Response to reviewers for "Spatio-temporal estimates of HIV risk group proportions for adolescent girls and young women across 13 priority countries in sub-Saharan Africa"

Corresponding author: Adam Howes (ath19@ic.ac.uk)

Contents

Reviewer 1

We thank the reviewer for their helpful comments regarding the statistical modelling.

1. In the manuscript, explain why you use the INLA not WINBUGS coding? The multinomial regression could be modeled in WINBUGS directly.

The reviewer is right to note that multinomial logistic regression models can be implemented in probabilistic programming languages like `WinBUGS` directly. However, for this application, Markov chain Monte Carlo approaches would be prohibitively expensive. For this reason, we chose to use integrated nested Laplace approximations via `R-INLA`, which have been shown to have comparable accuracy for latent Gaussian models in the realistic, non-asymptotic regime. We have added the following text to the manuscript to clarify this point: *Insert added text here.*

2. Is there no other potential covariate that could be used for better modeling?

Many of the covariates which one might expect to be most predictive of risk group proportions are themselves difficult to accurately measure, and can only lead to modest at best improvements in model performance. For example, despite the case for their being a clear link between the "proportion clients of FSW" covariate and the "proportion FSW" outcome we found only marginal benefits to inclusion.

3. Please explain sub-national effect more clearly. Why and how you used it?

Using district-level spatial random effects allowed us to account for subnational variation in risk group proportions. We considered both spatially unstructured (IID) and spatially structured (Besag) random effects, which can be implemented in `R-INLA` by setting `model = "iid"` or `model = "besag"` respectively.

4. Why the interaction term for spatiotemporal effect didn't consider in the modeling framework?

```
#' TODO Calculate the average Sobol index for survey random effects
sobol <- 1
```

Fitting each country individually, the proportion of variance (Sobol' index) attributable to the temporal (survey) random effects was on average 1/% (Figure B.3). This is corroborated by the lack of temporal trends in Figures B.5 through B.17 which show the modelled and direct estimates for each country individually, as well as the fact that unstructured (IID) rather than structured (AR1) temporal random effects were preferred in the model selection (Figure A.1).

We had wanted to include the spatiotemporal interaction terms regardless, first for statistical completeness, and second to demonstrate how complex three-way[1] interactions can be specified in `R-INLA` by careful use of the `group` and `replicate` options. However, we found that including these terms substantially increased the computational burden of the model when fitting all countries jointly, so in the end decided to exclude them.

[1] As all random effects in the model are interacted with category.

Reviewer 2

We thank the reviewer for their kind comments.

> My little concern is about using different data from UNAIDS Key Population Atlas apart from the DHS, which is the may source data for the analyses. I believe the two variants of data are based on different designs, and combining them may not result in dependable results. It would have been more attainable if the UNAIDS data had been used in their sensitivity analysis to confirm the results from the DHS data.

For the FSW risk group, we used age-disaggregated (Figure A.4) national-level estimates from Stevens et al. (2022) to inform the level, and household survey data to inform subnational variation. Estimates of hidden populations like FSW from household surveys have significant limitations due in part to stigma around disclosing membership, as well as potential for not being included in the sampling frame (Abdul-Quader, Baughman, and Hladik 2014). For this reason, we believe it is more appropriate to calibrate our estimates to Stevens et al. (2022), who as well as including the KP Atlas data integrate data from other FSW population size estimates using a Bayesian mixed effects model.

> Figure 1 is not clear. I recommend that the authors use a table as an alternative visualization.

Table B.3 in the appendix provides an alternative tabulation of the surveys we used, with sample size further broken down by age group. We are unsure on how to make Figure 1 might clearer. Other examples of similar figures in published papers include Giguère et al. (2021) (Figure 1), ...

References

Abdul-Quader, Abu S, Andrew L Baughman, and Wolfgang Hladik. 2014. "Estimating the Size of Key Populations: Current Status and Future Possibilities." *Current Opinion in HIV and AIDS* 9 (2): 107–14.

Giguère, Katia, Jeffrey W Eaton, Kimberly Marsh, Leigh F Johnson, Cheryl C Johnson, Eboi Ehui, Andreas Jahn, et al. 2021. "Trends in Knowledge of HIV Status and Efficiency of HIV Testing Services in Sub-Saharan Africa, 2000–20: A Modelling Study Using Survey and HIV Testing Programme Data." *The Lancet HIV* 8 (5): e284–93.

Stevens, Oliver, Keith Sabin, Sonia Arias Garcia, Kalai Willis, Abu Abdul-Quader, Anne McIntyre, Frances Cowan, et al. 2022. "Estimating key population size, HIV prevalence, and ART coverage for sub-Saharan Africa at the national level."