Identifiers: NCT ID not yet assigned

Unique Protocol ID: WHO-PEN@Scale

Brief Title: Strengthening Primary Healthcare Delivery for Diabetes and Hypertension in Eswatini

Date of document: November 10th 2019

Statistical analysis plan:

We will test the following two null hypotheses for each outcome: i) participants in the intervention arms (arm 2 [PEN] and arm 3 [ePEN] combined) have the same outcome as participants in the standard-of-care arm; and ii) participants in arm 3 (ePEN) have the same outcome as participants in arm 2 (PEN). We will not adjust for multiple hypothesis testing in these analyses because each comparison answers a fundamentally different question. That is, the comparison of the standard-of-care arm with arm 2 and 3 combined answers what the effect was of scaling up the WHO-PEN package in primary care in Eswatini, and the comparison of arm 2 (PEN) with arm 3 (ePEN) answers what the effect was of involving community health workers in the WHO-PEN scale-up compared to not involving community health workers in the scale-up. In secondary analyses, we will also compare the standard-of-care arm to arm 3 (ePEN).

All primary analyses will exclude clusters in the cities of Manzini and Mbabane because we expect a high degree of contamination between the study arms in these two cities given the close proximity of healthcare facilities in these locations. However, we will include these clusters in secondary analyses.

We will use ordinary least squares regression to compare mean HbA1c among adults with diabetes and mean systolic blood pressure among adults with hypertension between the study arms. All regression models will regress the outcome onto an indicator for study arm, and adjust standard errors for clustering at the level of the unit of randomization (the primary healthcare facility and its catchment area). In the primary analysis, we will include participants' socio-demographic characteristics as co-variates, which we do not expect will substantially affect the

point estimates but may well reduce the variance (and thus increase power). In secondary analyses, we will run these regressions without adjusting for participants' socio-demographic characteristics. We will use a significance level of $p<0.05$ for all analyses. As a robustness check, the p-values for the effect estimates obtained from these regressions will be calculated using randomization inference, which is a permutation method that, unlike regression analysis, does not rely on parametric assumptions [1, 2]. Secondary endpoints will be analyzed using the same approach except that we will use modified Poisson regression with a robust error structure for binary outcomes.[3]

All analyses will include a sub-group analysis that compares the effect of the intervention(s) between men and women, whether or not a household was ever visited by a community health worker, whether or not a household was visited by a community health worker in the past 12 months, ten-year age groups, rural versus urban areas, the different regions (Hhohho, Lubombo, Manzini, and Shiselweni), and categories of educational attainment. In addition to disaggregating effects by these sub-groups, we will employ random forest analyses to identify sub-groups that were particularly likely or unlikely to benefit from the intervention. These techniques are a non-parametric data-driven way of identifying sub-groups.[4, 5]

References:

1. Athey S, Imbens G. The Econometrics of Randomized Experiments. arXiv:160700698 [statME]. 2016.
2. Heß S. Randomization inference with Stata: A guide and software. Stata Journal. 2017;17(3):630-51.
3. Zou G. A modified poisson regression approach to prospective studies with binary data. American journal of epidemiology. 2004;159(7):702-6. Epub 2004/03/23. PubMed PMID: 15033648.
4. Athey S, Imbens G. Recursive partitioning for heterogeneous causal effects. Proceedings of the National Academy of Sciences. 2016;113(27):7353-60. doi: 10.1073/pnas.1510489113.
5. Wager S, Athey S. Estimation and Inference of Heterogeneous Treatment Effects using Random Forests. Journal of the American Statistical Association. 2017:1-15. doi: 10.1080/01621459.2017.1319839.