

International Journal of Population Data Science

Journal Website: www.ijpds.org



Quantifying bias due to linkage error: methods and application to South Africa's NHLS National HIV Cohort

Evelyn Lauren¹, Dickman Gareta², Khumbo Shumba³, Mazvita Muchengeti⁴, William MacLeod⁵, Matthew Fox⁶, Wendy Stevens⁷, Tshepo Molapo⁸, Musa Manganye⁸, and Koleka Mlisana⁹

¹Health Economics and Epidemiology Research Office, Department of Internal Medicine, School of Clinical Medicine, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa; Department of Biostatistics, Boston University School of Public Health, Boston, USA

²Africa Health Research Institute, Somkhele, South Africa; Institute of Social and Preventive Medicine, University of Bern, Bern, Switzerland; Graduate School for Health Sciences, University of Bern, Bern, Switzerland

³Health Economics and Epidemiology Research Office, Department of Internal Medicine, School of Clinical Medicine, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

⁴National Institute of Communicable Diseases, National Health Laboratory Service, Johannesburg, South Africa

⁵Health Economics and Epidemiology Research Office, Department of Internal Medicine, School of Clinical Medicine, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa; Department of Global Health, Boston University School of Public Health, Boston, USA

⁶Health Economics and Epidemiology Research Office, Department of Internal Medicine, School of Clinical Medicine, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa; Department of Global Health, Boston University School of Public Health, Boston, USA; Department of Epidemiology, Boston University School of Public Health, Boston, USA

⁷National Health Laboratory Service, Johannesburg, South Africa; Diagnostic Innovations Hub, University of Witwatersrand, Johannesburg, South Africa

⁸HIV/AIDS and STIs Cluster, National Department of Health, Pretoria, South Africa

⁹National Health Laboratory Service, Johannesburg, South Africa; School of Laboratory Medicine and Medical Sciences, University of KwaZulu Natal, Durban, South Africa

Background

Linkage error can bias downstream estimates. South Africa's National Health Laboratory Service (NHLS) National HIV Cohort, created through probabilistic linkage of routine laboratory records, offers a rare opportunity to examine viral suppression durability at scale. We propose a sensitivity analysis approach to assess robustness of findings to linkage error.

Methods

People living with HIV (PLHIV) aged 15–59 years in the NHLS cohort were followed for 18 months from their first viral load (VL) <200 copies/mL between July 2021–June 2022. We assessed: 1) 12-month monitoring (any VL 6–18 months post-baseline), and 2) 12-month transmission risk using WHO thresholds: <200 copies (zero risk), 200–999 copies (minimal risk), and ≥1000 copies (elevated risk). To assess robustness to linkage error, we repeated analyses in a subset of high-confidence patient clusters – clusters stable across linkage thresholds. We applied inverse probability weighting using demographics for representativeness. Assuming cluster selection is independent of outcomes, this yields minimal-error estimates for comparison with naïve estimates.

Results

The cohort was linked with 90.4% sensitivity (Sen) and 91.9% positive predictive value (PPV). We identified 3,584,204 PLHIV with baseline VL <200 copies, of which 80% were monitored in 12 months; of these, 91% had no transmission risk, 5% minimal risk, and 4% elevated risk. In the high-confidence subset (3,072,758 individuals), 78% were monitored in 12 months with 92% having no transmission risk, 5% minimal risk, and 3% elevated risk.

Conclusion

Viral suppression is highly durable among PLHIV in South Africa. Sensitivity analyses confirmed findings were robust despite non-trivial linkage error rates.

