**CODEBOOK AND DATA DIARY**

**Diagnostic accuracy study:** Evaluation of Xpert® MTB/RIF assay in induced sputum and gastric lavage/washing samples from young children with suspected tuberculosis from the MVA85A TB vaccine trial

In this study, we compare accuracy of theXpert® MTB/RIF assay, an automated nucleic acid amplification test, between induced sputum and gastric washing samples from young children in a South African TB endemic setting. The following provides a summary of our variables/codebook.

* Result of sputum smear microscopy test from first gastric washing sample (0=negative result, 1=positive result, .m=missing result, variable name **AAFB\_GW\_1**)
* Result of sputum smear microscopy test from second gastric washing sample (0=negative result, 1=positive result, .m=missing result, variable name **AAFB\_GW\_2**)
* Result of sputum smear microscopy test from first induced sputum sample (0=negative result, 1=positive result, .m=missing result, variable name **AAFB\_IS\_1**)
* Result of sputum smear microscopy test from second induced sputum sample (0=negative result, 1=positive result, .m=missing result, variable name **AAFB\_IS\_2**)
* Date of admission (variable name **Admission\_Date**)
* Admission episode (1=first admission, 2=second admission, 3=third admission, 4=fourth admission, variable name **Admission\_Event**)
* Date of birth (variable name **Date\_of\_Birth**)
* Result of mycobacterial culture for *Mycobacterium tuberculosis* (MTB) from first gastric washing sample (0=Negative, 1=positive, .m=missing, .c=contaminated, .n=Mycobacteria other than MTB (MOTT), variable name **MTB\_Culture\_GW\_1**)
* Result of mycobacterial culture for MTB from second gastric washing sample (0=Negative, 1=positive, .m=missing, .c=contaminated, .n=Mycobacteria other than MOTT, variable name **MTB\_Culture\_GW\_2**)
* Result of mycobacterial culture for MTB from first induced sputum sample (0=Negative, 1=positive, .m=missing, .c=contaminated, .n=Mycobacteria other than MOTT, variable name **MTB\_Culture\_IS\_1**)
* Result of mycobacterial culture for MTB from second induced sputum sample (0=Negative, 1=positive, .m=missing, .c=contaminated, .n=Mycobacteria other than MOTT, variable name **MTB\_Culture\_IS\_2**)
* Date of randomization into vaccine or placebo arm (variable name **Randomisation\_Date**)
* Sex of child (Female and Male, variable name **Sex**)
* Anonymized study identification number (variable name **Study\_Identification\_Number**)
* Arm (Placebo and Intervention arm, variable name **Arm**)
* Result of Xpert MTB/RIF assay, a nucleic acid test for *Mycobacterium tuberculosis*, from first gastric washing sample (0=negative, 1=positive, .i=invalid result as provided by GeneXpert device, .m=missing result (mainly due to stock-out of mitogen reagent during trial conduct, variable name **Xpert\_GW\_1**)
* Result of Xpert MTB/RIF assay from second gastric washing sample (0=negative, 1=positive, .i=invalid result as provided by GeneXpert device, .m=missing result (mainly due to stock-out of mitogen reagent during trial conduct, variable name **Xpert\_GW\_2**)
* Result of Xpert MTB/RIF assay from first induced sputum sample (0=negative, 1=positive, .i=invalid result as provided by GeneXpert device, .m=missing result (mainly due to stock-out of mitogen reagent during trial conduct, variable name **Xpert\_IS\_1**)
* Result of Xpert MTB/RIF assay from second induced sputum sample (0=negative, 1=positive, .i=invalid result as provided by GeneXpert device, .m=missing result (mainly due to stock-out of mitogen reagent during trial conduct, variable name **Xpert\_IS\_2**)