THE LANCET
Global Health

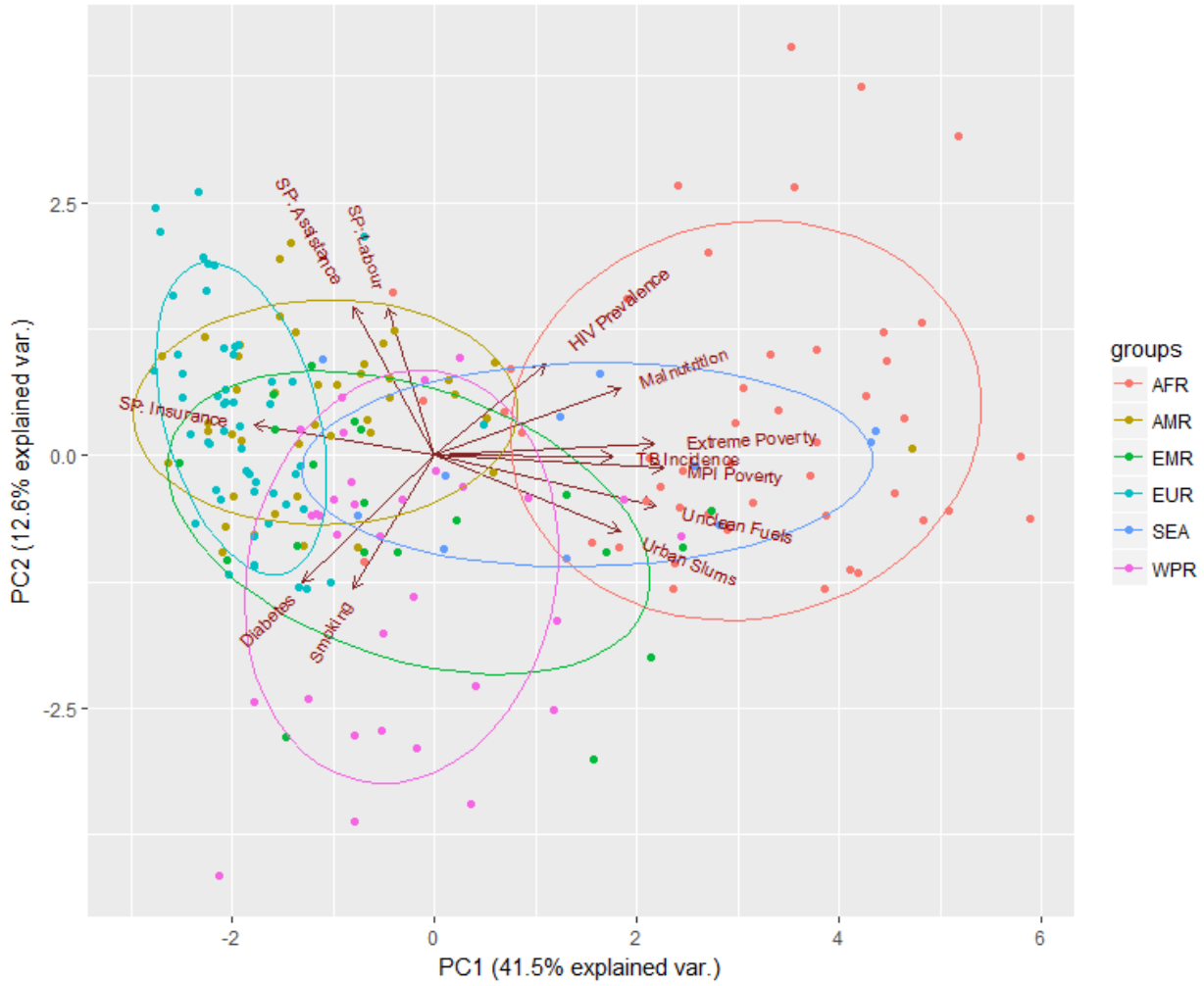
# Supplementary appendix

This appendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

Supplement to: Carter DJ, Glaziou P, Lönnroth K, et al. The impact of social protection and poverty elimination on global tuberculosis incidence: a statistical modelling analysis of Sustainable Development Goal 1. *Lancet Glob Health* 2018; published online March 23. http://dx.doi.org/10.1016/S2214-109X(18)30195-5.

Appendix 1 –PCA

To examine relationships between selected variables, a PCA was performed after conducting a single imputation by region using the regression substitution method. Five imputations of the dataset were generated using multiple imputation by chained equations and the data used to generate this biplot comes from a randomly selected imputed dataset.[20] Smoking and diabetes cluster together and not using clean fuels and living in an urban slum also cluster together, informing a decision to combine variables. Extreme poverty and MPI poverty are also closely related, suggesting they may be collinear. Both are quite closely related to TB incidence, as expected. Malnutrition is also quite strongly related to TB incidence. Social insurance is negatively correlated with TB incidence as expected, and closely associated with regions with high and middle income countries.



647
648 **Figure A**. Biplot of the first two principal components of variables in the conceptual framework. Closeness of the
649 arrows indicates the degree of correlation between the variables. The direction of the arrows points in the
650 direction of increasing values of that variable. Ellipses represent groupings by WHO region.
651
652
653

654 Appendix 2 - Missing Data
655
656 Table A presents the level of missingness in the data prior to the multiple imputation procedure. The
657 multiple imputation procedure works by running iterated regression analyses, taking one variable as the
658 outcome variable and the remaining variables in the dataset as the predictor variables. Doing this for all
659 variables results in a set of predicted values (imputed values) for each missing value [20]. Repeating this
660 procedure multiple times provides multiple datasets with no missing data which are pooled together to
661 perform analyses.
662
663 The advantage of undertaking this procedure is that the imputed values for a given country closely
664 resemble values for countries with similar characteristics in the other variables, and are thus more likely
665 to accurately reflect possible values of the missing data. Doing this multiple times ensures convergence
666 on a reasonable value. In this dataset, missingness on social insurance and missingness on absolute
667 poverty was not associated with any of the four proximal TB risk factors, nor was it associated with WHO
668 region.
669
670
671

| Subtarget Indicator Measures | % missing data |
|---|---|
| SDG 1.1: Absolute Poverty | 42.0% |
| SDG 1.2: Multidimensional Poverty Index (prevalence) | 52.2% |
| SDG 1.3: Labour Market Social Protection | 71.7% |
| SDG 1.3: Social Assistance | 48.3% |
| SDG 1.3: Social Insurance | 49.3% |
| | |
| Poverty Subconstruct Measures | % missing data |
| Malnourished % | 0.01% |
| HIV Prevalence | 22.9% |
| Health Behaviours | 39.0% |
| Housing Quality | 54.6% |

672 **Table A.** Proportion of missing data for each subtarget indicator and for each measure of poverty. Missingness in
673 the context of these data is that there was no recorded value after 2010.

Appendix 3 – Mediators

A simple analysis for mediation was performed by adjusting the estimate for the linear relationship between SDG 1 and TB for each of the risk factors in turn. Evidence for mediation is determined by the degree of change in the relationship between SDG 1 and TB is adjusted for with the risk factor under examination. All risk factors will show a degree of mediation for extreme poverty and social insurance.

| Indicator | | Crude | Adjusted | Mediation? |
|---|---|---|---|---|
| SDG 1.1: Extreme Poverty | Malnutrition | **3.23*** ** | **2.29*** ** | Moderate |
| | HIV | **3.23*** ** | **1.99*** ** | Moderate |
| | Health Behaviours | **3.23*** ** | **4.67*** ** | Moderate |
| | Housing Quality | **3.23*** ** | 2.43 | Moderate |
| | | | | |
| SDG 1.2: MPI Prevalence | Malnutrition | **1.94*** ** | **1.17*** | Weak |
| | Health Behaviours | **1.94*** ** | **2.32** ** | Weak |
| | Housing Quality | **1.94*** ** | -1.44 | Strong |
| | | | | |
| SDG 1.3: Social Assistance | Malnutrition | -0.81 | -0.49 | Weak |
| | Housing Quality | -0.81 | 0.38 | Moderate |
| | | | | |
| SDG 1.3: Social Insurance | Malnutrition | **-3.73*** ** | **-2.79** ** | Moderate |
| | HIV | **-3.73*** ** | **-2.00** ** | Moderate |
| | Health Behaviours | **-3.73*** ** | **-4.98*** ** | Moderate |
| | Housing Quality | **-3.73*** ** | -1.90 | Strong |

Stars indicate significant difference of coefficient from 0: ***p < 0.001, ** p < .01, * p < .05

682 **Table B.** Effect estimates for SDG indicators and TB incidence adjusting for a sole mediator, where included
683 mediators are associated with the SDG indicator in question and TB incidence (Table X; Figure B). Pathways are
684 assumed independent.
685
686
687

Appendix 4 – Income Analyses

GDP per capita and the Gini coefficient were entered into the multiple imputation model and included as a covariate in the adjusted linear models used for mediation analysis as a sensitivity analysis. Neither GDP nor Gini coefficient had a strong association with TB incidence adjusting for the four TB risk factors. However, adjustment for the Gini coefficient resulted in a loss of conventional statistical significance for the health behaviour pathway, and a gain of conventional statistical significance for the social insurance pathway though the direction and magnitude of effect are still consistent with the prior results without income adjustment. This loss of conventional statistical significance is expected given the higher prevalence of both smoking and diabetes in high and middle-income countries that is likely captured in part by GDP. This result is not unexpected as GDP and Gini coefficient are likely to be distal (upstream) drivers of the four more proximal risk factors and thus some of the association of TB with smoking and diabetes may be accounted for in adjustment for GDP and Gini. As the direction and magnitude of effect still hold, we conclude that even after adjustment for income effects, health behaviour is likely to remain a mediator of the relationship between the SDG subtargets and TB. The results may indicate that the health behaviour pathway is responsible for less of the total effect of the SDG subtargets and TB than malnutrition, HIV, or housing quality.

| Association with TB Incidence<br>(adjusting additionally for GDP per capita and Gini coefficient) | |
|---|---|
| Indicator | Adjusted Estimate (more and equally proximal; blue & orange layer) |
| SDG 1.1: Extreme Poverty | 0.87 (-0.89, 2.64) |
| SDG 1.2: MPI Prevalence | -0.35 (-2.13, 1.44) |
| SDG 1.3: Labour Protection | -1.09 (-4.58, 2.41) |
| SDG 1.3: Social Assistance | -0.09 (-0.87, 0.69) |
| SDG 1.3: Social Insurance | **-1.52 (-2.84, -0.20)** |
| | |
| Income Covariate | Adjusted Estimate |
| GDP per capita (in 1000s) | -1.36 (-2.60, 0.12) |
| Gini Coefficient | 0.29 (-2.37, 2.94) |

| Measure | Adjusted Estimate |
|---|---|
| Malnourishment | **2.75 (0.64, 4.85)** |
| Health Behaviour | 2.49 (-0.48, 5.45) |
| Housing Quality | **1.54 (0.55, 2.54)** |
| HIV Prevalence | **20.44 (13.25, 27.63)** |

Bold indicates significant difference of coefficient from 0 with $p < .05$

**Table C**. The association between each indicator or measure and TB incidence per 100,000 per year. The adjusted estimate represents the association adjusted for all variables more or equally proximal to TB incidence as well as GDP per capita and Gini coefficient for a pooling of 50 datasets from a multiple imputation procedure.

Appendix 5 – Sensitivity Analyses to Modelling Assumptions for selected pathways

Complete esults from the model at 50% coverage of social protection (an increase of ~30% from baseline) and 90% elimination of extreme poverty (an increase of ~10% from baseline) are presented in Table D. There remains a substantial decrease in TB incidence predicted even from partial attainment of one or both of the SDG subtargets.

Results from producing projections for 2035 incidence using only countries with complete information for social insurance and extreme poverty (n = 94) are presented in Table E. The estimates are broadly similar to the estimates produced with multiple imputation, but with wider credible intervals – this is expected given the reduced power from decreased sample size, but also because of the exponential component of the model. The upper bound of the credible interval, if close to baseline, will decrease much less rapidly than its lower bound counterpart.

**Table D. Projected impact on 2035 TB incidence by mid-level achievement of selected subtargets**

| Pathway | β (95% Confidence Intervals) | Estimated 2035 Incidence (/100,000) | 95% Credible Interval (/100,000) | Reduction from Baseline* (95% CIs) | Expected Annual Proportional Decline |
|---|---|---|---|---|---|
| *Pathway A*: Social Protection for All (50% coverage) | -2·13 (-3·36, -0·90) | 69·6 cases | (50·9, 94·8) cases | 41·6% (20·5%, 57·3%) | 3·3% |
| *Pathway C:* Extreme Poverty Elimination (90% Eliminated) | -2·37 (-3·64, -1·10) | 97·6 cases | (89·0, 110·1) cases | 18·1% (7·6%, 25·3%) | 1·3% |
| *Pathway A & Pathway C:* Social Protection and Poverty Elimination | --- | 56·7 cases | (36·9, 86·2) cases | 52·5% (27·7%, 69·6%) | 4·6% |

**Table D**. The projected TB incidence per 100,000 per year associated with the achievement of selected SDG 1 subtargets independent from the other targets – 50% social protection coverage and 90% extreme poverty elimination.

**Table E. Projected impact on 2035 TB incidence: Complete cases**

| Pathway | β (95% Confidence Intervals) | Estimated 2035 Incidence (/100,000) | 95% Credible Interval (/100,000) | Reduction from Baseline* (95% CIs) | Expected Annual Proportional Decline |
|---|---|---|---|---|---|
| *Pathway A*: Social Protection for All | -2·66 (-4·36, -0·95) | 76·7 cases | (57·8, 101·6) cases | 35·6% (14·8%, 51·6%) | 2·7% |

| | | | | | |
|---|---|---|---|---|---|
| (100% coverage) | | | | | |
| *Pathway C:* Extreme Poverty Elimination (100% Eliminated) | -2·01 (-3·96, -0·08) | 30·9 cases | (7·8, 112·5) cases | 74·0% (5·6%, 93·5%) | 8·2% |
| *Pathway A & Pathway C:* Social Protection and Poverty Elimination | --- | 19·3 cases | (3·4, 95·8) cases | 83·8% (19·7%, 97·2%) | 10·9% |

**Table E**. The projected TB incidence per 100,000 per year associated with the achievement of selected SDG 1 subtargets – 100% social protection coverage and 100% extreme poverty elimination. These numbers are from a data subset with complete information on social insurance and extreme poverty.