

Data set structure

OST_prescriptions

- **Total Records:** 6,163,698 records
- **Fields:**
 - **IAIN:** Unique patient identifier
 - **patient_date_of_birth:** Patient's date of birth
 - **patient_sex:** Patient's gender
 - **paid_month_end_date:** End date of the month when the prescription was paid
 - **patient_health_board_name:** Name of the NHS Health Board to which the patient belongs

paid_month_end_date

- **Minimum Value:** 2009-04-30
- **Maximum Value:** 2023-03-31

lookup_NRS_all_cause_deaths (All-Cause Deaths Dataset)

- **Number of Records:** 12,701 records
- **Time Range:** 2009-02-07 to 2023-12-27

Variable Name	Variable Name	Data Type	Description
IAIN	Anonymous Index Number	numeric	Unique personal identifier (anonymized)
NRS_DATE_OF_DEATH	Date of Death	Date	Records the date of death of an individual

NRS_RECORD_TYPE	Record Type	character	Type of death record (e.g., "99B")
NRS_DATE_OF_REGISTRATION	Date of Registration	Date	Date of death registration
NRS_UNDERLYING_CAUSE_OF_DEATH	Fundamental Cause of Death (ICD-10 code)	haven_labelled (character)	WHO ICD-10 code, records the primary cause of death (e.g., "X42" = accidental drug poisoning)
NRS_CAUSE_OF_DEATH_CODE_0 to NRS_CAUSE_OF_DEATH_CODE_9	Other Causes of Death Codes (ICD-10)	haven_labelled (character)	Other related causes of death (up to 10), may include complications, comorbidities, etc.

lookup_NRSdrd_drug_related_deaths (Drug-Related Deaths Dataset)

- **Number of Records:** 6,019 records
- **Time Range:** 2009-04-04 to 2023-03-31

Variable Name	Variable Name	Data Type	Description
IAIN	Anonymous Index Number	numeric	Unique personal identifier (consistent with the all-cause deaths dataset)
NRSdrd_HMMBimplic	Implication of Heroin/Morphine/Methadone/Buprenorphine	haven_labelled (numeric)	1 = "Yes" (92.77% of records involve these drugs)
NRSdrd_UKdefn	Compliance with UK Drug-Related Death Definition	haven_labelled (numeric)	1 = "Yes" (all records comply)

NRSdrd_UNDERLYING_CAUSE_OF_DEATH	Fundamental Cause of Death (ICD-10 code)	haven_labelled (character)	Primary ICD-10 code for drug-related deaths (e.g., "X41" = accidental antidepressant poisoning)
NRSdrd_DOD	Date of Death	Date	Records the date of death of an individual
NRSdrd_FYOD	Fiscal Year of Death	haven_labelled (numeric)	Fiscal year based on the date of death (e.g., 2022 = "2022/23")
NRSdrd_poison	Poisonous Drug	character	Listed drugs that directly caused or contributed to death (e.g., "METHADONE, COCAINE")
NRSdrd_alsopres	Other Drugs Present	character	Drugs found in the body but did not directly cause death (e.g., "Diazepam")
type_of_death	Type of Death	haven_labelled (numeric)	1 = "Suicide/Homicide" (X60-X64, X85) 2 = "Undetermined Intent" 3 = "Accidental" (main type)

lookup_SMR0104_hospital_stays (Hospital Stays Dataset)

1. Dataset Overview

- **Number of Records:** 286,984 records
- **Time Range:** 1994-06-07 to 2023-12-27 (admission date)
- **Main Features:**

- Includes records of acute and psychiatric hospital stays
- Each hospital stay can contain up to 40 episodes, with one primary diagnosis and five other diagnoses per episode
- Diagnoses are recorded using ICD-10 codes

2. Core Variables (English and Chinese)

Variable Name	Variable Name	Data Type	Description
IAIN	Anonymous Index Number	numeric	Unique personal identifier (anonymized)
RECORD_TYPE	Record Type	haven_labelled	"01B" = Acute Hospital Stay "04B" = Psychiatric Hospital Stay
CIS_MARKER	Continuous Hospital Stay Identifier	numeric	Unique identifier for the same continuous hospital stay period
ADMISSION_DATE	Admission Date	Date	Patient admission date (earliest from 1994)
DISCHARGE_DATE	Discharge Date	Date	Patient discharge date (from 2009 onwards)
flag_discharge_death	Discharge Death Flag	numeric	1 = Death at discharge 0 = Survived discharge
flag_no_disdate_last_episode	Last Episode No Discharge Date Flag	numeric	1 = No discharge date in the last episode

3. Episode Diagnosis Variables

- The dataset includes up to 40 episodes (EPI_1 to EPI_40), with each episode having:
 - Primary diagnosis (MAIN_CONDITION)
 - Other diagnoses 1-5 (OTHER_CONDITION_1 to 5)

Example Structure (for the first episode)

Variable Name	Variable Name	Data Type	Description
EPI_1_MAIN_CONDITION	Primary Diagnosis of Episode 1	character	ICD-10 code (e.g., "T402" = Heroin poisoning)
EPI_1_OTHER_CONDITION_1	Other Diagnosis 1 of Episode 1	character	Complications or other diagnoses
...
EPI_1_OTHER_CONDITION_5	Other Diagnosis 5 of Episode 1	character	Complications or other diagnoses

lookup_SMR0104drhs_drug_related_hospital_stays
(Drug-Related Hospital Stays Dataset)

1. Dataset Overview

- **Number of Records:** 97,850 records
- **Time Range:** 2005-10-10 to 2023-12-27 (admission date)
- **Main Features:**
 - Specifically records hospital stays related to drug use
 - Includes records of acute and psychiatric hospital stays
 - Diagnoses are recorded using ICD-10 codes, with a focus on drug-related diagnoses
 - The data structure is the same as the all-cause hospital stays dataset but filtered for drug-related cases

2. Core Variables (English and Chinese)

Variable Name	Variable Name	Data Type	Description

IAIN	Anonymous Index Number	numeric	Unique personal identifier (anonymized)
RECORD_TYPE	Record Type	haven_labelled	"01B" = Acute Hospital Stay "04B" = Psychiatric Hospital Stay
CIS_MARKER	Continuous Hospital Stay Identifier	numeric	Unique identifier for the same continuous hospital stay period
ADMISSION_DATE	Admission Date	Date	Patient admission date
DISCHARGE_DATE	Discharge Date	Date	Patient discharge date
flag_discharge_death	Discharge Death Flag	numeric	1 = Death at discharge 0 = Survived discharge
flag_no_disdate_last_episode	Last Episode No Discharge Date Flag	numeric	1 = No discharge date in the last episode

3. Episode Diagnosis Variables

- Same structure as the all-cause hospital stays dataset, including up to 40 episodes (EPI_1 to EPI_40), with each episode having:
 - Primary diagnosis (MAIN_CONDITION)
 - Other diagnoses 1-5 (OTHER_CONDITION_1 to 5)

4. Data Characteristics

a. Relationship with All-Cause Hospital Stays Dataset:

- This dataset is a subset of the previous dataset (97,850/286,984 \approx 34%)
- Records have been filtered to include those with drug-related ICD codes in the primary or other diagnoses

lookup_SDMD (Scottish Drug Misuse Database)

1. Dataset Overview

- **Number of Records:** 258,016 records
- **Time Range:** 2009-04-01 to 2021-03-31

2. Core Variables (English and Chinese)

Basic Information

Variable Name	Variable Name	Data Type	Description
IAIN	Anonymous Index Number	numeric	Unique personal identifier
anonymised_EOCN	Anonymous Treatment Number	integer	Unique identifier for the treatment record
recordtype	Record Type	haven_labelled	1 = Initial Assessment 2 = Follow-up
dateass	Assessment Date	Date	Initial assessment date
revdate	Review Date	Date	Follow-up review date

Drug Use Information

Variable Name	Variable Name	Data Type	Description
ill_anydrug	Use of Any Illegal Drug	haven_labelled	1 = Yes
ill_opioids	Use of Opioids	haven_labelled	1 = Yes
ill_cocaine	Use of Cocaine	haven_labelled	1 = Yes
drugfree	Drug-Free in the Past Month	haven_labelled	1 = Yes 2 = No

Detailed Records of Illegal Drugs (up to 5 types)

Variable Name	Variable Name	Data Type	Description
igr1-igr5	Classification of Illegal Drugs (Groups)	haven_labelled	Drug type code (e.g., 1 = Heroin)
i1-i5	Specific Code of Illegal Drugs	haven_labelled	Detailed drug code
mroute1-mroute5	Route of Administration	haven_labelled	e.g., 1 = Intravenous Injection

Prescription Drug Records

Variable Name	Variable Name	Data Type	Description
pres_opioids	Prescription Opioids	haven_labelled	1 = Yes
pres_ost	Opioid Substitution Therapy	haven_labelled	1 = Yes
pgr1-pgr5	Classification of Prescription Drugs (Groups)	haven_labelled	Drug type code
pres1-pres5	Specific Code of Prescription Drugs	haven_labelled	Detailed drug code

Other Information

Variable Name	Variable Name	Data Type	Description
newinj	Injection Behavior	haven_labelled	1 = Injected in the Past Month 3 = Never Injected

accom	Accommodation Type	haven_labelled	1 = Owned/Rented 4 = Homeless
-------	--------------------	----------------	---------------------------------------

3. Key Coding Information

Drug Classification Codes (igr/pgr)

Code	Drug Type
1	Heroin
2	Morphine
3	Methadone
10	Buprenorphine
20	Cocaine

Route of Administration Codes (mroute)

Code	Route
1	Intravenous Injection
3	Inhalation
11	Oral

lookup_DAlSy (Drug Abuse Information System Dataset)

1. Dataset Overview

- Number of Records: 32,939 records

- **Time Range:** 2020-10-23 to 2023-03-31
- **Main Features:**
 - Includes assessment and review records for drug abuse treatment
 - Detailed records of illegal and prescription drug use
 - Includes detailed information on routes of administration, injection behavior, etc.
 - The dataset structure is more detailed than SDMD, recording up to 10 types of drug use

2. Core Variables (English and Chinese)

Basic Information

Variable Name	Variable Name	Data Type	Description
IAIN	Anonymous Index Number	numeric	Unique personal identifier
anonymised_EOCN_DAI_Sy	Anonymous Treatment Number	integer	Unique identifier for the treatment record
assessment_type_assessment_category_description	Assessment Type	character	"Assessment" = Assessment "Review" = Review
assessment_date	Assessment Date	Date	Assessment/Review date

Drug Use Information

Variable Name	Variable Name	Data Type	Description
ill_anydrug	Use of Any Illegal Drug	haven_labelled	1 = Yes
ill_opioids	Use of Opioids	haven_labelled	1 = Yes (40% of records)
ill_cocaine	Use of Cocaine	haven_labelled	1 = Yes (15.9% of records)

assessment_drug_use_history_description	Drug Use History	character	e.g., "Used in the past month"
---	------------------	-----------	--------------------------------

Detailed Records of Illegal Drugs (up to 10 types)

Variable Name	Variable Name	Data Type	Description
drug_use_is_main_drug_1-10	Is Main Drug	logical	TRUE/FALSE
drug_category_description_1-10	Drug Category	character	e.g., "Opioids"
drug_type_description_1-10	Specific Drug	character	e.g., "Heroin"
drug_type_code_1-10	Drug Code	character	e.g., "1111" = Heroin
drug_route_description_1-10	Route of Administration	character	e.g., "Smoke"

Prescription Drug Records

Variable Name	Variable Name	Data Type	Description
pres_depdrug_opioids	Prescription Opioids	haven_labelled	1 = Yes (79.4% of records)
pres_depdrug_ost	Opioid Substitution Therapy	haven_labelled	1 = Yes (79.2% of records)
assessment_has_been_prescribed_meds_for_substance_abuse_description	Prescription Drug Use History	character	"Yes"/"No"
presc_drug_category_description_1-10	Prescription Drug Category	character	

presc_drug_type_description_1-10	Specific Prescription Drug	character	e.g., "Methadone"
----------------------------------	----------------------------	-----------	-------------------

Other Information

Variable Name	Variable Name	Data Type	Description
assessment_injection_history_description	Injection History	character	e.g., "Never"
assessment_accommodation_description	Accommodation Type	character	e.g., "Owned/rented - stable"

3. Key Data Characteristics

a. Drug Use Patterns:

- 61.25% of records show use of illegal drugs
- Heroin is the most common illegal drug (main drug 1 is often heroin)
- Inhalation is the most common route of administration

b. Treatment Features:

- Nearly 80% of patients receive opioid or substitution therapy
- Methadone is the most common prescription drug

c. Data Structure:

- Can record up to 10 types of drug use
- Contains more detailed descriptive information than the SDMD dataset