$$\begin{aligned}
E[T_I|T_I \leq t] &= \frac{1}{1 - e^{-\lambda_I t}} \int_0^t u \lambda_I e^{-\lambda_I u} du \\
&= \frac{1}{1 - e^{-\lambda_I t}} \left[-u e^{-\lambda_I u} - \frac{1}{\lambda_I} e^{-\lambda_I u} \right]_{u=0}^{u=t} \\
&= \frac{1}{1 - e^{-\lambda_I t}} \left[-t e^{-\lambda_I t} - \frac{1}{\lambda_I} e^{-\lambda_I t} + \frac{1}{\lambda_I} \right] \\
&= \frac{1}{\lambda_I} - \frac{t e^{-\lambda_I t}}{1 - e^{-\lambda_I t}}.
\end{aligned}$$

Thus, the expected person-time at risk for each unvaccinated

individual by study time t is

$$\begin{aligned}
E[T_I|T_I \leq t] \Pr(T_I \leq t) + t \Pr(T_I > t) \\
&= \left[\frac{1}{\lambda_I} - \frac{t e^{-\lambda_I t}}{1 - e^{-\lambda_I t}} \right] (1 - e^{-\lambda_I t}) + t e^{-\lambda_I t} \\
&= \frac{(1 - e^{-\lambda_I t})}{\lambda_I} - t e^{-\lambda_I t} + t e^{-\lambda_I t} \\
&= \frac{(1 - e^{-\lambda_I t})}{\lambda_I}.
\end{aligned}$$

A parallel derivation is used for vaccinated individuals.

© The Author(s) 2019. Published by Oxford University Press on behalf of the Johns Hopkins Bloomberg School of Public Health. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com.

THE AUTHORS REPLY

We thank Dr. Dean for her comments (1). The misleading appearance of time-dependent vaccine effectiveness (VE) is a problem under diverse study designs that do not ascertain time to infection or do not ascertain all infections (2). In our article (3), we highlighted that VE estimates under the test-negative design (TND) may be exposed to time-dependent bias when the vaccine effect on infection is nonnull. Differential depletion of person-time at risk among vaccinated and unvaccinated persons who acquire natural immunity through undetected infections may create the appearance of waning immunity.

Dr. Dean proposes a modification to the TND aiming to correct for this bias in VE estimates by analyzing a subset of data sampled from vaccinated and unvaccinated persons' influenza-susceptible time at risk (1). Censoring observations from persons with a known history of influenzavirus infection is a compelling design-level correction for the problems created by differential preinfection person-time among the vaccinated and unvaccinated. We believe strategies such as this should be sought urgently to salvage valid inferences from TND studies that present numerous other advantages, such as efficiency and the ability to correct for health-care-seeking bias.

Nonetheless, we caution that there are feasibility considerations to take into account for the specific approach suggested by Dean (1). The proposed strategy requires that all influenzavirus infections (or at least those preceding a test-negative visit) be observed by researchers—in other words, that each influenzavirus infection result in a person's seeking care from the study team, receiving an influenza diagnostic test, and having his or her infection status recorded. However, approximately 65%–85% of influenza infections are clinically inapparent (4); even if symptoms occur, it remains the case that only a minority of people seek medical attention and an even smaller fraction receive diagnostic tests (5). While the proportion varies by age and setting, previous studies have suggested that surveillance of and testing for acute respiratory illnesses will capture only 1 out of every 79 (95% confidence interval: 47, 148) cases of symptomatic influenza not requiring hospitalization (6). The

probability of detecting a previous influenzavirus infection, given that it occurred, is further reduced in TND studies that enroll only inpatients.

Various study designs have been undertaken to create data-collection streams that are amenable to such analyses for diseases with a high proportion of asymptomatic or mild infections. In the most notable instance, birth cohort studies coupling regular (weekly or biweekly) testing of asymptomatic stool for rotavirus infection with active surveillance of clinical gastroenteritis have yielded insights into the protective effect of previous rotavirus infection against future infection and disease (e.g., see Velázquez et al. (7)), but such designs necessitate prospective follow-up of participants, which the TND aims to circumvent.

What information, then, could enable researchers to censor naturally immune person-time in the retrospective context of TND studies, as Dean proposes (1)? We propose that testing for immunological evidence of recent influenza infection may present an attractive option. For instance, an assay which indicates recent infection (e.g., within the same season) and which discriminates between the response to natural infection and vaccination could provide a basis for censoring test-negative observations believed to have occurred after a naturally immunizing influenza infection. If it can be assumed that recent influenza infection deterministically prevents future infections, the non-censored observations will presumably represent samples from people's influenza-susceptible time at risk. Collection of sera for ascertainment of previous infection would impose less of a burden than prospective follow-up, although serological criteria for ascertaining recent infection should be considered carefully in the absence of a baseline specimen. The feasibility of this and other strategies, and any impact on inferences regarding influenza VE, should be explored in future studies.

ACKNOWLEDGMENTS

Conflict of interest: none declared.

REFERENCES

1. Dean NE. Re: "Measurement of vaccine direct effects under the test-negative design" [letter]. *Am J Epidemiol*. 2019;188(4):806–807.
2. Lopman BA, Pitzer VE. Waxing understanding of waning immunity. *J Infect Dis*. 2018;217(6):851–853.
3. Lewnard JA, Tedijanto C, Cowling BJ, et al. Measurement of vaccine direct effects under the test-negative design. *Am J Epidemiol*. 2018;187(12):2686–2697.
4. Leung NH, Xu C, Ip DK, et al. The fraction of influenza virus infections that are asymptomatic: a systematic review and meta-analysis. *Epidemiology*. 2015;26(6):862–872.
5. Reed C, Kim IK, Singleton JA, et al. Estimating influenza illnesses and hospitalizations averted by vaccination—United States, 2013–14 influenza season. *MMWR Morb Mortal Wkly Rep*. 2014;63(49):1151–1154.
6. Reed C, Angulo FJ, Swerdlow DL, et al. Estimates of the prevalence of pandemic (H1N1) 2009, United States, April–July 2009. *Emerg Infect Dis*. 2009;15(2):2004–2007.
7. Velázquez FR, Matson DO, Calva JJ, et al. Rotavirus infection in infants as protection against subsequent infections. *N Engl J Med*. 1996;335(14):1022–1028.

Joseph A. Lewnard¹, Christine Tedijanto², Benjamin J. Cowling³, and Marc Lipsitch² (e-mail: jlewnard@berkeley.edu)

¹ Division of Epidemiology and Biostatistics, School of Public Health, University of California, Berkeley, Berkeley, CA

² Center for Communicable Disease Dynamics, Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA

³ WHO Collaborating Centre for Infectious Disease Epidemiology and Control, School of Public Health, Li Ka Shing Faculty of Medicine, University of Hong Kong, Hong Kong Special Administrative Region, China

DOI: 10.1093/aje/kwz018; Advance Access publication: January 25, 2019

© The Author(s) 2019. Published by Oxford University Press on behalf of the Johns Hopkins Bloomberg School of Public Health. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com.

RE: "QUANTIFICATION OF HUMAN MICROBIOME STABILITY OVER 6 MONTHS: IMPLICATIONS FOR EPIDEMIOLOGIC STUDIES"

We read with keen interest the paper by Sinha et al. (1), in which they evaluated phylum-level microbiome stability across pairs of samples obtained 6 months apart from 3 separate studies. All 3 studies had replicate fecal samples ($n = 44$ –116 replicates), 2 studies had replicate saliva samples ($n = 42$ –94 replicates), and 1 study had replicates from 4 additional body sites. Sample collection, processing, and 16S ribosomal RNA sequencing methods differed between studies. The authors used intraclass correlation coefficients (ICCs) to evaluate microbiome stability, comparing phyla relative abundances, α -diversity measures (number of operational taxonomic units, Chao1, Shannon Index), and β -diversity measures (principal coordinates analysis (PCoA) coordinates from unweighted and weighted UniFrac distances) across the 2 time points. An ICC of 0 was interpreted to indicate no stability, whereas an ICC of 1 was taken to reflect perfect stability across replicate samples. Adjusting for age and sex (and sequencing center in one study), ICC values for feces (0.00–0.66 (median, 0.44)) and saliva (0.16–0.60 (median, 0.42)) were lower and more variable for relative abundance and α -diversity measures. In contrast, ICC values for β -diversity measures were consistently larger, exemplified by first principal coordinates analysis (PCoA1) (feces: 0.54–0.76 (median, 0.63); saliva: 0.44–0.84 (median, 0.71)). The investigators concluded that UniFrac distances are the preferred metric when assessing microbial community stability across repeated measurements.

The work by Sinha et al. highlights several key methodologic considerations that are important in the rapidly expanding field of microbiome research. Understanding the temporal stability of the microbiome is crucial when attempting to characterize the relationship between host characteristics or therapeutic

interventions and changes in microbiome composition and diversity. At present, there is a lack of standard methods to quantify characteristics of microbiome communities, cross-sectionally and longitudinally, that take into account the complexity of such data. Microbiome data has a compositional structure that can yield spurious correlations when statistics, like ICCs, are naively applied (2, 3). Aitchison and Greenacre (4, 5) developed log-ratio transformations that relieve spurious correlations and allow meaningful comparisons of subcompositions, a condition necessary for reproducibility, both temporally and between labs.

To further illustrate this important methodologic consideration, we gathered the same Human Microbiome Project (HMP) (6, 7) data set used in the analysis of Sinha et al. We identified subjects with samples collected at 2 timepoints from stool, saliva, and anterior nares body sites. After performing a centered log-ratio transformation, which accounts for the compositional structure of the data, we followed the analysis in Sinha et al. The between-individual (σ_b^2) and technical/temporal variabilities (σ_e^2) were estimated using linear mixed-effects models. ICCs were then calculated as the ratio of between-individual variability over total variability ($\text{ICC} = \sigma_b^2 / (\sigma_b^2 + \sigma_e^2)$). ICCs were bootstrapped 1,000 times with replacement to identify deviation from the mean. Based on our results, a different conclusion could be reached from that reported. We observed that across the 2 replicates obtained 6 months apart, and for all 3 body sites, the Human Microbiome Project data shows moderate to good consistency, with all average ICCs above 0.5 (Figure 1). Our conclusion is that taking the compositional nature of microbiome data into account increases the estimate of temporal stability. Relative abundances and UniFrac measures do not account for