# ****S1 Text. Disease standardisation****

Electronic health records are standardised to ICD10. Mandarin-speaking clinicians assign ICD10 codes to any free format disease description(s) using a recently developed bespoke standardisation tool. The ICD10 diagnoses are then verified by another clinician. A set of diagnoses are produced for each participant, showing which diseases they have developed and when. These data are then imported and linked into our research database in Oxford.