# Extended data: Statistical investigation of the UK Biobank biochemistry assays quality procedures - Inclusion criteria

Naomi E. Allen<sup>1,2\*</sup>, Matthew Arnold<sup>1\*</sup>, Sarah Parish<sup>3\*</sup>, Michael Hill<sup>3</sup>, Simon Sheard<sup>2</sup>, Howard Callen<sup>1,2</sup>, Daniel Fry<sup>2</sup>, Stewart Moffat<sup>3</sup>, Mark Gordon<sup>2</sup>, Samantha Welsh<sup>2</sup>, Paul Elliott<sup>4</sup>, Rory Collins<sup>1,2</sup>

<sup>&</sup>lt;sup>1</sup> Clinical Trial Service Unit & Epidemiological Studies Unit, Nuffield Department of Population Health, University of Oxford, Big Data Institute, Old Road Campus, Oxford, OX3 7LF.

<sup>&</sup>lt;sup>2</sup> UK Biobank, 1-2 Spectrum Way, Cheadle, Stockport, SK3 0LA.

<sup>&</sup>lt;sup>3</sup> MRC Population Health Research Unit, Nuffield Department of Population Health, University of Oxford, Big Data Institute, Old Road Campus, Oxford, OX3 7LF.

<sup>&</sup>lt;sup>4</sup> MRC Centre for Environment and Health, Department of Epidemiology and Biostatistics, School of Public Health, St Mary's Campus, Imperial College London, UK

<sup>\*</sup>joint first authors

### 1. Inclusion criteria for analyses

## All assays

- 1. Within reportable range at initial measurement
- 2. No error reported

## Assays on serum samples

- 3. Not in dip period
- 4. Not in "aliquot misclassification" period
- 5. Aliquot number 1-3, manual or missing

Inclusion rules for any dilution factor estimation (for other serum results the original result is retained):

- 6. In a restricted list of assays (see Table S1)
- 7. Not more than 15 SDs below mean in manual aliquots

Inclusion rules machine-tip dilution factor estimation (for other serum results the original result is retained):

- 7. Aliquot number 1-3, manual or missing
- 8. Has machine-tip information or was a manual sample
- 9. More than 1 result from machine-tip (one of the machines was recorded as being used only once for the resurvey data, so correction was not possible)
- 10. Not from visits in 2006, a year before the main baseline recruitment, because there were only 4 such samples with data (which is too sparse for this analysis)

Inclusion in sample reduction in concentration estimation:

11. With machine-tip dilution factor estimate but from a restricted list of assays (see Table S2)

#### Inclusion in date correction:

- 12. Results not excluded by 1-5 or 7 above
- 13. In restricted list of assays listed in point 6 plus HbA1c (for other results the original result is retained)

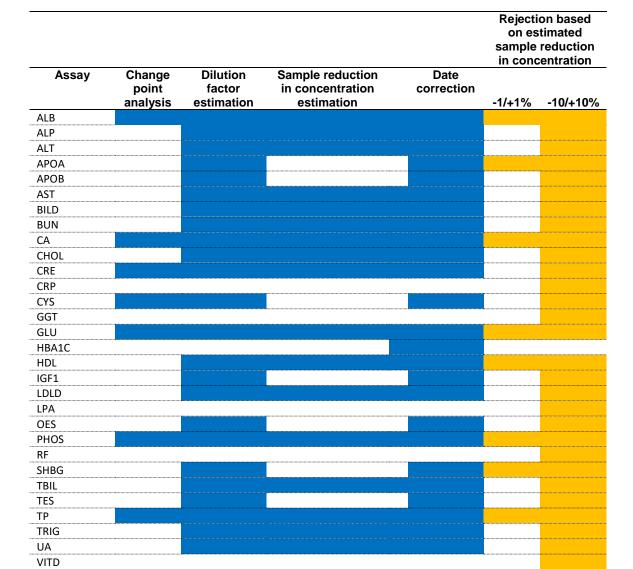


Table S1. Assays included in each analysis stage