Letter to the Editor

Akira Endo*

Department of Infectious Disease Epidemiology, London School of Hygiene & Tropical Medicine, London, UK.

* Correspondence to: akira.endo@lshtm.ac.uk

Dear Editor,

I read with interest the recent study by Yokota et al. [1] on polymerase chain reaction (PCR) tests for SARS-CoV-2 between nasopharyngeal swab (NPS) and saliva samples. The authors used a Bayesian latent-class model to simultaneously estimate the sensitivity and specificity of PCR tests along with the unknown prevalence in the study cohorts (contact tracing and airport screening), and reported the sensitivity of 86% and 92% respectively for NPS and saliva samples and over 99.9% specificity for both tests. They also reported high concordance between them and concluded that saliva samples can replace the conventional NPS. Compared with NPS, saliva tests are low-invasive and requires less involvement of healthcare professionals. The equivalent performance of saliva tests, also suggested in other studies [2–5], has a profound impact on the ongoing public health planning against COVID-19 where efficient and safe testing protocols are of paramount importance.

However, I noticed that the authors' model has limitations due to the following two conditions, which rendered the reported sensitivity of 86% and 92% merely reflective of positive agreement rates (PARs) [6] between NPS and saliva tests.

- NPS and saliva test results are assumed to be independent although the authors showed that their Ct values are highly correlated (with a Kendall's W [7] of 0.87).
- The prevalence p was jointly estimated with other parameters, which caused the sensitivity almost solely determined by PAR in the data.

I showed that posterior distributions of PAR obtained from the original data are in effect identical to those of sensitivity obtained from the full dataset; NPS: 86% (90% credible interval 77-93); saliva: 90% (82-96) (Figure 1A). Almost identical distributions were also obtained from data excluding individuals testing negative for both tests ('negative-negatives', 1872/1924 participants); NPS: 88% (79-94) and saliva: 91% (83-96). These indicate that the original sensitivity estimates almost solely relied on individuals receiving at least one positive result (47 from contact tracing and 5 from airport screening data) and that the rest of the dataset was barely informative, which was caused by the latent variable *p* freely optimised reflecting PARs and the independent assumption. By using simulation, I also showed that, in the presence of high correlation between (hypothetical) viral loads in NPS and saliva samples (Pearson's correlation 0.8), I can reproduce the observed contact tracing data and Kendall's *W* (0.86), which nonetheless suggest lower sensitivity (62 % for NPS and 67% for saliva) than the original study (Figures 1B and 1C).

These results highlight the inherent limitations of the original study relying on potentially intercorrelated single-point PCR tests without external validation of true infection statuses. Sensitivity and PAR correspond under the independence assumption, and this study estimated sensitivity based on that assumption. While I support the finding of concordance between NPS and saliva tests, the interpretation of their sensitivity estimates warrants caution. In practice, multiple factors can lead to viral loads of both NPS and saliva below the detectable limit, including samples taken too early or late in the infection course or simply of poor quality. PCR will remain to be the de facto gold-standard test; the possibility of a less-invasive sampling method without apparent loss in performance is an invaluable finding. However, any tests should be used with a proper understanding of their limitations in performance: e.g. how frequently and, more importantly, when they can fail, which this study was unfortunately not sufficient to inform.

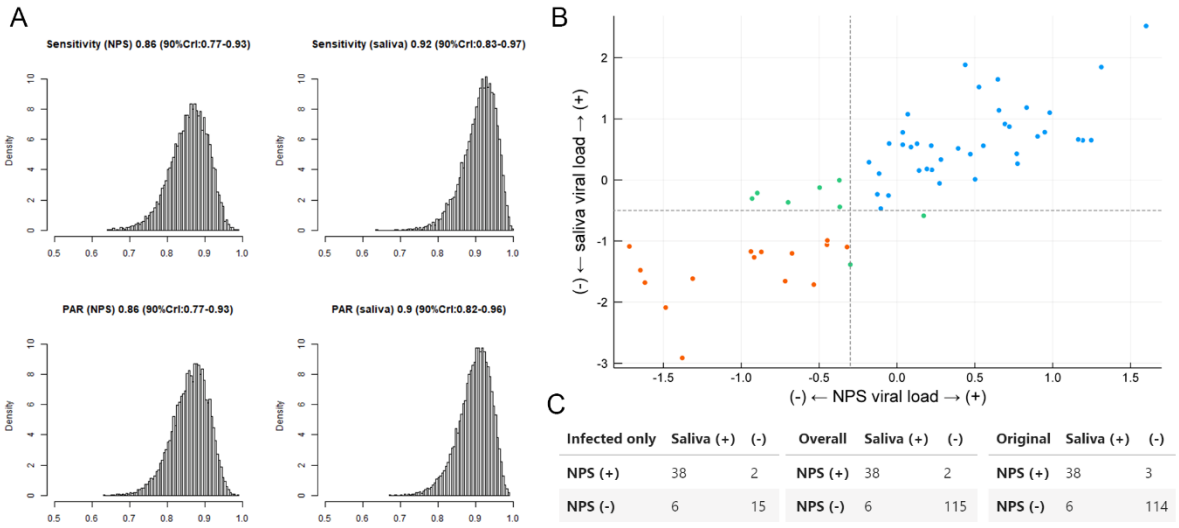

Figure 1. Posterior distributions and simulated dataset. (A) Estimated sensitivity and positive agreement rate (PAR) for nasopharyngeal swab (NPS) and saliva tests. Top panels: posterior distribution of sensitivity reproduced from the original method. Bottom panels: PARs obtained from only saliva-positive samples for NPS and NPS-positive samples for saliva. The histograms show 10,000 posterior samples. (B) Simulated hypothetical viral loads in NPS and saliva samples. Assuming a multivariate-normal distribution with a Pearson's correlation of 0.8, 61 samples were drawn and plotted. Detection limits for NPS and saliva tests were assumed to be -0.3 and -0.5, respectively (denoted by dashed lines) and each sample was coloured according to the test results (blue: both positive; green: one positive and one negative; red: both negative). Kendall's *W* for samples with at least one positive test (viral load above the detection limit) was 0.86. (C) Simulated and original data in 2-by-2 tables. The simulated viral loads were classified as positive/negative according to the assumed detection limits. The left table: simulated data restricted to truly infected individuals. The middle table: overall simulated data combined with 100 true negative individuals (assumed to be negative for both tests). The right table: the original contact tracing data.

Acknowledgements

AE is financially supported by The Nakajima Foundation.

Notes

Code availability: detailed methods and replication codes are available on GitHub

(https://github.com/akira-endo/reanalysis_Yokota2020).

Potential conflict of interest: AE received a research grant from Taisho Pharmaceutical Co., Ltd.

1. Yokota I, Shane PY, Okada K, et al. Mass screening of asymptomatic persons for SARS-CoV-2 using saliva. Clin Infect Dis **2020**; Available at: https://academic.oup.com/cid/advance-article/doi/10.1093/cid/ciaa1388/5911780.
2. Williams E, Bond K, Zhang B, Putland M, Williamson DA. Saliva as a noninvasive specimen for detection of sars-cov-2. J. Clin. Microbiol. 2020;
3. Pasomsub E, Watcharananan SP, Boonyawat K, et al. Saliva sample as a non-invasive specimen for the diagnosis of coronavirus disease 2019: a cross-sectional study. Clin Microbiol Infect **2020**; Available at: https://linkinghub.elsevier.com/retrieve/pii/S1198743X20302780.
4. Kojima N, Turner F, Slepnev V, et al. Self-Collected Oral Fluid and Nasal Swabs Demonstrate Comparable Sensitivity to Clinician Collected Nasopharyngeal Swabs for Covid-19 Detection. medRxiv **2020**; :2020.04.11.20062372. Available at: http://medrxiv.org/content/early/2020/04/15/2020.04.11.20062372.abstract.
5. Wyllie AL, Fournier J, Casanovas-Massana A, et al. Saliva or Nasopharyngeal Swab Specimens for Detection of SARS-CoV-2. N Engl J Med **2020**; 383:1283–1286. Available at: http://www.nejm.org/doi/10.1056/NEJMc2016359.
6. Obermeier P, Muehlhans S, Hoppe C, et al. Enabling Precision Medicine With Digital Case Classification at the Point-of-Care. EBioMedicine **2016**; 4:191–196. Available at: https://linkinghub.elsevier.com/retrieve/pii/S2352396416300068.
7. Kendall MG, Smith BB. The Problem of *m* Rankings. Ann Math Stat **1939**; 10:275–287. Available at: http://projecteuclid.org/euclid.aoms/1177732186.