



Medicines & Healthcare products  
Regulatory Agency



# **Mental Health Data Set and CPRD primary care data Documentation (set 18)**

**Version 1.4**

**Date: 06 December 2019**

# 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.

Version	Affected Areas Summary of Change	Prepared By	Reviewed By
1.0	Initial Draft	Helen Booth Eleanor Yelland	Rachael Williams
1.1	Modified	Helen Booth	Eleanor Yelland
1.2	Modified	Eleanor Yelland	Helen Booth
1.3	Updated for set 17	Eleanor Yelland	Achim Wolf
1.4	Updated for set 18	Rebecca Ghosh	Eleanor Yelland

## Summary of Changes

### Version 1.1

- Updated document for set 15
- Updated header and footer with new agency branding

### Version 1.2

- Updated document for set 16
- Updated to include CPRD Aurum

### Version 1.3

- Updated document for set 17

### Version 1.4

- Updated document for set 18

## **Mental Health Data Set (MHDS) and CPRD primary care data**

This document provides an overview of the Mental Health Data Set (MHDS), and the available subset that is linked to CPRD GOLD and CPRD Aurum.

### **What are the Mental Health Data Set (MHDS) data?**

The Mental Health Data Set is managed by NHS Digital and brings together data recorded by a range of secondary care mental health services providers as a part of routine patient care. It is a secondary uses dataset that was designed to record information about anyone who is thought to be suffering from a mental illness and is receiving specialist adult mental health care in a partially or wholly NHS-funded setting.

The structure of the Mental Health Data Set is shaped by the payment by results system and is regularly updated by NHS Digital, along with the name used to describe it. Originally, it was named the Mental Health Minimum Data Set (MHMDS) but was expanded in 2014 to include information about learning disabilities and autism services. It was subsequently referred to as the Mental Health and Learning Disabilities Data Set (MHLDS). The most recent version of the dataset (version 4) is called the Mental Health Services Data Set (MHSDS). For clarity, CPRD documentation will refer to these data as the Mental Health Data Set (MHDS).

The dataset covers secondary care mental health services provided in hospital, in community settings and at outpatient clinics. The dataset provides information about the type of care received, the length of care episodes, as well as information about contact with mental health and social care professionals and measures of health and social functioning. These include adult mental health services, learning disabilities and autism services (as of 2014), and early intervention care. In the most recent version of the dataset from NHS Digital (MHSDS v4) data for patients aged less than 18 years who are treated within an adult setting or an early intervention service are included. Diagnostic recording is not mandatory and is limited to data recorded within Health of the Nation Outcome Scales (HoNOS) scores or specific event or episode types.

The dataset does not include data about:

- child and adolescent mental health services (CAMHS)
- patients receiving treatment in a specialist mental health service, but not thought to be suffering from a mental illness (e.g. those receiving smoking cessation services)
- high secure services
- primary care mental health services
- independent mental health care providers (not wholly or partially NHS-funded)
- data considered sensitive such as uses of the Mental Health Act 1983 are not included
- Patients attending a service outside of England that is not commissioned by an English Clinical Commissioning Group or NHS England specialised commissioner

## Accessing the Mental Health Data Set linked to CPRD GOLD and CPRD Aurum

MHDS 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 MHRA Independent Scientific Advisory Committee (ISAC) approval. The cost of access depends upon the volume of data required.

Not all patients in CPRD GOLD or CPRD Aurum are eligible to be linked to MHDS, 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 MHDS data. A linkage coverage file (linkage\_coverage.txt) provides the start and end dates of MHDS encounter time.

### Linkage coverage period

Given the frequent structural changes of this dataset, CPRD currently provide the MHDS in two different formats. Format 1 holds data concerning events between **April 2007 and March 2011** and Format 2 holds data concerning events between **April 2011 and November 2015**. The most recent change in the dataset structure (MHSDS v4) from NHS Digital means that a further third format will be available in future holding data from 2015 onwards. Currently the latest release of MHDS data linked to CPRD primary care data (set 18) remains in formats 1 and 2 and covers the period **April 2007 – November 2015**. Further details of the structure of each format and the variables included are provided in the data structure and formatting section, and the data dictionary at the end of this document.

### Linkage algorithm and the match\_rank variable

Linkage between MHDS and CPRD primary care data uses an eight-step deterministic linkage algorithm based on four identifiers, shown in the Table 1 below. 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 MHDS data.

Table 1: NHS Digital 8 step linkage algorithm

Step	Match
1	Exact NHS number, sex, date of birth (DOB), postcode
2	Exact NHS number, sex, DOB
3	Exact NHS number, sex, postcode, partial DOB
4	Exact NHS number, sex, partial DOB
5	Exact NHS number, postcode
6	Exact sex, DOB and postcode (where NHS number does not contradict the match, the DOB is not 1st of January and the postcode is not on the communal establishment list)
7	Exact sex, DOB and postcode (where the NHS number does not contradict the match and the DOB is not 1st of January)
8	Exact NHS number

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. Linkage of CPRD and MHDS data is performed at the patient level. CPRD provides users with a match\_rank variable in the MHDS Patient file (mhds\_patient.txt) which corresponds to the step at which the match between a patient in CPRD and MHDS 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 MHDS dataset. Matching results are summarised in Table 2A and 2B below.

Table 2A: Number and proportion of **CPRD GOLD** patients matched to a MHDS patient at each step of the linkage algorithm in set 18.

match_rank	Frequency	Percent
1	314,878	35.7%
2	497,214	56.4%
3	2,599	0.3%
4	6,718	0.8%
5	20,057	2.3%
6	11,301	1.3%
7	784	0.1%
8	27,456	3.1%

Table 2B: Number and proportion of **CPRD Aurum** patients matched to a MHDS patient at each step of the linkage algorithm in set 18.

match_rank	Frequency	Percent
1	680,336	31.0%
2	1,330,596	60.7%
3	6,322	0.3%
4	18,086	0.8%
5	45,658	2.1%
6	29,435	1.3%
7	2,042	0.1%
8	81,376	3.7%

## Data structure and formatting

MHDS data provided by CPRD represents only a subset of the variables that are collected in the National MHDS dataset provided by NHS Digital. Fields which may lead to the potential re-identification of patients or practices have been pseudonymised. This includes codes for service providers and health commissioners. As far as possible, the linked MHDS data is supplied “as is”, without any modification or cleaning during processing by CPRD. Where CPRD has modified the MHDS data, details are described below.

To reflect changes over time in the structure of the data collected by NHS Digital, data are currently provided in two formats. Format 1 contains data collected between April 2007 and March 2011 (01/04/2007 - 31/03/2011) and is arranged into separate files relating to information on periods (financial year), spells of care, and Health of the Nation Outcome Scales (HoNOS). Format 2 contains data collected between April 2011 and November 2015 (01/04/2011 - 30/11/2015) and is arranged into files relating to information on records (spell of care), episodes of care, and events. There is one patient file that relates to the MHDS data across both formats.

For each patient cohort, MHDS data will be provided as separate text tab delimited files. Files can be uploaded 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 data has been modified for linked patients in the following ways:

- CPRD has generated the MHDS patient identifier (**gen\_mhds\_id**), a patient key to identify a unique patient in the MHDS data. An individual that has contributed data to more than one CPRD practice will have the same patient key (**gen\_mhds\_id**) in the MHDS patient file.
- Licensing obligations require that no attempts are made to re-identify patients in CPRD datasets. The original spell, record, episode, and event identifiers have been encoded by the CPRD to minimise the risk of breaching licensing conditions through linkage of these data to other data sources containing patient identifiable information. This means that these variables may differ in each release of the linked MHDS data.
- CPRD has provided a derived ethnicity variable (**gen\_ethnicity**) in the MHDS patient file which is the most commonly recorded ethnicity for each patient, among all MHDS data between 2007 and 2015. The ethnicity recorded (**ethnicity**) at the spell (pre-2011) or record level (post-2011) remains unchanged.

## Known issues

- HoNOS scores - there are a small number of HoNOS records where the date and type are identified as being the same, but the score itself is different. There is also some incorrect recording of HoNOS dates, with records occurring prior to the score being developed.
- Where commissioner codes were not entered identically they will be encrypted as different codes. It is therefore possible that different encrypted codes refer to the same commissioner.
- Recording of smoking status in the record file (format 2) contains coding errors.

## MHDS: Data dictionary

### 1. Patient file (MHDS\_PATIENT.txt)

<i>Column name</i>	<i>Description</i>	<i>Type</i>	<i>Format</i>
patid	Encrypted unique key given to a patient in CPRD GOLD or CPRD Aurum [primary key]	INTEGER	20
pracid	Encrypted unique key given to a practice in CPRD GOLD or CPRD Aurum	INTEGER	5
gen_mhds_id	MHDS patient key. A generated unique key assigned to a patient in the MHDS data. An individual that has contributed data to more than one CPRD practice has the same gen_mhds_id.	INTEGER	20
n_patid_mh	Number of individuals in CPRD GOLD or CPRD Aurum assigned the same gen_mhds_id	INTEGER	4
match_rank <sup>1</sup>	Indicates the quality of matching between a record in MHDS and CPRD primary care data and gives the level of confidence that an MHDS record has been correctly matched to a patient in CPRD GOLD or CPRD Aurum.	INTEGER	1
pre_2011	Patient present in Format 1 data (01/04/2007 and 31/03/2011)	INTEGER	1
post_2011	Patient present in Format 2 data (01/04/2011 and 30/11/2015)	INTEGER	1
gen_ethnicity	CPRD derived patient ethnicity	CHAR	10
gen_firstknown	CPRD derived variable indicating the year in which the patient is first known to have received care from any mental health service	INTEGER	4
gen_propsych_date	CPRD derived earliest record of prodromal psychosis, indicating the first noticeable change in behaviour or mental state prior to emergence of full-blown psychosis	DATE	dd/mm/yyyy
gen_emerpsych_date	CPRD derived earliest record of emergent psychosis indicating the first clear evidence of a positive psychotic symptom for the patient (i.e. delusion, hallucination, or thought disorder)	DATE	dd/mm/yyyy
gen_manpsych_date	CPRD derived earliest record of manifest psychosis. This is the date at which a positive psychotic symptom has lasted for a week	DATE	dd/mm/yyyy

---

<sup>1</sup> An eight-step process is used to match patients in CPRD primary care data (CPRD GOLD or CPRD Aurum) and MHDS 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.

gen_psychpresc_date	CPRD derived earliest record of a prescription for an anti-psychotic medication	DATE	dd/mm/yyyy
gen_psychtreat_start_date	CPRD derived earliest record of the patient starting prescribed anti-psychotic medication and remaining compliant for at least 75% of the time during the subsequent month (based on clinical judgment).	DATE	dd/mm/yyyy



## 2. Period file (MHDS1\_PERIOD.txt)

<i>Column name</i>	<i>Description</i>	<i>Type</i>	<i>Format</i>
patid	Encrypted unique key given to a patient in CPRD GOLD or CPRD Aurum [primary key]	INTEGER	20
period	The financial year that the record represents	INTEGER	4
cpa_standard_days	The number of days that the Care Programme Approach is at Standard level: 'requires the support or intervention of one agency or discipline; or requires only low-key support from more than one agency'	INTEGER	8
cpa_enhanced_days	The number of days that the Care Programme Approach is at the Enhanced level: 'multiple needs, including housing, employment etc, which requires inter-agency coordination'	INTEGER	8
sssa_number_community_care	Number of social services statutory assessments for community care within the reporting period.	INTEGER	8
bed_days_mh	Total bed days	INTEGER	8
bed_days_mh_medium_secure	Total number of bed days where the patient was admitted to an 'organisation site, service point or ward with a secure accommodation type classification b: 'medium secure accommodation, a secure facility providing care at a regional level under the care of a forensic psychiatrist' ' where the main treatment is for an adult mental illness. This excludes high security accommodation in hospital site approved to provide high security psychiatric services.	INTEGER	8
bed_days_mh_intensive	Total number of bed days where the patient was admitted to a 'ward with a clinical care intensity ('specially designated ward for patients needing containment and more intensive management') where the main treatment is for an adult mental illness. This includes hospital stays and care home stays in consultant care within the hospital provider spell.	INTEGER	8

care_days_acute_home_based	Total number of acute home-based care days	INTEGER	8
bed_days_nhs_community_care	Total number of bed days in a care home site operated and managed by an NHS organisation where the broad patient group is 'Patients with mental illness'.	INTEGER	8
residential_mh_non_nh_ccare_ind	Binary indicator of the occurrence of residential community mental health care provided by a non-NHS provider. (0 = no stays in non-NHS care home during the reporting period; 1= one or more stays in non-NHS care home during reporting period).	INTEGER	1
daycare_attend_mh_nhssite	Total number of attendances to a day care facility where the function classification is 'mental illness'. Facility are either solely NHS run or jointly run between NHS and non-NHS organisations.	INTEGER	8
daycare_attend_mh_nonnhssite_ind	Binary indicator of whether the patient attended a solely non-NHS funded day care facility with the function classification 'mental illness'.	INTEGER	1
sheltered_work_attendance_ind	Binary indicator of whether or not one or more sheltered work attendances or sessions occurred within the reporting period (0 = no attendance at a sheltered work facility; 1=one or more attendances at a sheltered work facility).	INTEGER	1
outpatient_attendance_consultant	Outpatient consultant appointment attendance	INTEGER	8
contacts_community_psych_nurse	Total number of contacts with the community psychiatric nurse group.	INTEGER	8
contacts_clinical_psychologist	Total number of contacts with a clinical psychologist within the reporting period	INTEGER	8
contacts_occupational_therapist	Total number of contact with an occupational therapist within the reporting period	INTEGER	8
social_worker_involvement	Binary indicator of whether a patient's care plan includes involvement with local authority social services social workers. (0 = no involvement of social worker; 1 = involvement of social worker)	INTEGER	1
home_help_visit_ind	Binary indicator of whether a home help visit occurred during the reporting period (0=no home help visit occurred; 1= one or more home help visits occurred)	INTEGER	1

ect_treatments_administered	Total number of Electro-Convulsive Therapy (ECT) administrations within the reporting period.	INTEGER	8
admissions	Total number of admissions within the reporting period	INTEGER	8
discharges	Total number of discharges within the reporting period	INTEGER	8
contacts_physiotherapist	Total number of contacts with a physiotherapist within reporting period.	INTEGER	8
contacts_consultant_psychotherapy	Total number of contacts with a consultant psychotherapist within reporting period.	INTEGER	8
contacts_social_worker	Total number of face to face contacts with a social worker within the reporting period	INTEGER	8
outpatient_dna_mh	Total number of 'Did not attend - no advance warning given' classifications on outpatient appointments within the reporting period	INTEGER	8
day_care_dna_mh_nhssite	Total number of 'Did not attend - no advance warning given' classifications on day care sessions within the reporting period	INTEGER	8
contacts_nhsdirect_mh	Total number of contacts with NHS direct that NHS direct are required to share with a Mental Health Trust within the reporting period	INTEGER	8
care_programme_approach_reviews	Total number of care programme approach reviews within the reporting period	INTEGER	8

### 3. MHDS Spell (MHDS1\_SPELL.txt)

<i>Column name</i>	<i>Description</i>	<i>Type</i>	<i>Format</i>
patid	Encrypted unique key given to a patient in CPRD GOLD or CPRD Aurum [primary key, in combination with e_care_spell_identifier]	INTEGER	20
e_care_spell_identifier	CPRD generated unique identifier for each spell of care [primary key, in combination with patid]	INTEGER	12
start_date_mhcs	Mental health care spell start date	DATE	dd/mm/yyyy
end_date_mhcs	Mental health care spell end date	DATE	dd/mm/yyyy
spell_definition_type	Defines methods used to derive start and end date of spell (A: derived from the succession of recorded contacts; D: declared explicitly; DX: declared explicitly but the dates have been adjusted to accommodate contacts that occurred outside of declared or derived dates; E: derived from the collation of recorded episodes; EX: derived from the collation of recorded episodes but adjusted for contacts that occurred outside these dates)	CHAR	9
e_code_of_provider	Encoded code for care provider	INTEGER	8
e_code_of_commissioner	Encoded code for commissioner	INTEGER	8
marital_status	Recorded marital status	CHAR	9
Ethnicity	Recorded ethnicity	CHAR	10

#### 4. MHDS HoNOS file (MHDS1\_HONOS.txt)

<i>Column name</i>	<i>Description</i>	<i>Type</i>	<i>Format</i>
patid	Encrypted unique key given to a patient in CPRD GOLD or CPRD Aurum [primary key]	INTEGER	20
honos_date	Date of HoNOS score	DATE	dd/mm/yyyy
honos_type	Type of HoNOS event reported on 1 = First, 2= Most recent in period, 3=Worst ever, 4=Best in last 12 months	INTEGER	1
honos	HoNOS score encoded as 13-character string (Each character position refers to a given category. E.g. The 1st character refers to 'overactive, aggressive, disruptive or agitated behaviour'. Each character value encodes the severity of the problem. The category of the most severe clinical problem being rated by items 8-12 is indicated by a letter at position 8). See glossary for HoNOS scoring here: <a href="http://bjp.rcpsych.org/content/bjprcpsych/174/5/432.full.pdf">http://bjp.rcpsych.org/content/bjprcpsych/174/5/432.full.pdf</a>	CHAR	13

## 5. Record file (MHDS2\_RECORD.txt)

<i>Column name</i>	<i>Description</i>	<i>Type</i>	<i>Format</i>
patid	Encrypted unique key given to a patient in CPRD GOLD or CPRD Aurum [primary key, in combination with e_care_record_spell]	INTEGER	20
e_care_record_spell	CPRD generated unique identifier for each record spell of care [primary key, in combination with patid]	INTEGER	12
start_date_mhcs	Start date of mental health care spell	DATE	dd/mm/yyyy
end_date_mhcs	End date of mental health care spell	DATE	dd/mm/yyyy
health_care_spell_end_reason	Present if mental health care spell ended before end date. It provides the reason for ending the care spell. (00 = Finished on professional advice; 01 = Finished against professional advice; 02 = Finished by patient's non-attendance; 03 = Patient died; 21 = Transfer to medium secure; 23 = Transfer to high secure; 25 = Transfer to other health provider)	CHAR	9
e_code_of_provider	Encoded code for care provider	INTEGER	8
marital_status	Marital status	CHAR	9
ethnicity	Original ethnicity	CHAR	10
gen_crisisplancreat_date	Crisis plan creation date – CPRD derived using latest crisisplancreat_date within spell	DATE	dd/mm/yyyy
gen_crisisplanupdate_date	Crisis plan update date - CPRD derived using latest crisisplanupdate_date within spell	DATE	dd/mm/yyyy
gen_statusdate_smoking	Smoking status date - CPRD derived from latest statusdate_smoking within spell	DATE	dd/mm/yyyy
smokstat	Smoking status (1 = Current smoker 2= Ex-smoker 3= Non-smoker – history unknown 4= Never smoked Z= Not stated). Added 2014	CHAR	9
pcp_disability_score	Patient provided score of perception of disability impact (Disability Impact Perception) (01 = Yes -limited a lot; 02 = Yes - limited a little; 03 = No- not limited; 04 = Prefer not to say). Added 2014	CHAR	9
pcp_behem_score	Patient provided score of perception of behaviour and emotional disability impact (Disability Impact Perception) (01 = Yes - limited a lot; 02 = Yes - limited a little; 03 = No- not limited; 04 = Prefer not to say). This disability category is used where the patient has times when they lack control over their feelings or actions. Added 2014	CHAR	8
pcp_hearing_score	Patient provided score of perception of hearing disability impact (Disability Impact Perception) (01 = Yes -limited a lot; 02 =	CHAR	9

	Yes - limited a little; 03 = No- not limited; 04 = Prefer not to say). This disability category is used where the patient has difficulty hearing, or need hearing aids, or need to lip-read what people say. Added 2014		
pcp_manual_score	Patient provided score of perception of disability impact (Disability Impact Perception) (01 = Yes -limited a lot; 02 = Yes - limited a little; 03 = No- not limited; 04 = Prefer not to say). This disability category is used where the patient experiences difficulty performing tasks with their hands. Added 2014	CHAR	9
pcp_memorychild_score	Patient provided score of disability impact (Disability Impact Perception) (01 = Yes - limited a lot; 02 = Yes - limited a little; 03 = No- not limited; 04 = Prefer not to say). This disability category is used where the patient has difficulty with memory or ability to concentrate, learn or understand which began before the age of 18. Added 2014	CHAR	9
pcp_memoryadult_score	Patient provided score of perception emotional disability impact (Disability Impact Perception) (01 = Yes -limited a lot; 02 = Yes - limited a little; 03 = No- not limited; 04 = Prefer not to say). This disability category is used where the patient has difficulty with memory or ability to concentrate, learn or understand which began after the age of 18. Added 2014	CHAR	9
pcp_mobility_score	Patient provided score of perception of mobility disability impact (Disability Impact Perception) (01 = Yes -limited a lot; 02 = Yes - limited a little; 03 = No- not limited; 04 = Prefer not to say). This disability category is used where the patient has difficulty getting around physically without assistance or needs aids like wheelchairs or walking frames; or where the patient has difficulty controlling how their arms, legs or head move. Added 2014	INTEGER	8
pcp_danger_score	Patient provided score of perception disability impact (Disability Impact Perception) (01 = Yes -limited a lot; 02 = Yes - limited a little; 03 = No- not limited; 04 = Prefer not to say). This disability category is used where the patient has difficulty understanding that some things, places or situations can be dangerous and could lead to a risk of injury or harm. Added 2014	CHAR	9
pcp_personal_score	Patient provided score of perception of disability impact (Disability Impact Perception) (01 = Yes -limited a lot; 02 = Yes - limited a little; 03 = No- not limited;	CHAR	8

	04 = Prefer not to say). This disability category is used where the patient has difficulty keeping clean and dressing the way they would like to. Added 2014		
pcp_progressive_score	Patient provided score of perception of disability impact (Disability Impact Perception) (01 = Yes -limited a lot; 02 = Yes - limited a little; 03 = No- not limited; 04 = Prefer not to say). This disability category is used where the patient has any illness which affects what they can do, or which is making the more ill, which is getting worse, and which is going to continue getting worse. Added 2014	INTEGER	8
pcp_sight_score	Patient provided score of perception of disability impact (Disability Impact Perception) (01 = Yes -limited a lot; 02 = Yes - limited a little; 03 = No- not limited; 04 = Prefer not to say). This disability category is used where the patient has difficulty seeing signs or things printed on paper, or seeing things at a distance. Added 2014	INTEGER	8
pcp_speech_score	Patient provided score of perception of disability impact (Disability Impact Perception) (01 = Yes -limited a lot; 02 = Yes - limited a little; 03 = No- not limited; 04 = Prefer not to say). This disability category is used where the patient has difficulty speaking or using language to communicate or make their needs known. Added 2014	CHAR	9
pcp_autism_score	Patient provided score of perception of disability impact (Disability Impact Perception) (01 = Yes -limited a lot; 02 = Yes - limited a little; 03 = No- not limited; 04 = Prefer not to say). Added 2014	INTEGER	8
pcp_other_score	Patient provided score of perception of disability impact (Disability Impact Perception) (01 = Yes -limited a lot; 02 = Yes - limited a little; 03 = No- not limited; 04 = Prefer not to say). This disability category is used where the patient has any other important health issue. Added 2014	INTEGER	8



## 6. Episode file (MHDS2\_EPISODE.txt)

<i>Column name</i>	<i>Description</i>	<i>Type</i>	<i>Format</i>
patid	Encrypted unique key given to a patient in CPRD GOLD or CPRD Aurum [primary key, in combination with e_care_record_spell & e_episode_id]	INTEGER	20
e_care_record_spell	CPRD generated unique identifier for each record spell of care [primary key, in combination with patid & e_episode_id]	INTEGER	12
e_episode_id	CPRD generated unique identifier for each episode [primary key, in combination with patid & e_care_record_spell]	INTEGER	12
epitype <sup>2</sup>	Episode type	CHAR	9
epistart_date	Episode start date. Replaced with delayed_dischstart_date if missing, and epitype is 'DDISC', or refrec_date if missing and epitype is 'REF'.	DATE	dd/mm/yyyy
epiend_date	Episode end date. Replaced with delayed_dischend_date if missing.	DATE	dd/mm/yyyy
ic_inactive_timestamp		DATE	dd/mm/yyyy
data1	Dependent on epitype <sup>6</sup>	CHAR	9
data2	Dependent on epitype <sup>6</sup>	CHAR	9

---

<sup>2</sup> Each episode type may be associated with up to two data fields. Content of each data field is dependent on the episode type – the fields may contain raw data values or may be encoded values that represent text. The file epitype\_descriptions.xls contains information on all episode types, description of the data in each field, and information needed to decode the data.

## 7. Event file (MHDS2\_EVENT.txt)

<i>Column name</i>	<i>Description</i>	<i>Type</i>	<i>Format</i>
patid	Encrypted unique key given to a patient in CPRD GOLD or CPRD Aurum [primary key, in combination with e_care_record_spell & e_event_id]	INTEGER	20
e_care_record_spell	CPRD generated unique identifier for each record spell of care [primary key, in combination with patid & e_event_id]	INTEGER	12
e_event_id	CPRD generated unique identifier for each event [primary key, in combination with patid & e_care_record_spell]	INTEGER	12
eventtype <sup>3</sup>	Type of event recorded	CHAR	9
eventdate	Date of event	DATE	dd/mm/yyyy
data1	Dependent on event type <sup>8</sup>	CHAR	9
data2	Dependent on event type <sup>8</sup>	CHAR	9
data3	Dependent on event type <sup>8</sup>	CHAR	9
data4	Dependent on event type <sup>8</sup>	CHAR	9
data5	Dependent on event type <sup>8</sup>	CHAR	9
data6	Dependent on event type <sup>8</sup>	CHAR	9
data7	Dependent on event type <sup>8</sup>	CHAR	9
data8	Dependent on event type <sup>8</sup>	CHAR	9
data9	Dependent on event type <sup>8</sup>	CHAR	9
data10	Dependent on event type <sup>8</sup>	INTEGER	8
data11	Dependent on event type <sup>8</sup>	INTEGER	8
data12	Dependent on event type <sup>8</sup>	INTEGER	8
data13	Dependent on event type <sup>8</sup>	INTEGER	8
data14	Dependent on event type <sup>8</sup>	INTEGER	8
data15	Dependent on event type <sup>8</sup>	INTEGER	8
data16	Dependent on event type <sup>8</sup>	INTEGER	8
data17	Dependent on event type <sup>8</sup>	INTEGER	8
data18	Dependent on event type <sup>8</sup>	INTEGER	8
data19	Dependent on event type <sup>8</sup>	INTEGER	8
data20	Dependent on event type <sup>8</sup>	INTEGER	8
data21	Dependent on event type <sup>8</sup>	CHAR	9
data22	Dependent on event type <sup>8</sup>	INTEGER	8

---

<sup>3</sup> Each event type may be associated with up to 22 data fields. Content of each data field is dependent on the event type – the fields may contain raw data values or may be encoded values that represent text. The file eventtype\_descriptions.xls contains information on all event types, description of the data in each field, and information needed to decode the data.