Statistics in Biosciences (2020) 12:263–266
https://doi.org/10.1007/s12561-020-09296-x

Introduction to Special Issue on 'Statistical Methods for HIV/AIDS Research'

Ying Qing Chen[1]

Published online: 19 October 2020

The acquired immunological deficiency syndrome (AIDS), caused by the human immunodeficiency virus (HIV), is deadly and has had a devastating impact on public health globally. It has claimed almost 33 million lives by mid-2020, and there were 38 million people estimated living with HIV by the end of 2019 [1]. In the U.S. alone, where the first case of AIDS was recognized in 1981, most recent statistics reported in 2019 that by 2018 about 700,000 people died of AIDS since then, and 1.1 million people were living with HIV [2].

For the past four decades, ground-breaking scientific advances have led to great progresses via epidemiologic studies, clinical trials, and implementation projects in HIV/AIDS research, e.g., discovering the HIV, developing highly efficacious therapeutic regimens, and inventing effective prevention tools. Along with these progresses, novel statistical methods and theory also have played pivotal roles. They were developed and applied to almost every aspect of the HIV/AIDS research, including but not limited to study design consideration and challenging data analysis.

Nowadays, as HIV/AIDS have become more manageable, ending HIV/AIDS as a global pandemic is no longer inconceivable. Key factors leading to such an optimism are the broad use of highly effective combination antiretroviral therapy (cART) for treatment and pre-exposure prophylaxis (PrEP), in addition to screening of blood product, condom use, and voluntary medical male circumcision, among others [3]. Nevertheless, to realize the optimism still faces a daunting reality—in 2018 only, 1.7 million people became newly infected and 770,000 people died of AIDS-related illnesses [1]. There is much yet to be done, e.g., developing efficacious vaccine and strengthening implementation programs supported by the President's Emergency Plan for AIDS Relief (PEPFAR).

Like what they have contributed toward the past successes in the HIV/AIDS research, additional novel statistical methods and theory are in need to deal with

✉ Ying Qing Chen
yqchen@fredhutch.org

[1] Vaccine and Infectious Disease Division and Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA

new challenges in the effort of ending the pandemic. In this special issue, we thus have assembled eleven thought-provoking articles to address a wide spectrum of statistical issues arising from the research for HIV/AIDS-related epidemiologic and clinical studies.

Brookmeyer and Wu [4] developed a novel method to estimate the proportion of living HIV-infected patients that have been diagnosed, in order to help with tracking the HIV epidemic and planning the needed resources. However, a methodological challenge exists as those who are not diagnosed tend to be "hidden" and hence are not directly observable. Their method is innovative by utilizing the population subject to mandatory tests to estimate the size of the undiagnosed HIV infections. It does not require historical HIV/AIDS surveillance data or biomarkers such as the patients' CD4 counts.

Sheng et al. [5] compared the HIV prevalence measured by antenatal clinic (ANC) unlinked anonymous testing (UAT) and ANC routing testing (RT) among 15 countries where both UAT and RT data were available for the years in common, and studied a novel linear mixture model to estimate the RT-to-UAT calibration parameter. The method studied in this paper is timely, as new guidance released by the World Health Organization (WHO) and UNAIDS recommended countries to transition from UAT- to RT-tracking HIV prevalence, because the RT-tracking was deemed to be more consistent and economic.

Zhuang and Chen [6] proposed a new measure, the so-called 'population surrogacy fraction of treatment effect' or simply the ρ-measure, in the setting of clinical trials, to supplement existing statistical measures that would be used to measure and rank the relative surrogacy of potential biomarkers quantitatively. The new ρ-measure has an appealing population impact interpretation, and is applied to the HIV Prevention Trial Network 052 Study, a landmark trial for the HIV treatment-as-prevention strategy, to assess the relative surrogacy of an HIV-infected patient's viral load used as the surrogate biomarker for his or her HIV-uninfected partner's possible seroconversion.

Huang et al. [7] proposed two novel simulation approaches to generating event times following the Cox proportional hazards models with both time-independent covariates and continuous cyclic and piecewise time-dependent covariates. One was based on simulating survival data under a single-dose regimen first before data are aggregated over multiple dosing cycles, and the other was based on simulating survival data directly under a multiple-dose regimen. Three simulation experiments were carried out with an application to two Antibody Mediated Prevention (AMP) Phase 2b efficacy trials.

Zheng and Chen [8] proposed a shape-invariant hazard regression model that allows to estimate a multiplicative treatment effect with adjustment of covariates that have non-multiplicative effects. This new model can also be used for risk prediction. When applied to the HIVNET 012 Study of single-dose nevirapine in prevention of mother-to-child transmission, the new model showed that the single-dose nevirapine treatment would improve 18-month infant survival significantly with appropriate adjustment of the maternal CD4 counts and the virus load.

Huang and Tarkhan [9] proposed some novel algorithms for selecting immune response endpoints to be used in regimen down-selection, based on importance

weights assigned to individual endpoints and their correlation structure. Via extensive simulation studies, it was shown that pre-selection of endpoints can substantially improve performance of the subsequent regimen down-selection process. The algorithms were demonstrated for down-selection of vaccine regimens in HIV vaccine trials.

Saegusa et al. [10] considered variable selection for threshold regression, an alternative to the Cox proportional hazards regression model when the proportional hazards assumption is not met. It extends the broken adaptive ridge (BAR) method, originally designed for variable selection for one regression function, to simultaneous variable selection for both regression functions needed in the threshold regression model. The proposed variable selection method was used to identify risk factors for drug non-adherence in an HIV PrEP study with oral regimens.

Ren et al. [11] studies model-based and design-based approaches for the analysis of data arising from a stepped wedge randomized design that may be used to evaluate novel HIV prevention and treatment tools. Different scenarios were assessed and compared in robustness, efficiency, Type I error rate, and power for the leading analytical options, including generalized estimating equations (GEE) and linear mixed model (LMM)-based approaches.

Qian et al. [12] proposed semiparametric estimators for the survival probability in double-sampling designs by generalizing the deductive and computerizable estimation approach. In particular, the semiparametric estimators were based on a discretized support structure, which approximates the possibly continuous observed data distribution and circumvents the derivation of the mixture distribution. The proposed estimators were used to estimate the mortality rate in a double-sampling design component of the PEPFAR program.

Benkeser et al. [13] discussed the motivation for a sequentially multi-arm randomized trial (SMART) design to evaluate different combinations of mHealth intervention aiming to reduce risk behaviors among HIV-negative transgender youths and young adults. Novel robust methods were proposed to analyze the SMART designs as well.

Janes et al. [14] considered four potential trial designs for future trials of candidate HIV vaccines, with a focus on accommodating the newest addition to the prevention package—ART-based oral PrEP, as the newly added non-vaccine prevention modalities were anticipated to yield reduction in HIV incidence, and trial participants would have individual choices among these modalities. The Run-In designs, enrolling all but randomizing only those who would decline PrEP or show inadequate adherence to PrEP after one or multiple run-in period, were advocated.

As summarized above, the 11 articles of this Special Issue cover a wide range of research topics in HIV/AIDS epidemiology, clinical trials, implementation, and statistical designs. Although their statistical considerations were mainly motivated and aimed to be applied in HIV/AIDS research, their applications can certainly go beyond the field of HIV/AIDS research. With the ongoing COVID-19 pandemic caused by the SARS-CoV-2, the fundamental ideas and principles presented by these papers shall apply to the COVID-19 research as well.

References

1. The World Health Organization (2020) HIV/AIDS fact sheets. https://www.who.int/news-room/fact-sheets/detail/hiv-aids. Accessed 15 Oct 2020
2. The US CDC (2019) HIV Surveillance Supplemental Report on Estimated HIV Incidence and Prevalence in the United States 2014–2018. https://www.cdc.gov/hiv/pdf/library/reports/surveillance/cdc-hiv-surveillance-supplemental-report-vol-25-1.pdf. Accessed 15 Oct 2020
3. Fauci AS, Lane HC (2020) Four decades of HIV/AIDS - much accomplished, much to do. N Engl J Med 383:1–4
4. Brookmeyer R, Wu Z (2019) A method for estimating the proportion of HIV infected persons that have been diagnosed and application to China. Stat Biosci. https://doi.org/10.1007/s12561-019-09240-8
5. Sheng B, Eaton JW, Mahy M, Bao L (2020) Comparison of HIV prevalence among antenatal clinic attendees estimated from routine testing and unlinked anonymous. Stat Biosci. https://doi.org/10.1007/s12561-020-09265-4
6. Zhuang R, Chen YQ (2019) Measuring surrogacy in clinical research with an application to studying surrogate markers for HIV treatment-as-prevention. Stat Biosci. https://doi.org/10.1007/s12561-019-09244-4
7. Huang Y, Zhang Y, Zhang Z, Gilbert PB (2019) Generating survival times using Cox proportional hazards models with cyclic and piecewise time-varying covariates. Stat Biosci. https://doi.org/10.1007/s12561-020-09266-3
8. Zheng C, Chen YQ (2019) On a shape-invariant hazard regression model with application to an HIV prevention study of mother-to-child transmission. Stat Biosci. https://doi.org/10.1007/s12561-019-09260-4
9. Huang Y, Tarkhan A (2020) Methods for feature selection in down-selection of vaccine regimens based on multivariate immune response endpoints. Stat Biosci. https://doi.org/10.1007/s12561-020-09275-2
10. Saegusa T, Ma T, Li G, Chen YQ, Lee M-LT (2020) Variable selection in threshold regression model with applications to HIV drug adherence data. Stat Biosci. https://doi.org/10.1007/s12561-020-09284-1
11. Ren Y, Hughes JP, Heagerty PJ (2019) A simulation study of statistical approaches to data analysis in the stepped wedge design. Stat Biosci. Original: https://doi.org/10.1007/s12561-019-09259-x. Correction: https://doi.org/10.1007/s12561-020-09289-w
12. Qian T, Frangakis C, Yiannoutsos C (2019) Deductive semiparametric estimation in double-sampling designs with application to PEPFAR. Stat Biosci. https://doi.org/10.1007/s12561-019-09262-2
13. Benkeser D, Horvath K, Reback CJ, Rusow J, Hudgens M (2020) Design and analysis considerations for a sequentially randomized HIV prevention trial. Stat Biosci. https://doi.org/10.1007/s12561-020-09274-3
14. Janes H, Zhu Y, Brown ER (2020) Designing HIV vaccine efficacy trials in the context of highly effective non-vaccine prevention modalities. Stat Biosci. https://doi.org/10.1007/s12561-020-09292-1