**RECOVERY generalizability and representativeness project**

**Hospital Episode Statistics derivation methodology and codelists**

Note – the HES APC data dictionary is available at <https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/hospital-episode-statistics/hospital-episode-statistics-data-dictionary#download-hes-data-dictionaries>

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Characteristic | HES field | ICD-10/ OPCS-4 codes  (NB all codes are ICD-10 unless otherwise specified) | Derivation rule | Notes | References |
| Age | admiage | N/A | Value recorded in index spell |  | HES APC data dictionary |
| Gender | sex | N/A | Value recorded in index spell |  | HES APC data dictionary |
| Ethnicity | ethnos | N/A | Value recorded in index spell |  | HES APC data dictionary |
| COVID-19 coding | diag\_01 to diag\_20 (any position) | U071, U072 | If present in index spell (NB: the presence of these codes is used to define the national HES cohort; for the RECOVERY HES cohort we have used presence of a COVID-19 code, or else date of randomisation) | WHO documentation | <https://www.who.int/standards/classifications/classification-of-diseases/emergency-use-icd-codes-for-covid-19-disease-outbreak> |
| Index of multiple deprivation (score) | lsoa11 | N/A | Value recorded in index spell | Derived from the lsoa11 (lower-super-output-area) field rather than IMD as IMD provided in HES at the moment refers to the 2010 indices (most recent are 2019) | https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019 |
| Index of multiple deprivation (quintiles) | lsoa11 | N/A | Value recorded in index spell |
| Charlson score (aggregated) | diag\_01 to diag\_20  (any position) | (see individual parameters) | Any record in the 5 years before index date (epistart of the index spell) | Charlson score calculated from ICD-10 codes using published and validated methodology and codelists | 10.1097/01.mlr.0000182534.19832.83 |
| Myocardial infarction (Charlson score) | I21, I22, I25.2 |
| Congestive heart failure (Charlson score) | I09.0, I11.0, I13.0, I13.2, I25.5, I42.0, I42.5, I42.6, I42.7, I42.8,  I42.9, I43, I50, P29.0 |
| Peripheral vascular disease (Charlson score) | I70, I71, I73.1, I73.8, I73.9, I77.1, I79.0, I79.2, K55.1, K55.8,  K55.9, Z95.8, Z95.9 |
| Cerebrovascular disease (Charlson score) | G45 |
| Chronic pulmonary disease (Charlson score) | I27.8, I27.9, J40, J41, J42, J43,  J44, J45, J46, J47, J60, J61, J62,  J63, J64, J65, J66, J67, J68.4, J70.1, J70.3 |
| Rheumatic disease (Charlson score) | M05, M06, M31.5, M32, M33, M34, M35.1, M35.3, M36.0 |
| Dementia (Charlson score) | F00, F01, F02, F03, F05.1, G30,  G31.1 |
| Peptic ulcer disease (Charlson score) | K25, K26, K27, K28 |
| Liver disease, mild (Charlson score) | B18, K70.0, K70.1, K70.2, K70.3, K70.9, K71.3, K71.4, K71.5, K71.7, K73, K74, K76.0, K76.2,  K76.3, K76.4, K76.8, K76.9,  Z94.4 |
| Liver disease, moderate-severe (Charlson score) | I85.0, I85.9, I86.4, I98.2,  K70.4, K71.1, K72.9, K76.5,  K76.6, K76.7 |
| Diabetes mellitus, with chronic complications (Charlson score) | E10.2, E10.3, E10.4, E10.5, E10.7, E11.2, E11.3, E11.4, E11.5, E11.7, E12.2, E12.3,  E12.4, E12.5, E12.7, E13.2,  E13.3, E13.4, E13.5, E13.7,  E14.2, E14.3, E14.4, E14.5,  E14.7 |
| Diabetes mellitus, without chronic complications  (Charlson score) | E10.0, E10.1, E10.6, E10.8,  E10.9, E11.0,E11.1, E11.6,  E11.8, E11.9, E12.0, E12.1,  E12.6, E12.8, E12.9, E13.0,  E13.1, E13.6, E13.8, E13.9,  E14.0, E14.1, E14.6, E14.8,  E14.9 |
| Renal disease  (Charlson score) | I21.0, I13.1, N03.2, N03.3, N03.4, N03.5, N03.6, N03.7,  N05.2, N05.3, N05.4, N05.5,  N05.6, N05.7, N18, N19, N25.0, Z49.0, Z49.1, Z49.2, Z94.0, Z99.2 |
| Any solid tumour, excluding malignant neoplasm of skin  (Charlson score) | C00-C26, C30-C34, C37-C41,  C43, C45-C58, C60—C76, C88.2, C88.7, C88.9, C90.0, C90.2, C90.3, C96-C97 |
| Metastatic solid tumour (Charlson score) | C77-C80 |
| Hemiplegia or paraplegia (Charlson score) | G04.1, G11.4, G80.1, G80.2, G81, G82, G83.0, G83.1, G83.2, G83.3, G83.4, G83.9 |
| AIDS /HIV  (Charlson score) | B20, B21, B22, B24 |
| Lymphoma  (Charlson score) | B21.1, C81-C86, C88.0, C88.3, C88.4 | Codelists built based on clinical review of the ICD-10 terminology by two clinicians (no publicly available codelists for these criteria) | Bespoke |
| Leukemia  (Charlson score) | C90.1, C91-C95, M36.1, D47.1, N16.1, D47.5 |
| Heart disease (any) | diag\_01 to diag\_20  (any position) | I08, I09, I11, I13, I20, I21, I22, I23, I24, I25, I27.1, I27.8, I27.9,  I30, I31, I32, I33, I34, I35, I36, I37, I38, I39, I40, I41, I42, I43, I44, I45, I46, I47, I48, I49, I50, I51, I52 | Any record in the 5 years before index date (epistart of the index spell) | RECOVERY trial cause-of-death derivation (“cardiac” category) | Reference? |
| Diabetes (any) | E10-E14 | Merged both Charlson diabetes codelists (codes above), manual review of the ICD-10 terminology, and literature review | https://bmjopen.bmj.com/content/6/8/e009952.long |
| Chronic liver disease | B18, K70.0, K70.1, K70.2, K70.3, K70.9, K71.3, K71.4, K71.5, K71.7, K73, K74, K76.0, K76.2,  K76.3, K76.4, K76.8, K76.9,  Z94.4, I85.0, I85.9, I86.4, I98.2,  K70.4, K71.1, K72.9, K76.5,  K76.6, K76.7 | Merged mild and moderate-severe liver disease lists | N/A |
| Severe mental illness | F20, F21, F22, F23, F24, F25, F28, F29, F30, F31, F32.2, F32.3  F32.8, F32.9, F33.2, F33.3  F33.4, F33.8, F33.9 | Publicly-available codelist used for the QCOVID NHS risk calculator | https://www.datadictionary.nhs.uk/Covid19PRA/Severe\_Mental\_Illness.html |
| Alcohol-attributable diseases | F10.3-F10.9, F10.0, F10.1,  F10.2, G62.1, G31.2, G72.1,  I42.6, K29.2, K70.0-K70.4,  K70.9, K85.2, K86.0, Q86.0,  P04.3 | Publicly-available codelist published by the US CDC | https://www.cdc.gov/alcohol/ardi/alcohol-related-icd-codes.html |
| Obesity | E66 | Any record in the 5 years before index date (epistart of the index spell) | ICD-10 terminology review, literature review | https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6748036/ |
| Immunosuppression (cancer with active chemotherapy) | diag\_01 to diag\_20 for ICD-10 codes  (any position)  opertn\_01 to opertn \_21 for OPCS-4 codes (any position) | **Cancer (ICD-10):**  Check full list in codelist repository file (spreadsheet)  **Chemotherapy (ICD-10):**  Y43.1, Y43.3, Z51.1  **Chemotherapy (OPCS-4):**  Check full list in codelist repository file (spreadsheet) | Cancer diagnosis recorded in the 5 years prior to index date AND chemotherapy recorded in 6 months prior to index date | Definition according to the UK Green Book criteria (and the Chief Medical Officer’s recommendations for the shielded-patient list) | <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1057798/Greenbook-chapter-14a-28Feb22.pdf>  https://digital.nhs.uk/coronavirus/shielded-patient-list/methodology/rule-logic |
| Immunosuppression (haematological malignancy or bone marrow transplant) | **Haematological malignancy (ICD-10):**  Check full list in codelist repository file (spreadsheet)  **Haematological transplant (ICD-10):**  T86.0  **Haematological transplant (OPCS-4):**  W34, X33.6 | Haematological malignancy or haematological transplant recorded in the 2 years prior to index date |
| Immunosuppression (solid organ transplant) | **Solid organ transplant (ICD-10):**  N16.5, T86.1, T86.2, T86.3, T86.4, T86.8, T86.9, Y83.0, Z94  **Solid organ transplant (OPCS-4):**  Check full list in codelist repository file (spreadsheet) | Solid organ transplant recorded in the 5 years prior to index date |
| Immunosuppression (hyposplenism) | **Hyposplenism (ICD-1):**  D73.0, Q89.0, D57.0, D57.1  **Splenectomy (OPCS-4):**  J69.2 | Hyposplenism or splenectomy recorded in the 5 years prior to index date |
| Immunosuppression (other long-term) | D829, D83, D830, D831, D832,  D838, D839, D84, D840, D841, D848, D849, D71 | Any record within 5 years prior to index date | Manual terminology review and code selection | N/A |
| Immunosuppression (other short-term) | **ICD-10:** D70, Y43.3  **OPCS:** X92.1, X95.1 | Any record within 6 months prior to index date |
| Renal replacement therapy  (peritoneal dialysis) | **Peritoneal dialysis (ICD-10):** Z49.2  **Acute kidney injury (ICD-10):** N17  **Peritoneal dialysis (OPCS-4):** X41.1, X40.2, X40.5, X40.6 | Occurrence of any admission with a peritoneal dialysis code (without diagnosis of acute kidney injury) in the 5 years before index date (epistart of the index spell) | reference | reference |
| Renal replacement therapy  (maintenance haemodialysis) | **Dialysis (ICD-10):**  E85.3, Y60.2, Y61.2, Y62.2, Y84.1, Z99.2, T82.4, Z49.1  **Dialysis (OPCS-4):**  X40.1, X40.3, X40.4 | Occurrence of a dialysis code in the 5 years before index date (epistart of the index spell) of the index spell in a patient who has had: |
| **ESKD (ICD-10):**  N18.0, N18.5, Q60.1 | (a) a diagnostic code for ESKD any prior time; OR |
| **Fistula/graft (OPCS-4):**  L74.1, L74.2, L74.6, L74.8, L74.9 | (b) the insertion of an arteriovenous fistula or graft any prior time. |
| Renal replacement therapy  (probable maintenance haemodialysis) | **Dialysis (ICD-10):**  E85.3, Y60.2, Y61.2, Y62.2, Y84.1, Z99.2, T82.4, Z49.1  **Dialysis (OPCS-4):**  X40.1, X40.3, X40.4 | Occurrence of at least two episodes containing a dialysis code, with at least 90 days between the start of the first recorded dialysis, and the start of any subsequent dialysis (without diagnosis of acute kidney injury) in the 5 years before index date (epistart of the index spell) |
| Hospital frailty score | diag\_01 to diag\_20 for ICD-10 codes  (any position); | Codes and scoring used as per reference publication | Last 2 years including index admission | https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(18)30668-8/fulltext | https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(18)30668-8/fulltext |