**Supplementary Material 2. Long-term conditions in USoc**

We explored the survey measure of long-term conditions from USoc and compared the prevalence and positive predictive value of the responses compared with long-term condition prevalence recorded in secondary care electronic healthcare records. We used data from USoc linked to Scottish Morbidity Records (linkage data from England, Wales and NI were not available).[[1]](#footnote-1)

We included in this analysis all USoc participants for whom linked secondary healthcare data were available (n=4,307, n=2,376 with long-term conditions measured in 2009). Long-term conditions in USoc were recorded if they had ever appeared in survey responses between 2009 and 2018. In linked data we included all secondary care inpatient admissions until 2018. We mapped survey response options to the appropriate ICD9 and ICD10 codes.

The survey question from USoc was initially developed from an existing survey tool (National Health and Nutrition Examination Survey, NHANES) with further long-term conditions added in 2017 and in 2020 in the COVID-19 waves. Only long-term conditions from the earlier waves are included in this validation although all responses from the COVID-19 version of the question were used in this main analysis.

We initially described the completeness of the linkage by age (Supplementary Material 2, Table 1), compared the prevalence of long-term conditions in USoc with hospital records (Supplementary Material 2, Table 2), and calculated the positive predictive value (PPV) of both for the other (Supplementary Material 2, Figure 1). If a patient's hospital record has a record for a particular long-term condition, then the positive predictive value of the hospital record measure of the long-term condition is the probability that among people with this condition in their hospital record, the response to USoc includes the same long-term condition. If a patient's USoc survey response has a record for a particular long-term condition, then the positive predictive value of the USoc measure of the long-term condition is the probability among people with this long-term condition recorded in the survey that the hospital record includes the same long-term condition.

We found that linkage is more complete at older ages (Table 1). This reflects the fact that the probability of inpatient hospital admission increases with age.

**Supplementary Material 2, Table 1. Completeness of linkage by year of birth**

|  |  |
| --- | --- |
| **Year of birth** | **Number (unweighted %) with a linked record** |
| 1910-1919 | 3 (100.0) |
| 1920-1929 | 38 (97.4) |
| 1930-1939 | 218 (97.7) |
| 1940-1949 | 413 (96.4) |
| 1950-1959 | 567 (89.1) |
| 1960-1969 | 646 (88.7) |
| 1970-1979 | 664 (87.7) |
| 1980-1989 | 557 (83.5) |
| 1990-1999 | 529 (75.2) |
| 2000-2009 | 277 (63.9) |
| 2010- | 395 (50.4) |

The differences between the two data sets reflect the differences in how the measures were captured, reflected in differences in prevalence shown in Table 2 and differences in PPV in Figure 1. USoc data provide a self-reported measure of long-term health conditions, whereas secondary care diagnoses are only recorded for people who have had an inpatient stay in hospital. For example, cancer is more likely to be recorded in hospital records than these survey data reflected in the higher hospital compared with USoc PPV and prevalence, while asthma, which is frequently managed in primary care and not associated with a hospital admission, has higher prevalence and PPV in USoc compared with hospital admission data.

**Supplementary Material 2, Table 2. Long-term condition prevalence in USoc and Secondary care records**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Long-term condition** | **ICD9** | **ICD10** | **Scotland (ever)** | **IP admission (ever)** |
| Asthma | 493 | J45 | 360 (15.2) | 149 (5.9) |
| Arthritis | 711-716 | M00-M19 | 473 (18.3) | 195 (7.9) |
| Congestive heart failure | 428 | I50 | 28 (1.2) | 66 (2.7) |
| Coronary heart disease | 410-414 | I20-I25 | 68 (2.9) | 182 (7.4) |
| Angina | 411, 413 | I20 | 108 (4.4) | 98 (4.1) |
| Heart attack or myocardial infarction | 410 | I21 I22 I23 | 86 (3.5) | 66 (2.7) |
| Stroke | 430-438 | I60-I69 | 81 (3.5) | 73 (3.4) |
| Emphysema | 492 | J43 | 43 (1.7) | 10 (0.3) |
| Hyperthyroidism  | 242 | E05 | 22 (0.9) | 15 (0.6) |
| Hypothyroidism  | 244 | E00-E03 | 114 (4.4) | 50 (1.9) |
| Chronic bronchitis / COPD | 491, 496 | J44 | 65 (2.5) | 87 (3.3) |
| Any kind of liver condition | 570-573 | K70-K77 | 59 (2.4) | 47 (1.8) |
| Cancer or malignancy | 140-239 | C00-C97 | 118 (4.8) | 301 (12.4) |
| Diabetes | 249, 250 | E10-E14 | 137 (5.3) | 115 (4.4) |
| Epilepsy | 345 | G40 G41 | 50 (2.2) | 37 (1.4) |
| High blood pressure | 401-405 | I10-I15 | 572 (23.5) | 291 (12.1) |
| Clinical depression | 296, 311, 298 | F32 F33 F34 F38 F39 | 283 (10.8) | 52 (2.0) |

**Supplementary Material 2, Figure 1. PPV of USoc and secondary care recording of long-term health conditions**



Although with the caveat that the data only came from Scotland, rather than the whole of the UK, this analysis highlights that the long-term health conditions recorded in USoc are those which are self-defined by an individual, and managed in the community, and do not reflect the burden of long-term health conditions seen and managed in secondary care. This would suggest that the measurement of long-term conditions in USoc is appropriate as the patterns seen compared with EHR data are those that would be expected. For this evaluation, which explores multimorbidity in a primary care intervention, this analysis highlights the appropriateness of the long-term condition measure in USoc, as primary care in where most long-term conditions in the community are managed.

1. in house analysis of Understanding Society linked Scottish Morbidity Records, ISER, University of Essex (2019) [↑](#footnote-ref-1)