Documentation Control Sheet

Over time, it may be necessary to issue amendments or clarifications to parts of this document. This form must be updated whenever changes are made.

<table>
<thead>
<tr>
<th>Version</th>
<th>Affected Areas</th>
<th>Prepared By</th>
<th>Reviewed By</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>Initial Draft</td>
<td>Tarita Murray-Thomas</td>
<td>Wilhelmine Meeraus, Arlene Gallagher</td>
</tr>
<tr>
<td>1.1</td>
<td>Modified</td>
<td>Tarita Murray-Thomas</td>
<td>Shivani Padmanabhan</td>
</tr>
<tr>
<td>1.2</td>
<td>Modified</td>
<td>Jenny Campbell</td>
<td>Rebecca Ghosh</td>
</tr>
<tr>
<td>1.3</td>
<td>Modified</td>
<td>Jenny Campbell</td>
<td>Elizabeth Crellin</td>
</tr>
<tr>
<td>1.4</td>
<td>Modified</td>
<td>Elizabeth Crellin</td>
<td>Jenny Campbell</td>
</tr>
<tr>
<td>1.5</td>
<td>Modified</td>
<td>Elizabeth Crellin</td>
<td>Jenny Campbell</td>
</tr>
<tr>
<td>1.6</td>
<td>Modified</td>
<td>Elizabeth Crellin</td>
<td>Jenny Campbell</td>
</tr>
<tr>
<td>1.7</td>
<td>Modified</td>
<td>Susan Hodgson</td>
<td>Jenny Campbell</td>
</tr>
<tr>
<td>1.8</td>
<td>Modified</td>
<td>Tarita Murray-Thomas</td>
<td>Susan Hodgson</td>
</tr>
</tbody>
</table>

Summary of Changes

Version 1.1
- Updated document version number, date and HES set
- Added the HES AE coverage dates for this release
- Added explanation of changed definition of the derived ethnicity variable
- Added changes introduced in set 13
- Updated references to reflect change of name from HSCIC to NHS Digital

Version 1.2
- Updated document version number, date and HES set for the release of set 14.
- Updated match rank table for set 14

Version 1.3
- Updated document version number, date and HES set for the release of set 15.
- Updated match rank table for set 15
- waitdays variable is no longer available and has been removed from the documentation.

Version 1.4
- Updated document version number, date and HES set for the release of set 16.
- Updated match rank table for set 16
- Updated to include CPRD Aurum
- diagscheme variable has been added

Version 1.5
- Updated document version number, date and HES set for the release of set 17.
- Updated match rank table for set 17
- Added table showing percentage capture of national A&E attendances by year
- Corrected data type for four variables in Attendance file

Version 1.6
- Updated document version number, date and HES set for the release of set 18
- Updated match rank table for set 18

Version 1.7
- Updated document version number, date and HES set for the release of set 19
- Updated match rank table for set 19
- Added sections on impact of COVID-19 pandemic and the introduction of the Emergency Care Dataset on HES A&E reporting.

Version 1.8
- Updated document version number, date and HES set for the release of set 21, added DOIs
HES Accident & Emergency (A&E) data linked to CPRD primary care data

This document provides an overview of HES Accident and Emergency (HES A&E) data, and the available subset that is linked to CPRD GOLD and CPRD Aurum.

Impact of the COVID-19 pandemic

What are the HES Accident and Emergency data?
HES A&E data consist of individual records of patient care administered in the accident and emergency setting in England. These data are a subset of national A&E data collected by NHS England to monitor the national standard that 95% of patients attending A&E for care should wait no longer than 4 hours from arrival to admission, transfer or discharge. A&E data are submitted by providers of all types of A&E services in England - Type 1, Type 2, Type 3, Type 4 departments and urgent care centres that average more than 50 attendances per week (Annex A).

The collection of HES A&E was first started in April 2007 on an experimental basis and continued until 2012/2013 when the experimental label was lifted. Table 1 shows the percentages of national A&E attendances that were captured in HES A&E data by year, excluding planned follow-up attendances:

Table 1: Capture of HES A&E attendances

<table>
<thead>
<tr>
<th>Year</th>
<th>Experimental status</th>
<th>Percentage attendances captured</th>
</tr>
</thead>
<tbody>
<tr>
<td>2007/08</td>
<td>Experimental</td>
<td>62%</td>
</tr>
<tr>
<td>2008/09</td>
<td>Experimental</td>
<td>68%</td>
</tr>
<tr>
<td>2009/10</td>
<td>Experimental</td>
<td>74%</td>
</tr>
<tr>
<td>2010/11</td>
<td>Experimental</td>
<td>74%</td>
</tr>
<tr>
<td>2011/12</td>
<td>Experimental</td>
<td>80%</td>
</tr>
<tr>
<td>2012/13</td>
<td></td>
<td>83%</td>
</tr>
<tr>
<td>2013/14</td>
<td></td>
<td>83%</td>
</tr>
<tr>
<td>2014/15</td>
<td></td>
<td>86%</td>
</tr>
</tbody>
</table>

Records in the HES A&E database are called ‘attendances’ and each A&E attendance relates to a single visit by an individual to A&E. Where follow up care is required and provided by the A&E department, a second planned attendance is recorded. A&E data collected includes details about patients’ attendance, outcome of the visit to A&E, waiting times, referral source, A&E diagnosis (not ICD-10 coded), A&E treatment (drug prescribing not recorded), A&E investigations (not OPCS coded) and Health Resource Group.

The impact of the introduction of the Emergency Care Data Set on HES A&E reporting

Overall coverage in HES has increased from 2018-19, however the data completeness for a number of key fields has reduced since the phased introduction of the new Emergency Care Data Set (ECDS) commenced in October 2017. As such, Annual data for certain fields and reported activity is now no longer directly comparable.


The methodological change notice paper published by NHS Digital is available at: https://digital.nhs.uk/binaries/content/assets/website-assets/publications/publications-admin-pages/methodological-changes/methchange20171212_hes.pdf

Accessing HES Accident and Emergency data linked to CPRD GOLD and CPRD Aurum

HES A&E data can only be accessed as part of a data extract linked to CPRD primary care data (CPRD GOLD or CPRD Aurum). Access is provided by the CPRD for a fee subject to protocol approval.

Not all patients in CPRD GOLD or CPRD Aurum are eligible to be linked to HES, for example, due to the region in which they reside (outside England), or the lack of a valid NHS identifier. Source files (linkage_eligibility.txt) are provided to allow researchers to identify the subset of patients who are eligible to have linked HES data.

Linkage coverage period


DOIs

Please cite in any publications using these data:

CPRD GOLD HES A&E August 2021 - https://doi.org/10.48329/xtwg-gp32

CPRD Aurum HES A&E August 2021 - https://doi.org/10.48329/2qy5-xx60

Linkage algorithm and the match_rank variable

Linkage between HES A&E and CPRD primary care data uses an eight-step deterministic linkage algorithm based on four identifiers, shown in Table 2. The linkage is undertaken by NHS Digital, acting as a trusted-third-party, on behalf of CPRD. No personal identifiers are held by CPRD, or included in the CPRD GOLD, CPRD Aurum, or linked HES A&E data.
Table 2: NHS Digital 8 step linkage algorithm

<table>
<thead>
<tr>
<th>Step</th>
<th>Match</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Exact NHS number, sex, date of birth (DOB), postcode</td>
</tr>
<tr>
<td>2</td>
<td>Exact NHS number, sex, DOB</td>
</tr>
<tr>
<td>3</td>
<td>Exact NHS number, sex, postcode, partial DOB</td>
</tr>
<tr>
<td>4</td>
<td>Exact NHS number, sex, partial DOB</td>
</tr>
<tr>
<td>5</td>
<td>Exact NHS number, postcode</td>
</tr>
<tr>
<td>6</td>
<td>Exact sex, DOB, and postcode (where NHS number does not contradict the match, the DOB is not 1st of January &amp; the postcode not on the communal establishment list)</td>
</tr>
<tr>
<td>7</td>
<td>Exact sex, DOB, and postcode (where the NHS number does not contradict the match and the DOB is not 1st of January)</td>
</tr>
<tr>
<td>8</td>
<td>Exact NHS number</td>
</tr>
</tbody>
</table>

The matching steps are applied sequentially. If a CPRD GOLD or CPRD Aurum patient record is matched in one step, it is no longer available for matching in subsequent steps. Matching results are summarised in Table 3A and 3B.

Table 3A: Number and proportion of **CPRD GOLD** patients matched to a HES patient* at each step of the linkage algorithm in set 21.

<table>
<thead>
<tr>
<th>Linkage step (match rank)</th>
<th>Frequency</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5,679,805</td>
<td>68.6%</td>
</tr>
<tr>
<td>2</td>
<td>2,310,259</td>
<td>27.9%</td>
</tr>
<tr>
<td>3</td>
<td>13,415</td>
<td>0.2%</td>
</tr>
<tr>
<td>4</td>
<td>18,075</td>
<td>0.2%</td>
</tr>
<tr>
<td>5</td>
<td>3,447</td>
<td>0.0%</td>
</tr>
<tr>
<td>6</td>
<td>234,817</td>
<td>2.8%</td>
</tr>
<tr>
<td>7</td>
<td>14,425</td>
<td>0.2%</td>
</tr>
<tr>
<td>8</td>
<td>6,494</td>
<td>0.1%</td>
</tr>
</tbody>
</table>

*includes patients in all HES datasets (Admitted patient care, Outpatient, and A&E)

Table 3B: Number and proportion of **CPRD Aurum** patients matched to a HES patient* at each step of the linkage algorithm in set 21.

<table>
<thead>
<tr>
<th>Linkage step (match rank)</th>
<th>Frequency</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>23,508,210</td>
<td>65.9%</td>
</tr>
<tr>
<td>2</td>
<td>10,772,078</td>
<td>30.2%</td>
</tr>
<tr>
<td>3</td>
<td>49,245</td>
<td>0.1%</td>
</tr>
<tr>
<td>4</td>
<td>75,154</td>
<td>0.2%</td>
</tr>
<tr>
<td>5</td>
<td>13,358</td>
<td>0.0%</td>
</tr>
<tr>
<td>6</td>
<td>1,129,843</td>
<td>3.2%</td>
</tr>
<tr>
<td>7</td>
<td>65,212</td>
<td>0.2%</td>
</tr>
<tr>
<td>8</td>
<td>32,727</td>
<td>0.1%</td>
</tr>
</tbody>
</table>

*includes patients in all HES datasets (Admitted patient care, Outpatient, and A&E)

CPRD provides users with a match_rank variable which corresponds to the step at which the match was established. In general, a lower value for the match_rank is considered stronger evidence for a positive match. Note that only patients with a match_rank of 5 or less are considered definitive matches and are
included in the linked HES A&E dataset. Patients matched on steps 6-8 have been retained in separate files. We envisage that the retained records will primarily be of interest to methodological researchers. If you are interested in these data, please speak to a member of the CPRD Observational Research team prior to submission of your protocol.

A linkage coverage file (linkage_coverage.txt) provides the start and end dates of HES encounter time.

A minority of patients are linked to multiple HESIDs. These patients are removed from the HES A&E dataset. However, the data have been retained and are available on request. If you are interested in these data, please speak to a member of the CPRD Observational Research team prior to submission of your protocol.

As far as possible, the linked HES A&E data is supplied “as is”, without any modification or cleaning during processing by CPRD. Where CPRD has modified the HES data, these are detailed below.

Data structure and formatting
HES A&E data provided by the CPRD represents only a subset of the variables that are collected in the National HES A&E dataset provided by NHS Digital. Fields such as organisation fields which may lead to the potential re-identification of patients or practices are not collected by the CPRD and/or not supplied to users.

The data are arranged into files relating to A&E attendance, diagnosis, treatment, and investigations undertaken in the A&E setting. Each record represents a single A&E attendance at a single provider. The HES ‘aekey’, the record identifier created by HES, is unique in combination with the CPRD patient identifier (patid). The patient identifier (patid) may be used to link together A&E attendance records for a single patient with CPRD HES admitted patient care and/or outpatient records.

For each patient cohort, HES A&E data will be provided as separate text tab delimited files. Files can be imported into statistical software such as Stata or SAS, or into data management packages such as Microsoft Access, for further data processing and analysis.

The format of the HES data has been modified for linked patients in the following ways:

- CPRD has introduced the HES patient identifier (gen_hesid) in A&E data. This is unique across all CPRD linked HES datasets including HES admitted patient care (APC), HES Outpatient (OP) and HES A&E data. An individual that has contributed data to more than one CPRD practice will have the same gen_hesid but this may change between linkage sets.
- The aekey variable has been altered so that it is unique (by patient identifier) across all A&E data.
- CPRD has provided a derived ethnicity variable (gen_ethnicity) in the HES A&E patient file which is the most commonly recorded ethnicity for each patient, among all HES data including HES APC, HES OP and HES A&E. The ethnicity recorded at A&E attendance (ethnos) remains unchanged.

Changes introduced in HES A&E sets

Set 12
Licensing obligations require that no attempts are made to re-identify patients in CPRD datasets. The aekey has been encoded by the CPRD to minimise the risk of breaching licensing conditions through
linkage of these data to other HES data sources containing patient identifiable information. What this means is that from set 12, the aekey variable is different from that of previous sets and will differ in each future release of HES A&E linkage sets.

Set 13
The definition of the derived ethnicity variable (gen_ethnicity) in the patient file has been changed so that ethnicity is specified where at least one episode has a specific ethnicity recorded but the majority of values are “unknown”.

Set 16
The diagscheme variable has been added to the CPRD linked dataset.

Known issues
- Data coverage is incomplete in comparison to national A&E data attendances. Analyses of patterns of missing data may be required when using these data.
- Some variables have high levels of missing data e.g. duration to initial assessment (initdur)
- Provisional HES A&E data are monthly publications of HES data. These data may be incomplete or contain errors for which no adjustments have yet been made by HES. Counts produced from provisional data are likely to be lower than those generated for the same period in the final dataset. It is also probable that clinical data are not complete, which may affect the last two months of any given period. There may also be errors due to coding inconsistencies that have not yet been investigated and corrected. At the end of the fiscal year there is a “month 13” annual refresh which corrects known data quality issues prior to locking the annual published data.
- Diagnosis is recorded using a specific coding system for HES A&E data. However, there may be some ICD-10 and Read codes within the A&E data resulting from experimental recording in a small percentage of A&E departments. The diagscheme variable provides an indication of the coding scheme used.

Look-up files
Lookup files relating to the use of HES A&E data will not be provided by the CPRD. These can be obtained online from NHS Digital using this link: https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/hospital-episode-statistics/hospital-episode-statistics-data-dictionary
# Annex A: Type of A&E service providers

<table>
<thead>
<tr>
<th>A&amp;E Service type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type 1</td>
<td>A&amp;E department = A consultant led 24-hour service with full resuscitation facilities and designated accommodation for the reception of accident and emergency patients</td>
</tr>
<tr>
<td>Type 2</td>
<td>A&amp;E department = A consultant led single specialty accident and emergency service (e.g. ophthalmology, dental) with designated accommodation for the reception of patients</td>
</tr>
<tr>
<td>Type 3 / Type 4 / Urgent Care Centre</td>
<td>Other type of A&amp;E/minor injury units (MIUs)/Walk-in Centres (WiCs)/Urgent Care Centre primarily designed for the receiving of accident and emergency patients. A type 3 department may be doctor led or nurse led. It may be co-located with a major A&amp;E or sited in the community. A defining characteristic of a service qualifying as a type 3 department is that it treats at least minor injuries and illnesses (sprains for example) and can be routinely accessed without appointment.</td>
</tr>
</tbody>
</table>
## HES A&E: Data dictionary

### 1. Patient (hesae_patient.txt)

<table>
<thead>
<tr>
<th>Column name</th>
<th>Description</th>
<th>Type</th>
<th>Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>patid</td>
<td>Encrypted unique key given to a patient in CPRD GOLD or CPRD Aurum [primary key]</td>
<td>INTEGER</td>
<td>20</td>
</tr>
<tr>
<td>pracid</td>
<td>Encrypted unique key given to a practice in CPRD GOLD or CPRD Aurum</td>
<td>INTEGER</td>
<td>5</td>
</tr>
<tr>
<td>gen_hesid¹</td>
<td>A generated unique key assigned to a patient across all CPRD linked HES datasets within a linkage set. An individual that has contributed data to more than one CPRD practice has the same gen_hesid but this may change between linkage sets.</td>
<td>INTEGER</td>
<td>20</td>
</tr>
<tr>
<td>n_patid_hes¹</td>
<td>Number of individuals in CPRD GOLD or CPRD Aurum assigned the same gen_hesid (unique patient identifier generated in HES)</td>
<td>INTEGER</td>
<td>3</td>
</tr>
<tr>
<td>gen_ethnicity¹</td>
<td>Patient’s ethnicity derived from all HES data (including HES outpatient, HES admitted patient care and HES A&amp;E)</td>
<td>CHAR</td>
<td>10</td>
</tr>
<tr>
<td>match_rank²</td>
<td>Indicates the quality of matching between a record in HES and CPRD primary care data and gives the level of confidence that an HES record has been correctly matched to a patient in CPRD GOLD or CPRD Aurum.</td>
<td>INTEGER</td>
<td>1</td>
</tr>
</tbody>
</table>

¹ Variable generated by CPRD.
² An eight-step process is used to match patients in CPRD primary care data (CPRD GOLD or CPRD Aurum) and HES using some or all of the following: NHS number, date of birth, sex and postcode. Only data for patients matched using steps 1-5 has been provided.
## ATTENDANCE (hesae_attendance.txt)

<table>
<thead>
<tr>
<th>Column name</th>
<th>Description</th>
<th>Type</th>
<th>Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>patid</td>
<td>Encrypted unique key given to a patient in CPRD GOLD or CPRD Aurum [primary key, in combination with aekey]</td>
<td>INTEGER</td>
<td>20</td>
</tr>
<tr>
<td>aekey³</td>
<td>Record identifier (unique in combination with patid) [primary key, in combination with patid]</td>
<td>INTEGER</td>
<td>20</td>
</tr>
<tr>
<td>arrivaldate</td>
<td>The arrival date of a patient in the A&amp;E department</td>
<td>DATE</td>
<td>dd/mm/yyyy</td>
</tr>
<tr>
<td>aepatgroup</td>
<td>The reason for an A&amp;E episode</td>
<td>INTEGER</td>
<td>2</td>
</tr>
<tr>
<td>aearrivaldate</td>
<td>The arrival date of a patient in the A&amp;E department</td>
<td>DATE</td>
<td>dd/mm/yyyy</td>
</tr>
<tr>
<td>aeadpttype</td>
<td>A classification of A&amp;E department type according to the activity carried out</td>
<td>INTEGER</td>
<td>2</td>
</tr>
<tr>
<td>aerefsource</td>
<td>The source of referral for each A&amp;E episode</td>
<td>INTEGER</td>
<td>2</td>
</tr>
<tr>
<td>aequalctype</td>
<td>Classification of the place where the incident occurred that led to an A&amp;E episode</td>
<td>INTEGER</td>
<td>2</td>
</tr>
<tr>
<td>aeatenddisp</td>
<td>The way in which an A&amp;E attendance might end</td>
<td>INTEGER</td>
<td>2</td>
</tr>
<tr>
<td>initdur</td>
<td>The time (expressed as a whole number of minutes) between the patient’s arrival and their initial assessment</td>
<td>INTEGER</td>
<td>8</td>
</tr>
<tr>
<td>tretdur</td>
<td>The time (expressed as a whole number of minutes) between the patient’s arrival and the start of their treatment</td>
<td>INTEGER</td>
<td>8</td>
</tr>
<tr>
<td>concldur</td>
<td>The time (expressed as a whole number of minutes) between the patient’s arrival and conclusion of their attendance or treatment (whichever is later)</td>
<td>INTEGER</td>
<td>8</td>
</tr>
<tr>
<td>depdur</td>
<td>The time (expressed as a whole number of minutes) between the patient’s arrival, and the time the A&amp;E attendance has concluded, and the department is no longer responsible for the care of the patient</td>
<td>INTEGER</td>
<td>8</td>
</tr>
<tr>
<td>ethnos</td>
<td>Ethnic category recorded at attendance</td>
<td>CHAR</td>
<td>2</td>
</tr>
</tbody>
</table>

³ This variable has been altered by the CPRD so that it is unique within and across all HES years.
3. **DIAGNOSIS** (hesae_diagnosis.txt)

<table>
<thead>
<tr>
<th>Column name</th>
<th>Description</th>
<th>Type</th>
<th>Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>patid</td>
<td>Encrypted unique key given to a patient in CPRD GOLD or CPRD Aurum [primary key, in combination with aekey]</td>
<td>INTEGER</td>
<td>20</td>
</tr>
<tr>
<td>aekey³</td>
<td>Record identifier (unique in combination with patid) [primary key, in combination with patid]</td>
<td>INTEGER</td>
<td>20</td>
</tr>
<tr>
<td>diag</td>
<td>A&amp;E diagnosis - 6 characters. A 6-character code made up of diagnosis condition (n2), sub-analysis (n1), anatomical area (n2) and anatomical side (an1). Only certain diagnoses contain a sub-analysis</td>
<td>CHAR</td>
<td>6</td>
</tr>
<tr>
<td>diag2</td>
<td>A&amp;E diagnosis - 2 characters. Includes the diagnosis condition (n2) of the 6-character diagnosis code</td>
<td>CHAR</td>
<td>2</td>
</tr>
<tr>
<td>diag3</td>
<td>A&amp;E diagnosis - 3 characters. Includes diagnosis (n2) and the sub-analysis (n1) of the 6-character diagnosis code. If no sub-analysis has been provided, or is not applicable, then the 2-character description is displayed if available.</td>
<td>CHAR</td>
<td>3</td>
</tr>
<tr>
<td>diaga</td>
<td>A&amp;E diagnosis - anatomical area</td>
<td>CHAR</td>
<td>2</td>
</tr>
<tr>
<td>diags</td>
<td>A&amp;E diagnosis - anatomical side</td>
<td>CHAR</td>
<td>1</td>
</tr>
<tr>
<td>diagscheme</td>
<td>Coding scheme in use</td>
<td>INTEGER</td>
<td>1</td>
</tr>
<tr>
<td>diag_order¹</td>
<td>Ordering of diagnosis at attendance, within range 1-12</td>
<td>INTEGER</td>
<td>2</td>
</tr>
</tbody>
</table>

4. **INVESTIGATION** (hesae_investigation.txt)

<table>
<thead>
<tr>
<th>Column name</th>
<th>Description</th>
<th>Type</th>
<th>Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>patid</td>
<td>Encrypted unique key given to a patient in CPRD GOLD or CPRD Aurum [primary key, in combination with aekey]</td>
<td>INTEGER</td>
<td>20</td>
</tr>
<tr>
<td>aekey³</td>
<td>Record identifier (unique in combination with patid) [primary key, in combination with patid]</td>
<td>INTEGER</td>
<td>20</td>
</tr>
<tr>
<td>invest</td>
<td>A&amp;E investigation - 6 characters. A 6-character code made up of investigation (n2) and local sub-analysis (up to an4)</td>
<td>CHAR</td>
<td>6</td>
</tr>
<tr>
<td>invest2</td>
<td>A&amp;E investigation - 2 characters. Consists of the investigation (n2) of the 6-character investigation code</td>
<td>CHAR</td>
<td>2</td>
</tr>
<tr>
<td>invest_order¹</td>
<td>Ordering of investigation at attendance, within range 1-12</td>
<td>INTEGER</td>
<td>2</td>
</tr>
</tbody>
</table>

¹ Variable generated by CPRD.
³ This variable has been altered by the CPRD so that it is unique within and across all HES years.
## 5. TREATMENT (hesae_treatment.txt)

<table>
<thead>
<tr>
<th>Column name</th>
<th>Description</th>
<th>Type</th>
<th>Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>patid</td>
<td>Encrypted unique key given to a patient in CPRD GOLD or CPRD Aurum [primary key, in combination with aekey]</td>
<td>INTEGER</td>
<td>20</td>
</tr>
<tr>
<td>aekey³</td>
<td>Record identifier (unique in combination with patid) [primary key, in combination with patid]</td>
<td>INTEGER</td>
<td>20</td>
</tr>
<tr>
<td>treat</td>
<td>A&amp;E Treatment - 6 characters. Treatment code made up of treatment (n2), sub-analysis (n1) and a local use section (up to an3)</td>
<td>CHAR</td>
<td>6</td>
</tr>
<tr>
<td>treat2</td>
<td>A&amp;E Treatment - 2 characters. Consists of treatment (n2) of the 6-character treatment code</td>
<td>CHAR</td>
<td>2</td>
</tr>
<tr>
<td>treat3</td>
<td>A&amp;E Treatment - 3 characters. Consists of treatment (n2) and the sub-analysis (n1) of the 6-character treatment code. If no sub-analysis has been provided, or is not applicable, then the 2-character description is displayed if available.</td>
<td>CHAR</td>
<td>3</td>
</tr>
<tr>
<td>treat_order¹</td>
<td>Ordering of treatment at attendance, within range 1-12</td>
<td>INTEGER</td>
<td>2</td>
</tr>
</tbody>
</table>

## 6. HEALTH RESOURCE GROUP TABLE (hesae_hrg.txt)

<table>
<thead>
<tr>
<th>Column name</th>
<th>Description</th>
<th>Type</th>
<th>Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>patid</td>
<td>Encrypted unique key given to a patient in CPRD GOLD or CPRD Aurum [primary key, in combination with aekey]</td>
<td>INTEGER</td>
<td>20</td>
</tr>
<tr>
<td>aekey³</td>
<td>Record identifier (unique in combination with patid) [primary key, in combination with patid]</td>
<td>INTEGER</td>
<td>20</td>
</tr>
<tr>
<td>domproc</td>
<td>Dominant Procedure</td>
<td>CHAR</td>
<td>6</td>
</tr>
<tr>
<td>hrgns</td>
<td>Trust derived HRG value</td>
<td>CHAR</td>
<td>3</td>
</tr>
<tr>
<td>hrgnhsvn</td>
<td>Version number of trust derived HRG</td>
<td>CHAR</td>
<td>3</td>
</tr>
<tr>
<td>sushrg</td>
<td>The SUS² Payment by Result (PbR) derived healthcare resource group (HRG) code</td>
<td>CHAR</td>
<td>6</td>
</tr>
<tr>
<td>sushrgvers</td>
<td>SUS² generated HRG version number</td>
<td>NUMERIC</td>
<td>4</td>
</tr>
</tbody>
</table>

## 7. PATIENT PATHWAY (hesae_pathway.txt)

<table>
<thead>
<tr>
<th>Column name</th>
<th>Description</th>
<th>Type</th>
<th>Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>patid</td>
<td>Encrypted unique key given to a patient in CPRD GOLD or CPRD Aurum [primary key, in combination with aekey]</td>
<td>INTEGER</td>
<td>20</td>
</tr>
<tr>
<td>aekey³</td>
<td>Record identifier (unique in combination with patid) [primary key, in combination with patid]</td>
<td>INTEGER</td>
<td>20</td>
</tr>
<tr>
<td>rttperstart⁵</td>
<td>The start date, for the referral to treatment period</td>
<td>DATE</td>
<td>dd/mm/yyyy</td>
</tr>
<tr>
<td>rttperend⁵</td>
<td>The end date, for the referral to treatment period</td>
<td>DATE</td>
<td>dd/mm/yyyy</td>
</tr>
</tbody>
</table>

¹ Variable generated by CPRD
³ This variable has been altered by the CPRD so that it is unique within and across all HES years
⁴ Secondary User Services
⁵ In sets 19 and 21, for the 2019/20 data, we observe that a proportion of the ‘rttperstart’ and ‘rttperend’ variables have a value ‘1900-01-01’. We are advised that some Patient Administration Systems default to this date when no date is entered, although this is not an official default value, and should be treated as invalid.