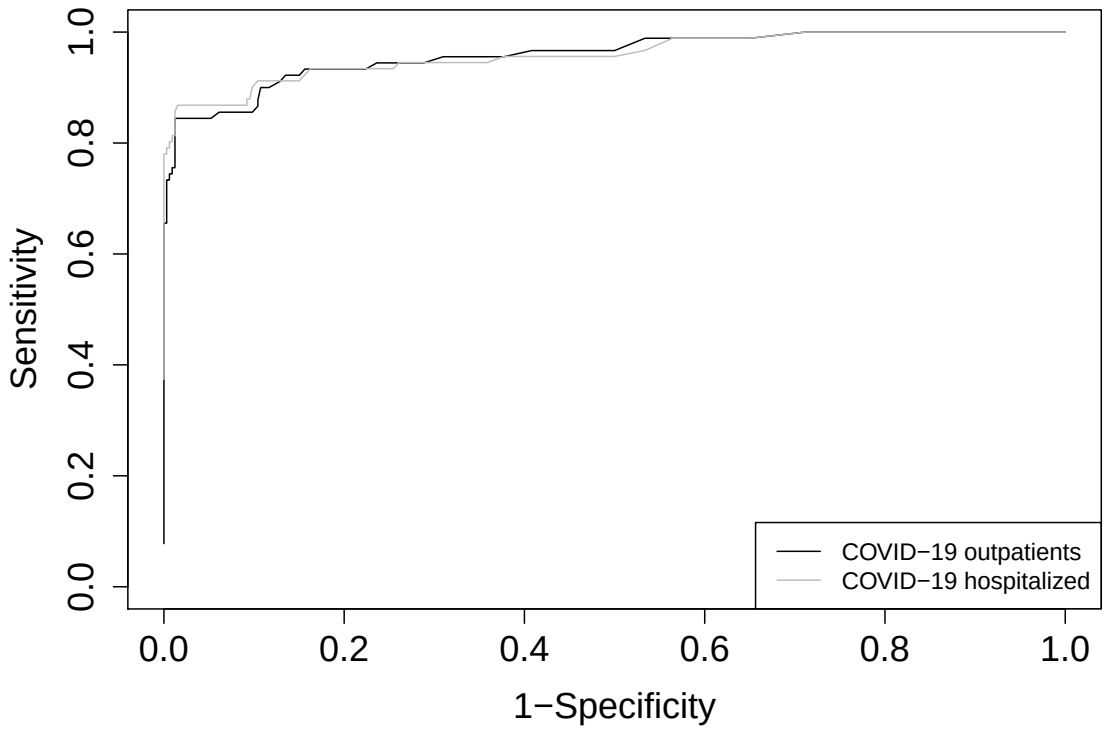

Figure S1: ROC-curve of the validation data for both the COVID-19 outpatient controls (black) and COVID-19 hospitalized controls (grey).