Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Institution - role (18 variables)	**A1.1**	**Yes**	To Keep	**Institution ID**	A persistent identifier for an institution. This may be created by the catalogue maintainer on first entry of an institution. It provides links across tables and between catalogues.		Manual entry or an existing identifier or creation by the catalogue maintainer on first entry of the institution to the catalogue.	A
Institution - role	**A1.2**	**Yes**	To keep	**Institution full name**	Official name of the institution or organisation as used in EU projects.	Free text	Manual entry	A
Institution - role	**A1.3**	**Yes**	Not a must	**Institution acronym**	Official acronym of the institution, if applicable.	Free text	Manual entry	A
Institution - role	**A1.4**	**Yes**	To Keep	**Type of institution**	In which sector is the institution? Select one of the following	Academic institution/ civic authority/ government agency/ healthcare payer/ network of primary care practices/ non-profit organisation/ pharmaceutical industry/ private organisation/ public health authority/ research centre/ statistical authority/CRO	Manual entry	A
Institution - role	**A1.4.1**	**Yes**	Not a must	**Type of organisation, other**	If other, description of type of organisation.	Free text	Manual entry	A
Institution - role	**A1.5**	**Yes**	To Keep	**Institution country**	Country in which organisation head office or coordinating centre is located.	Select from list of countries, links to ISO 3166-1 country codes (https://www.iso.org/obp/ui/#search/code/)	Manual entry	A
Institution - role	**A1.5.1**	**Yes**	Not a must	**Institution full address**	Full address where the organisation head office or coordinating centre is located.	Free text	Manual entry	A
Institution - role	A1.6	No	To Keep	Institution website	If applicable, a dedicated website for the organisation can be entered.	URL	Manual entry	A

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Institution - role	**A1.7**	**Yes**	To Keep	**Institution email contact**	Contact email address for person responsible for institution entry in catalogue.	Free text; must take form "prefix@domain"	Manual entry	A
Institution - role	**A1.8**	**Yes**	To Keep	**Institution role**	Roles of the institution in connection with data sources and data banks in the catalogue. Select one or more of the following:	Data originator/ research/ surveillance/ other	Manual entry	A
Institution - role	**A1.8.1**	**Yes**	Not a must	**Institution role, other**	If other, description of the primary role.	Free text	Manual entry	A
Institution - role	A1.9	No	Not a must	Institution expertise	Free text summary of experience and expertise. Could relate to 1 or more data sources in the catalogue.	Free text	Manual entry	A
Institution - role	**A1.10**	**Yes**	To Keep	**Institution networks**	Is the institution a member of an existing research network or consortium? If so, provide a list of the names of the networks or consortia the institution participates in.	Should take values within E1.1. If required value is not present in E1.1, an entry must be made in table E1-Network-Overview	Manual entry	A
Institution - role	A1.11	No	Not a must	Institution resources	Does the institution team responsible for working with data sources in the catalogue include members with any of the following expertise? Select all that apply.	Epidemiologist/ pharmacoepidemiologist/ statistician/ clinician/ clinical pharmacologist/ pharmacist/ geneticist/ pharmacogeneticist/ IT specialist/ ethics expertise/ legal expertise/ regulatory expertise	Manual entry	A

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Institution - role	A1.12	No	Not a must	Medical specialty	Does the institution or member have a medical specialty in any of the following therapeutic/disease areas? Select all that apply.	Anaesthesia/ Cardiovascular diseases/ Congenital Malformations/ Devices/ Disorders of the central nervous system/ Ear, nose and oropharynx disorders/ Endocrine disorders/ Eye disorders/ Gastrointestinal tract/ Geriatrics/ Gynaecology/ Immunological products and vaccines/ Immunosuppression/ Infectious diseases/ Liver disease/ Malignant disease/ Musculoskeletal and joint diseases/ Neonates/ Nutrition and blood/ Osteoporosis/ Paediatrics/ Poisoning/ Overdose/ Pregnancy/ Psychiatry/ Renal impairment/ Respiratory diseases/ Skin disorders/ Urinary tract disorders/ Other	Manual entry	A
Institution - role	A1.13	No	Not a must	Medical specialty, other	If other, list any additional specialty therapeutic/disease areas.	Free text	Manual entry	A
Institution - role	A1.14	No	Not a must	Study design experience	Does the institution or member have expertise using any of the following study designs? Select all that apply.	Cohort study/ case control/ case only/ drug utilisation/ interventional trial/ routinely collected data-based pragmatic trial/ meta-analysis/ other	Manual entry	A
Institution - role	A1.15	No	Not a must	Study design experience, other	If other, provide a description study design experience.	Free text	Manual entry	A
Institution - data source (13 variables)	**A2.1**	**Yes**	Not a must	**Experience with data sources**	With which of the following data sources present in the catalogue does the institution have experience or expertise? Select one or more data sources from the list	Should take values within B1.1.	Manual entry	A

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Institution - data source	A2.1.1	**Yes**	Not a must	**Experience with data banks**	With which of the following data banks present in the catalogue does the institution have experience or expertise? Select one or more data banks from the list	Should take values within C1.1.	Manual entry	A
Institution - data source	A2.2	**Yes**	To Keep	**Is DAP**	Does the institution have permission to access (an extract of) one or more data sources or data banks in the catalogue for research purposes?	No/yes	Manual entry	A
Institution - data source	A2.2.1	**Yes**	To Keep	**Data sources**	Name of data sources that can currently be accessed	Should take values within B1.1. If required value is not present in B1.1, an entry must be made in table B1-Data source-Underlying population	Manual entry	A
Institution - data source	A2.2.2	**Yes**	Not a must	**Data banks**	Name of data banks that can currently be accessed	Should take values within C1.1. If required value is not present in C1.1, an entry must be made in table C1-Data bank-Underlying population	Manual entry	A
Institution - data source	A2b.2.3	**Yes**	To Keep	**Data bank population subset**	Can the institution access the full data bank or a subset of the data bank subpopulation?	Full/subset	A separate entry is made manually for each data source	A, C
Institution - data source	A2b.2.4	**Yes**	Not a must	**Data bank population subset specification**	If the institution can only access a subset, which subsets can it access? Select all that apply.	Preterm new-borns/ term new-borns (0-27 days)/ infants and toddlers (28 days - 23 months)/ children (2-11 years)/ adolescents (12-17 years)/ adults (18-74 years)/ adults (75 years and over)/ pregnant women/ renal impaired/ hepatic impaired/ immunocompromised/ other	A separate entry is made manually for each data source	A, C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Institution - data source	A2b.2.5	**Yes**	Not a must	**Data bank subset other**	If 'other' is selected above, describe the subset that can be accessed.	Free text	A separate entry is made manually for each data source	A, C
Institution - data source	A2b.2.6	**Yes**	Not a must	**Data access level**	Can the institution access individual level data or only aggregate level data (where individual level data refer to individual records and aggregate level data refer to summary tables)?	Individual/aggregate	A separate entry is made manually for each data source	A, C
Institution - data source	A2.2.7	**Yes**	Not a must	**Data access extract**	Can the institution access the whole data bank or only an extract of the data bank?	Whole/extract	A separate entry is made manually for each data source	A, C
Institution - data source	A2.2.8	No	Not a must	Data access platform	Can the institution access (an extract of) the data bank on their local systems, or is access via a remote access system?	Local/remote	A separate entry is made manually for each data source	A, C
Institution - data source	A2.2.9	No	Not a must	Data access platform details	If data can only be accessed via a remote access system, describe the system and/or software used to access/process data.	Free text	A separate entry is made manually for each data source	A, C
Institution - data source	A2.2.10	No	Not a must	Data security	Describe the data encryption and security protections for the data bank that are put in place by the DAP	Free text	A separate entry is made manually for each data source	A, C
Institution - access contracts (7 variables)	A3.1	**Yes**	Not a must	**Reason for access**	What is the reason for being able to access (an extract of) the data bank? Select one or more from the following:	Originator/ research/ surveillance/ other	A separate entry is made manually for each data source selected in A2.3.1	A, C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Institution - access contracts	**A3.1.1**	**Yes**	Not a must	**Description of other access**	If other, provide a short description.	Free text	A separate entry is made manually for each data source selected in A2.3.1	A, C
Institution - access contracts	**A3.2**	**Yes**	Not a must	**Access permission**	Is permission from an external organization (e.g., a data controller) required for your institute to access (an extract of) the data bank, in order to process the data and publish results of data processing? For example, if approval of a study protocol is needed from an external organization.	No/ always/ study-dependent	A separate entry is made manually for each data source selected in A2.3.1	A, C
Institution - access contracts	**A3.3**	**Yes**	Not a must	**Access agreement termination**	The date of termination of the access agreement, if any.	DDMMYYYY	A separate entry is made manually for each data source selected in A2.3.1	A, C
Institution - access contracts	A3.4	No	Not a must	Access process	If permission to access data bank is required, what is the process for obtaining access for a study?	Free text/URL	A separate entry is made manually for each data source selected in A2.3.1	A, C
Institution - access contracts	**A3.5**	**Yes**	Not a must	**Access process time**	On average, how long does it take for approval/access to be obtained following an application for data bank?	Unit: day	A separate entry is made manually for each data source selected in A2.3.1	A, C
Institution - access contracts	**A3.6**	**Yes**	Not a must	**Access fees**	Are there fees for data access?	No/yes	A separate entry is made manually for each data source selected in A2.3.1	A, C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Institution - studies	A4.1	No	To Keep	Institution studies	A list of IDs of studies that have been conducted by the institution that are recorded in the catalogue	Should take values within F1.1.	This will be automatically populated via the link between Institution-studies and Study-institutions, by linking the Institution ID and Study ID.	A
Data source - underlying population (13 variables)	**B1.1**	**Yes**	To Keep	**Data source ID**	A persistent identifier for a data source. This may be created by the catalogue maintainer on first entry of an institution. It provides links across tables.			B
Data source - underlying population	**B1.2**	**Yes**	To Keep	**Data source name**	Name of data source used in European projects	Free text	Manual entry	B
Data source - underlying population	**B1.2.1**	**Yes**	Not a must	**Data source local name**	If different from above, name of data source in the national language of the data source	Free text	Manual entry	B
Data source - underlying population	**B1.3**	**Yes**	to keep	**Data source acronym**	Official acronym of the institution, if applicable.	Free text	Manual entry	B
Data source - underlying population	**B1.4**	**Yes**	Not a must	**Data source type**	Type of data source. Select one or more of the following:	Biobank/ administrative claims/ electronic health record/ registry/ other	Manual entry	B
Data source - underlying population	**B1.4.1**	**Yes**	Not a must	**Data source type, other**	If other, describe the type of data source.	Free text	Manual entry	B
Data source - underlying population	B1.5	No	Not a must	Description of underlying population	Summary description of the data source population (maximum 100 words) or URL to description	Free text/URL	Manual entry	B

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data source - underlying population	**B1.6**	**Yes**	Manda-tory	**Data source countries**	Countries where data originate (1 or more)	Select from list of countries, links to ISO 3166-1 country codes (https://www.iso.org/obp/ui/#search/code/)	Manual entry; may be automatically populated with entries in C1.5 when data bank entries are made	B
Data source - underlying population	**B1.6.1**	**Yes**	To Keep	**Data source regions**	Geographical regions that data source covers	Select from list of regions, links to ISO 3166-2 region codes (https://www.iso.org/obp/ui/#search/code/)	Manual entry; may be automatically populated with entries in C1.5.1 when data bank entries are made	B
Data source - underlying population	**B1.7**	**Yes**	To Keep	**Care setting**	Type of care setting(s) covered by data source.	General population/ primary care/ secondary care/ tertiary care/ other	Manual entry or may be automatically populated when C1.14 is populated for data banks in B4.1	B
Data source - underlying population	**B1.7.1**	**Yes**	Not a must	**Care setting, other**	If 'other,' description of the setting of care.	Free text	Manual entry or may be automatically populated when C1.14.1 is populated for data banks in B4.1	B
Data source - underlying population	B1.8	No	Not a must	Data bank population overlap	Describe in text the overlap between the underlying populations of data banks within the data source.	Free text	Manual entry; left blank for data sources with only one data bank	B

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data source - access	**B2.1**	**Yes**	Not a must	**DAPs**	List of DAPs that are listed in the catalogue as having (conditional) permission to access (an extract of) the data source	Should take values from A1.1	Automatically populated when a contributor enters A2.3 Ins-data sources-Is DAP as yes and a data source with corresponding Data source ID is selected	B
Data source - access	B2.2	No	To Keep	Governance details	If available, provide documents or webpages that describe the overall governance of the data source (governing data access or utilisation for research purposes by existing DAPs)	Link to documents/URL	Manual entry	B
Data source - access	B2.2.1	No	Not a must	Committee data access	Is there a committee that determines access to the data source (policy data access, committee data access) that are already listed?	No/yes	Manual entry	B
Data source - access	B2.3	No	Not a must	Policy data access	If available, provide documents or webpages that describe the policy for access to the data source.	Link to documents/URL	Manual entry. If data source consists of a single data bank, this may be automatically populated when C2.4 is populated	B
Data source - access	**B2.4**	**Yes**	To Keep	**Informed consent**	Is informed consent required for use of the data for research purposes? (for study)	Not required / Required for general use/ required for all studies/ required for intervention studies/ waiver/ other	Manual entry. If data source consists of a single data bank, this may be automatically populated when C2.4 is populated	B, A

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data source - access	**B2.4.1**	**Yes**	To Keep	**Informed consent other**	If other, describe the conditions when informed consent is required.	Free text	Manual entry. If data source consists of a single data bank, this may be automatically populated when C2.4.1 is populated	B, A
Data source - access	**B2.5**	**Yes**	To Keep	**Access to identifiable data**	Can identifiable data be accessed in the data source by the DAPs that are already listed as having access (for example, patient name/practitioner name/practice name)?	No/yes	Manual entry. If data source consists of a single data bank, this may be automatically populated when C2.5 is populated	B, A
Data source - access	B2.5.1	No	To Keep	Route to access to identifiable data	If yes above, what is the route to access or process this information? What permission is required?	Free text	Manual entry. If data source consists of a single data bank, this may be automatically populated when C2.5.1 is populated	B, A
Data source - access	**B2.6**	**Yes**	To Keep	**Access to subject details**	Can individual patients/practitioners/practices be contacted in the data source?	No/yes	Manual entry. If data source consists of a single data bank, this may be automatically populated when C2.6 is populated	B, A

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data source - access	B2.6.1	No	To Keep	Route to access to subject details	If yes above, what is the route to access this information? What permission is required?	Free text	Manual entry. If data source consists of a single data bank, this may be automatically populated when C2.6.1 is populated	B, A
Data source - access	B2.7	No	To Keep	Audit possibility	Are external parties allowed to audit the data? For example, is it possible for an external party to audit the quality or validity of the data source?	No/yes	Manual entry	B, A
Data source - access	B2.8	No	Manda-tory	Approval for publication	Is the catalogue contributor able to publish results from the data source without interaction from an external organization, such as approval of a protocol?	No/yes	This is entered manually per institution that makes an entry for the data source in the catalogue	B, A
Data source - qualification	**B3.1**	**Yes**	To Keep	**Data source qualification**	Has the data source successfully undergone a formal qualification process (e.g., from the EMA, or ISO or other certifications)?	No/yes	Manual entry	B
Data source - qualification	**B3.1.1**	**Yes**	To Keep	**Description of qualifications**	If yes, describe the qualification(s) granted.	Free text	Manual entry	B
Data source - data banks	**B4.1**	**Yes**	To Keep	**Data banks**	A list of the data banks that might be included in this data source	Takes values from C1.1	Manual entry; selected from a list of data bank names (C1.2) linked to identifier C1.1	B

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data source - data banks	B4.2	Yes	Not a must	Data bank families	A list of the families of data banks visible per data bank included in the data source	Hospital discharge records/ primary care medical records/ pharmacy dispensation records/ birth registry/ induced terminations registry/ congenital anomaly registry/ inhabitant registry/ registration with healthcare system/ outpatient visit records/ emergency care discharge records/ exemptions from copayment/ diagnostic tests or procedures reimbursement/ death registry/ cancer registry/ other disease registry/ vaccine registry/ biobank/ administrative claims/ spontaneous reporting of adverse drug reactions	Manual entry or automatically populated when data banks are entered in B4.1 with values from C5.1	B, C
Data source - data banks	B4.2.1	Yes	Not a must	Data bank families, other	If 'other' selected above, describe the type of data bank.	Free text	Manual entry or automatically populated when data banks are entered in B4.1 with values from C5.1.1	B, C
Data source - data banks	B4.3	Yes	to keep	Geographic regions	Are data on geographic region available?	No/yes	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	B4.4	Yes	to keep	Family linkage	Is any linkage available in the data bank between any of the following persons?	Household/ mother-child/ father-child/ sibling	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data source - data banks	B4.4.1	No	To Keep	Family linkage routine	Is familial linkage created routinely or can linkage be performed on an ad-hoc basis?	Routine/ad hoc	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	B4.4.2	No	Not a must	Family linkage availability	Is familial linkage available for all of the data source population or a subset of individuals (or records)?	Full/subset	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	B4.4.3	No	Not a must	Family linkage availability details	If a familial linkage is available for a subset of individuals (or records) only, provide a short description of the subset.	Free text	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	**B4.5**	**Yes**	To Keep	**Sociodemo-graphics included**	Which of the following sociodemographic information is included? Select all that apply.	None/ age/ sex/ gender/ ethnicity/ country of origin/ socioeconomic status/ marital status/ education level/ type of residency/ living in rural area/ health area/ deprivation index/ pharmaceutical copayment/ other	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	**B4.6**	**Yes**	To Keep	**Lifestyle factors included**	Which of the following lifestyle information is included? Select all that apply.	None/ tobacco use/ alcohol use/ frequency of exercise/ other	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data source - data banks	**B4.7**	**Yes**	To Keep	**Diagnosis codes included**	Are diagnostic codes included?	No/yes	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	B4.7.1	No	To Keep	Diagnosis code vocabulary	If yes, what vocabulary is used? Select from the following:	ICPC/ ICD9/ ICD10/ ICD1/ Read/ SNOMED/ SNOMED CT/ MedDRA/ OPCS/ CCS/ EDC/ Not coded (Free text)/ Other	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	B4.7.2	No	To Keep	Diagnosis code not coded	If Not coded (Free text), approximately what percentage of records are not coded?	Unit: %, not applicable	Manual entry	B
Data source - data banks	B4.7.3	No	Not a must	Diagnosis, other	Is other information on diagnoses of patient outcomes recorded, not in the form of diagnostic codes?	Free text	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	**B4.8**	**Yes**	To Keep	**Hospital admission discharge**	Is information on hospital admission and/or hospital discharge available?	No/yes	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data source - data banks	**B4.8.1**	**Yes**	To Keep	**ICU admission**	Is information on intensive care unit admission available?	No/yes	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	**B4.9**	**Yes**	To Keep	**Cause of death**	Is cause of death available?	No/yes	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	B4.9.1	No	To Keep	Cause of death vocabulary	If yes, what vocabulary is used? Select from the following:	ICPC/ ICD9/ ICD10/ ICD1/ Read/ SNOMED/ SNOMED CT/ MedDRA/ OPCS/ CCS/ EDC/ Not coded (Free text)/ Other	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	B4.9.2	No	To Keep	Cause of death not coded	If Not coded (Free text), approximately what percentage of records are not coded?	Unit: %, not applicable	Manual entry	B
Data source - data banks	B4.9.3	No	To Keep	Cause of death code, other	If 'other,' what vocabulary is used?	Free text	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data source - data banks	**B4.10**	**Yes**	To Keep	**Rare diseases**	Are rare diseases (<1 in 2000 people) captured?	No/yes	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	**B4.11**	**Yes**	To Keep	**Prescriptions included**	Is information on prescription of medications included?	No/yes	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	B4.11.1	No	To Keep	Prescriptions vocabulary	If yes, what vocabulary is used? Select from the following:	ATC/ RxNorm/ EphMRA/ ALT/ DrugBank/ Not coded (Free text)/ Other	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	B4.11.2	No	To Keep	Prescriptions not coded	If Not coded (Free text), approximately what percentage of records are not coded?	Unit: %, not applicable	Manual entry	B
Data source - data banks	B4.11.3	No	Not a must	Prescriptions vocabulary, other	If 'other,' what vocabulary is used?	Free text	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data source - data banks	**B4.12**	**Yes**	To Keep	**Dispensings included**	Is information on dispensing of medications included?	No/yes	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	B4.12.1	No	To Keep	Dispensings vocabulary	If yes, what vocabulary is used? Select from the following:	ATC/ RxNorm/ EphMRA/ ALT/ DrugBank/ Not coded (Free text)/ Other	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	B4.12.2	No	Not a must	Dispensings not coded	If Not coded (Free text), approximately what percentage of records are not coded?	Unit: %, not applicable	Manual entry	B
Data source - data banks	B4.12.3	No	Not a must	Dispensings vocabulary, other	If 'other,' what vocabulary is used?	Free text	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	**B4.13**	**Yes**	To Keep	**Product information**	Is product information available? Select all items that apply.	Brand name/ batch number/ formulation/ strength/ package size	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data source - data banks	B4.13.1	No	To Keep	Product information, product vocabulary	If product information is included, what vocabulary is used to describe the product?	ART 57/ IFA GmbH/ EQDM/ SPN/ MTHSPL/ Not coded (Free text)/ Other	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	B4.13.2	No	Not a must	Product vocabulary, other	If 'other,' what vocabulary is used?	Free text	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	B4.13.3	No	To Keep	Product information, substance vocabulary	If substance information is included, what vocabulary is used to describe the substance?	ART 57/ IFA GmbH/ EQDM/ SPN/ MTHSPL/ Not coded (Free text)/ Other	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	B4.13.4	No	Not a must	Substance vocabulary, other	If 'other,' what vocabulary is used?	Free text	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	B4.13.5	No	To Keep	Other medicines information	Is other information on medicines or products included?	No/yes	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data source - data banks	B4.13.6	No	Not a must	Other medicines information details	If yes to the above, select all the items that apply.	Dose/ dosage regime/ route of administration/ product information	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	**B4.14**	**Yes**	To Keep	**ATMP**	Is information on advanced therapy medicinal products included?	No/yes	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	**B4.15**	**Yes**	To Keep	**Contraception**	Is information on contraceptives included?	No/yes	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	B4.15.1	No	Not a must	Emergency contraception	If yes, is information on emergency contraceptives included?	No/yes	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	**B4.16**	**Yes**	To Keep	**Indication for use**	If medicines information is captured, is the indication for use recorded?	No/yes/not applicable	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data source - data banks	B4.16.1	No	To Keep	Indication vocabulary	If yes, what vocabulary is used? Select from the following:	ICPC/ ICD9/ ICD10/ ICD1/ Read/ SNOMED/ SNOMED CT/ MedDRA/ OPCS/ CCS/ EDC/ Not coded (Free text)/ Other	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	B4.16.2	No	To Keep	Indication not coded	If Not coded (Free text), approximately what percentage of records are not coded?	Unit: %, not applicable	Manual entry	B
Data source - data banks	B4.16.3	No	To Keep	Indication vocabulary, other	If 'other,' what vocabulary is used?	Free text	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	**B4.17**	**Yes**	To Keep	**Administration vaccines included**	Is information on the administration of vaccines included?	No/yes	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	**B4.17.1**	**Yes**	Not a must	**Administration other injectable medicines**	Is information on the administration of other injectable medicines included?	No/yes	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data source - data banks	**B4.18**	**Yes**	To Keep	**Medical devices**	Is information on medical devices included?	No/yes	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	**B4.19**	**Yes**	To Keep	**Procedures included**	Is information on procedures (e.g., diagnostic tests, therapeutic, surgical) included?	No/yes	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	B4.19.1	No	To Keep	Procedures vocabulary	If yes, what vocabulary is used? Select from the following:	ICPC/ ICD9/ ICD10/ ICD1/ Read/ SNOMED/ SNOMED CT/ MedDRA/ OPCS/ CCS/ EDC/ Not coded (Free text)/ Other	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	B4.19.2	No	To Keep	Procedures not coded	If Not coded (Free text), approximately what percentage of records are not coded?	Unit: %, not applicable	Manual entry	B
Data source - data banks	B4.19.3	No	Not a must	Procedure vocabulary, other	If 'other,' what vocabulary is used?	Free text	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data source - data banks	B4.19.4	No	Not a must	Procedure results	Are results of procedures available?	No/yes	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	B4.19.5	No	Not a must	Procedure results vocabulary	If yes, what vocabulary is used? Select from the following:	ICPC/ ICD9/ ICD10/ ICD1/ Read/ SNOMED/ SNOMED CT/ MedDRA/ OPCS/ CCS/ EDC/ Not coded (Free text)/ Other	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	B4.19.6	No	Not a must	Procedure results not coded	If Not coded (Free text), approximately what percentage of records are not coded?	Unit: %, not applicable	Manual entry	B
Data source - data banks	B4.19.7	No	Not a must	Procedure results vocabulary, other	If 'other,' what vocabulary is used?	Free text	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	**B4.20**	**Yes**	To Keep	**Clinical measures included**	Is information on clinical measures included (e.g., BMI, blood pressure)?	No/yes	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data source - data banks	**B4.21**	**Yes**	Not a must	**Healthcare centre**	Is information on the healthcare centre (e.g., pharmacy, general practice, hospital) available?	No/yes	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	**B4.21.1**	**Yes**	To Keep	**Healthcare provider**	Is information on the person providing healthcare (e.g., physician, pharmacist, specialist) available?	No/yes	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	**B4.22**	**Yes**	To Keep	**Genetic data**	Is genotyping information available?	No/yes	Manual entry or automatically populated when data banks are entered in B4.1 with values from	B
Data source - data banks	B4.22.1	No	To Keep	Genetic data vocabulary	If yes, what vocabulary is used? Select from the following:	OGG/ FG/ GO/ EGO/ SOPHARM/ PHARE/ other	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	B4.22.2	No	To Keep	Genetic data vocabulary, other	If 'other,' what vocabulary is used?	Free text	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data source - data banks	B4.22.3	No	Not a must	Genetic data summary	A short summary of how genetic information is captured in the data source.	Free text	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	**B4.23**	**Yes**	To Keep	**Biomarker data**	Is biomarker information available?	No/yes	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	B4.23.1	No	To Keep	Biomarker data vocabulary	If yes, what vocabulary is used? Select from the following:	BMO/ SMASH/ FOBI/ other	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	B4.23.2	No	Not a must	Biomarker data vocabulary, other	If 'other,' what vocabulary is used?	Free text	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	B4.23.3	No	Not a must	Biomarker data summary	A short summary of how biomarker information is captured in the data source.	Free text	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data source - data banks	**B4.24**	**Yes**	To Keep	**Patient-generated data**	Is patient-generated information (e.g., from wearable devices) available?	No/yes	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	B4.24.1	No	To Keep	Patient-generated data summary	A short summary of how patient-generated information (e.g., from wearable devices) is captured in the data source.	Free text	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	**B4.25**	**Yes**	To Keep	**Quality of life included**	Are quality of life measures available? Select all that apply.	AQoL-8D/ QOLS/ MQOL/ MQOL-E/ HRQOL/ WHOQOL/ EQ5D/ 15D/ SF-36/ SF-6D/ HUI/ other	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	**B4.25.1**	**Yes**	To Keep	**Quality of life, other**	If other, specify additional quality of life measures.	Free text	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	**B4.26**	**Yes**	To Keep	**Costs**	Is information on costs of healthcare available?	No/yes	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data source - data banks	B4.27.1	No	Not a must	Use of NLP	Is natural language processing used to process free text information into standardised information?	No/yes/not applicable	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	B4.27.2	No	Not a must	Details of NLP	If natural language processing is used, provide a summary description of how natural language processing is used.	Free text	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	B4.27.3	No	Not a must	Deidentified text	Are procedures in place to de-identify or remove personal information from free text?	No/yes/not applicable	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - data banks	**B4.27**	**Yes**	Not a must	**Free text**	Is free text information available?	No/yes	Manual entry or automatically populated when data banks are entered in B4.1 with values from C6	B
Data source - linkage	B5.1	No	Not a must	Linkage description	Provide a high-level description of the linkages that are currently available between data banks in the data source (max 100 words).	Free text	Manual entry	B
Data source - linkage	**B5.2**	**Yes**	Manda-tory	**Linkage strategy**	The linkage method that was used to link data banks. One entry per data bank.	Deterministic/ probabilistic/ combination/ other	Manual entry	B

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data source - linkage	B5.2.1	Yes	To Keep	Linkage variable	If a single variable (or linkage key) is used to link a data bank to others, a name and description of the variable is provided.	Free text	Manual entry	B, C
Data source - linkage	B5.2.2	Yes	Not a must	Linkage variable unique	If a single variable is used to link a data bank to others, is the variable a unique identifier? One entry per data bank.	No/yes	Manual entry	B, C
Data source - linkage	B5.3	Yes	To Keep	Linkage completeness	Provide a high-level description of the completeness of linkages that are currently available between data banks in the data source (max 100 words).	Free text	Manual entry	B
Data source - quantitative descriptors	B6.1	Yes	To Keep	Population size	Total number of unique individuals with records captured in the data source (most recent count). In the catalogue, this will accommodate counts per year.	Numeric	Manual entry	B
Data source - quantitative descriptors	B6.1.1	Yes	Not a must	Population size active	Number of unique, active, or currently registered individuals with records captured in the data source (most recent count). In the catalogue, this will accommodate counts per year.	Numeric	Manual entry	B
Data source - quantitative descriptors	B6.2	Yes	Not a must	Number of individuals with biological samples	Number of unique individuals with records of biological samples (e.g., blood, urine) (most recent count). In the catalogue, this will accommodate counts per year.	Numeric	Manual entry	B

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data source - quantitative descriptors	**B6.3**	**Yes**	To Keep	**Median observation time**	Median time for which unique individuals with records captured in the data source are observable (most recent count).	Unit years	Manual entry; in future implementations scripts or functions could work together with a common data model to extract this information directly from a data source	B
Data source - quantitative descriptors	**B6.3.1**	**Yes**	To Keep	**Median observation time active**	Median time for which unique, active, or currently registered individuals with records captured in the data source are observable (most recent count).	Unit: years	Manual entry; in future implementations scripts or functions could work together with a common data model to extract this information directly from a data source	B
Data source - quantitative descriptors	B6.4	No	To Keep	Population size by age	Total number of unique individuals with records captured in the data source (most recent count), stratified by age	Numeric, for each of the following categories: newborn infants (0 to 27 days), infants and toddlers (28 days to 23 months), children (2 to 11 years), adolescents (12 to 17 years), adults (18 to 45 years), adults (46 to 64 years), adults (65 to 75 years), adults (76 years and over)	Manual entry; in future implementations scripts or functions could work together with a common data model to extract this information directly from a data source	B

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data source - quantitative descriptors	B6.4.1	No	To Keep	Population size active by age	Number of unique, active, or currently registered individuals with records captured in the data source (most recent count), stratified by age	Numeric, for each of the following categories: newborn infants (0 to 27 days), infants and toddlers (28 days to 23 months), children (2 to 11 years), adolescents (12 to 17 years), adults (18 to 45 years), adults (46 to 64 years), adults (65 to 75 years), adults (76 years and over)	Manual entry; in future implementations scripts or functions could work together with a common data model to extract this information directly from a data source	B
Median age of underlying population	B6.5	No	Not a must	Median age of underlying population	Median age of the underlying population (e.g., in the geographical population that the data source covers, based on external estimates, e.g., census)	Numeric	Manual entry; in future implementations scripts or functions could work together with a common data model to extract this information directly from a data source	B
Median age data source	**B6.5.1**	**Yes**	Not a must	**Median age data source**	Median age of individuals within data source	Numeric	Manual entry; in future implementations scripts or functions could work together with a common data model to extract this information directly from a data source	B

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data source - quantitative descriptors	B6.6	No	Not a must	Proportion female in the underlying population	Proportion of females in the underlying population (based on external estimates, e.g., census)	Unit: %	Manual entry; in future implementations scripts or functions could work together with a common data model to extract this information directly from a data source	B
Data source - quantitative descriptors	**B6.6.1**	**Yes**	Not a must	**Proportion female data source**	Proportion of females in the data source	Unit: %	Manual entry; in future implementations scripts or functions could work together with a common data model to extract this information directly from a data source	B
Data source - quantitative descriptors	B6.7	No	Not a must	Proportion ethnicity in the underlying population	Proportion of people per ethnicity in the underlying population (based on external estimates, e.g., census)	Unit: %, per ethnicity group	Manual entry; in future implementations scripts or functions could work together with a common data model to extract this information directly from a data source	B

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data source - ETL	B6.7.1	Yes	Not a must	**Proportion ethnicity data source**	Proportion of people per ethnicity in the data source	Unit: %, per ethnicity group	Manual entry; in future implementations scripts or functions could work together with a common data model to extract this information directly from a data source	B
Data source - ETL	B7.1	Yes	Not a must	**Data source ETL**	Has one or more data banks of the data source been ETL-d to a common data model?	No/yes		B
Data source - ETL	B7.2	Yes	Not a must	**Data source ETL CDM**	Name of CDM(s) to which the data source has been ETL-d	Should take values within D1.1		B
Data source - ETL	B7.3	No	Not a must	Data source ETL CDM version	Version(s) of CDM(s) to which the data source has been ETL-d	Free text	Manual entry; could automatically take values from F4.3 study-mappings	B
Data source - ETL	B7.4	Yes	Not a must	**Data source ETL specifications**	If available, list of ETL specification documents uploaded to the catalogue. These documents describe the mapping of the data source to the CDM.	URL/uploaded documents	Manual entry; could automatically take values from F4.4 study-mappings	B
Data source - ETL	B7.4.1	No	Not a must	Data source ETL code	If available, the ETL code or scripts uploaded to the catalogue. These are the code or scripts that were used to transform data from the original source to the CDM.	URL/uploaded documents	Manual entry; could automatically take values from F4.4 study-mappings	B

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data source - studies	**B8.1**	**Yes**	To Keep	**Data source studies**	A list of IDs of studies that have been conducted using the data source that are recorded in the catalogue	Should take values in F1.1	Automatically populated when the data source ID is entered in F3.3 - Study - data sources and data banks	B
Data source - studies	B8.2	No	Not a must	Regulatory use	List the types of regulatory studies that the data source has been used for (e.g., PASS, PAES, …).	Should take values in F2.4 - study - protocol	Automatically populated with values from F2.4 - Type of study, when studies are entered, and the data source ID is entered in F3.3.	B
Data source - publications	B9.1	No	To keep	Data source website	If applicable, a dedicated website for the data source	URL	Manual entry	B
Data source - publications	B9.2	No	To Keep	Data source publications	A list of peer-reviewed papers or documents describing the data source (validation, data elements, representativeness) or its use for pharmacoepidemiologic research	Allow individual AMA format citations / URL to list of publications.	Automatically populated when an entry is made in F7.2-Study report, using the study IDs (F1.1) where the data source is selected in F3.3- Data sources and data banks	B
Data source - publications	B9.3	No	Not a must	Data source, other information	A list of URLs to other relevant resources describing the data source or its use in pharmacoepidemiologic studies	URL	Manual entry	B
Data bank - underlying population	**C1.1**	**Yes**	To Keep	**Data bank ID**	A persistent identifier for a data bank. This may be created by the catalogue maintainer on first entry of a data bank. It provides links across tables.			C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - underlying population	C1.2	**Yes**	Manda-tory	**Data bank name**	Name of data bank used in European projects	Free text	Manual entry	C
Data bank - underlying population	C1.2.1	**Yes**	Not a must	**Data bank local name**	If different from above, name of data bank in the national language of the data source	Free text	Manual entry	C
Data bank - underlying population	C1.3	**Yes**	To Keep	**Data bank acronym**	If applicable	Free text	Manual entry	C
Data bank - underlying population	C1.4	**Yes**	Manda-tory	**Description of underlying population**	Provide a summary description of the underlying population (maximum 100 words) or URL to a description	Free text/URL	Manual entry	C
Data bank - underlying population	C1.5	**Yes**	Manda-tory	**Data bank countries**	Countries where data originate (1 or more)	Select from list of countries, links to ISO 3166-1 country codes (https://www.iso.org/obp/ui/#search/code/)	Manual entry	C
Data bank - underlying population	C1.5.1	**Yes**	Manda-tory	**Data bank regions**	Geographical regions that data bank covers	Select from list of regions, links to ISO 3166-2 region codes (https://www.iso.org/obp/ui/#search/code/)	Manual entry	C
Data bank - underlying population	C1.6	**Yes**	To Keep	**<u>Causes</u> Event triggering of population entry**	Select the possible causes/ event for a person to enter the underlying population of the data bank.	Birth/ immigration/ residency obtained/ insurance coverage start/ disease diagnosis/ treatment started/ practice registration/ other	Manual entry	C
Data bank - underlying population	C1.6.1	**Yes**	To Keep	**<u>Causes</u> Event triggering of population entry other**	If other, specify the causes /event determining population exit of entry to the underlying population.	Free text	Manual entry	C
Data bank - underlying population	C1.7	**Yes**	Manda-tory	**Causes of population exit**	Select the possible causes / /event determining population exit for a person to leave the underlying population of the data bank.	Death/ emigration/ insurance coverage end/ practice deregistration/ loss to follow up/ other / End of treatment	Manual entry	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - underlying population	C1.7.1	Yes	To Keep	<u>Causes</u> Event triggering of population exit, other	If other, specify the causes / /event determining population exit of exit from the underlying population.	Free text	Manual entry	C
Data bank - underlying population	C1.8	Yes	To Keep	Population age groups	Which population age groups are captured in the data bank? Select all that are relevant.	Numeric, for each of the following categories: newborn infants (0 to 27 days), infants and toddlers (28 days to 23 months), children (2 to 11 years), adolescents (12 to 17 years), adults (18 to 45 years), adults (46 to 64 years), adults (65 to 75 years), adults (76 years and over)	Manual entry	C
Data bank - underlying population	C1.9	Yes	To Keep	Population pregnancy and neon	Does the data bank collect information on pregnant women and/or neonatal subpopulation?	No/yes	Manual entry	C
Data bank - underlying population	C1.10	Yes	To Keep	Disease subpopulations	Does the data bank collect information on a specific disease subpopulation c)?	No/yes	Manual entry	C
Data bank - underlying population	C1.10.1	Yes	To Keep	Disease subpopulations details	If yes, which disease subpopulations does the data bank collect information on?	Free text	Manual entry	C
Data bank - underlying population	C1.11	Yes	To Keep	Population coverage	Are there any people from the underlying population that the data bank covers whose data are not collected in the data bank (e.g., people without a national insurance number, people who are registered for private care)?	No/yes	Manual entry	C
Data bank - underlying population	C1.11.1	Yes	Manda-tory	Population coverage details	If yes to the question above, which subpopulation(s) are not covered by the data bank?	Free text	Manual entry	C
Data bank - underlying population	C1.11.2	Yes	Manda-tory	Population coverage, percentage	What percentage of the underlying population is covered by the data bank?	Unit: %	Manual entry	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - underlying population	C1.12	**Yes**	Manda-tory	**Time span - first collection**	On what date did records first start entering the data bank?	DDMMYYYY	Manual entry	C
Data bank - underlying population	C1.13	**Yes**	Manda-tory	**Time span - last collection**	If data collection in the data bank has ceased, on what date did new records last enter the data bank?	DDMMYYYY	Manual entry	C
Data bank - underlying population	C1.14	**Yes**	Manda-tory	**Care setting**	Type of care setting(s) covered by data bank.	General population/ primary care/ secondary care/ tertiary care/ other	Manual entry	C
Data bank - underlying population	C1.14.1	**Yes**	Not a must	**Care setting, other**	If 'other,' description of the setting of care.	Free text	Manual entry	C
Data bank - access	C2.1	**Yes**	To Keep	**DAPs**	List of DAPs that are listed in the catalogue as having (conditional) permission to access (an extract of) the data bank	Should take values from A1.1	Automatically populated when a contributor enters A2.3 Ins-data sources and data banks-Is DAP as yes and a data bank with corresponding Data bank ID is selected	C
Data bank - access	C2.2	**Yes**	To Keep	**Data controller**	The name of the institution or organisation that is responsible for governance of the data bank.	Free text	Manual entry	C
Data bank - access	C2.3	No	To Keep	Governance details	If available, provide documents or webpages that describe the overall governance of the data bank (governing data access or utilisation for research purposes by existing DAPs)	Link to documents/URL	Manual entry	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - access	C2.3.1	No	Not a must	Committee data access	Is there a committee that determines access to the data source by the DAPs that are already listed?	No/yes	Manual entry	C
Data bank - access	C2.4	No	Not a must	Policy data access	If available, provide documents or webpages that describe the policy for access to the data bank.	Link to documents/URL	Manual entry	C
Data bank - access	**C2.5**	**Yes**	To Keep	**Informed consent**	Is informed consent required for use of the data for research purposes?	Required for general use/ required for all studies/ required for intervention studies/ waiver/ other	Manual entry	C, A
Data bank - access	**C2.5.1**	**Yes**	To Keep	**Informed consent other**	If other, describe the conditions when informed consent is required.	Free text	Manual entry	C, A
Data bank - access	**C2.6**	**Yes**	To Keep	**Access to identifiable data**	Can identifiable data be accessed in the data bank (including patient/practitioner name/practice name)?	Free text	Manual entry	C, A
Data bank - access	C2.6.1	No	To Keep	Route to access to identifiable data	If yes above, what is the route to access this information? What permission is required?	No/yes	Manual entry	C, A
Data bank - access	**C2.7**	**Yes**	To Keep	**Access to subject details**	Can individual patients/practitioners/practices be contacted in the data bank?	No/yes	Manual entry	C, A
Data bank - access	C2.7.1	No	To Keep	Route to access to subject details	If yes above, what is the route to access this information? What permission is required?	Free text	Manual entry	C, A
Data bank - access	C2.8	No	To Keep	Audit possibility	Are external parties allowed to audit the data? For example, is it possible for an external party to audit the quality or validity of the data bank?	No/yes	Manual entry	C, A

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - access	C2.9	No	Manda-tory	Approval for publication	Is the catalogue contributor able to publish results from the data bank without interaction from an external organization, such as approval of a protocol?	No/yes	This is entered manually per institution that makes an entry for the data source in the catalogue	C, A
Data bank - access	**C2.10**	**Yes**	To Keep	**Third-party access**	Can (an extract of) the data bank be accessed with permission by a third party?	No/yes	Manual entry	C
Data bank - access	**C2.10.1**	**Yes**	To Keep	**Third-party access conditions**	If yes above, describe the conditions under which third-party access may be granted (another organisation that is not a data originator or a DAP)	Free text	Manual entry	C
Data bank - access	**C2.11**	**Yes**	Not a must	**Non–EU EEA access**	Can (an extract of) the data bank be accessed with permission by a non-EU/EEA institution?	No/yes	Manual entry	C
Data bank - access	**C2.11.1**	**Yes**	Not a must	**Non–EU EEA access conditions**	If yes above, describe the conditions under which third-party access may be granted	No/yes	Manual entry	C
Data bank - access	C2.12	No	To Keep	Standard operating procedures	Is there a standard operating procedure document that defines the processes and procedures for data capture and management?	No/yes	Manual entry	C
Data bank - access	C2.13	No	To Keep	Biospecimen access	If the data bank contains biospecimens (e.g., tissue samples), can these be retrieved?	No/yes	Manual entry	C
Data bank - access	C2.13.1	No	To Keep	Biospecimen access conditions	If yes above, describe the conditions under which permission to retrieve biospecimens may be granted.	Free text	Manual entry	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - qualification	C3.1	Yes	To Keep	Data bank qualification	Has the data bank successfully undergone a qualification process (e.g., from the EMA)?	No/yes	Manual entry	C
Data bank - qualification	C3.1.1	Yes	To Keep	Description of qualifications	If yes, describe the qualification(s) granted.	Free text	Manual entry	C
Data bank - originator	C4.1	Yes	To Keep	Originator	Name of the organisation that maintains the collection of records.	Free text	Manual entry	C
Data bank - originator	C4.1.1	No	Not a must	Originator email contact	Contact email address of the originator organization	Free text; must take form "prefix@domain"	Manual entry	C
Data bank - originator	C4.2	Yes	Not a must	Type of organisation	Description of the type of organisation that is originator of the data bank.	Academic institution/ civic authority/ government agency/ healthcare payer/ network of primary care practices/ non-profit organisation/ pharmaceutical industry/ private organisation/ public health authority/ research centre/ statistical authority	Manual entry	C
Data bank - originator	C4.2.1	Yes	Not a must	Type of organisation, other	If other, description of the type of originator.	Free text	Manual entry	C
Data bank - originator	C4.3	No	To Keep	Description of data collection	Describe the process of collection and recording of data in the data bank. This could include the tools, such as surveys, used, or a description of the system that the originator uses to gather data and store it the data bank.	Free text	Manual entry	C
Data bank - originator	C4.4	Yes	Not a must	Reason data bank is sustained	Description of the reason why the data bank is sustained by the organisation (e.g., for surveillance, clinical purposes, financial or administrative purposes, research purposes).	Free text	Manual entry	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - originator	**C4.5**	**Yes**	Not a must	**Date established**	Date that the data bank was first established.	DDMMYYYY	Manual entry	C
Data bank - originator	C4.6	No	To Keep	Finance source	Specify the main financial support sources for the data bank in the last 3 years. Select all that apply.	Funding by own institution/ national, regional, or municipal public funding/ European public funding/ funding from industry or contract research/ funding from public-private partnership/ funds from patients organisations, charity and foundations/ other	Manual entry	C
Data bank - originator	C4.6.1	No	Not a must	Finance source, other	Specify	Free text	Manual entry	C
Data bank - prompt	**C5.1**	**Yes**	To Keep	**Data bank family**	Which of the following families of data banks best describe this data bank?	Hospital discharge records/ primary care medical records/ pharmacy dispensation records/ birth registry/ induced terminations registry/ congenital anomaly registry/ inhabitant registry/ registration with healthcare system/ outpatient visit records/ emergency care discharge records/ exemptions from copayment/ diagnostic tests or procedures reimbursement/ death registry/ cancer registry/ other disease registry/ vaccine registry/ biobank/ administrative claims / spontaneous reporting of adverse drug reactions	Manual entry	C
Data bank - prompt	C5.1.1	**Yes**	To Keep	**Data bank family, other**	If 'other,' describe the type of the data bank.	Free text	Manual entry	C
Data bank - prompt	**C5.2**	**Yes**	To Keep	**Data bank prompt**	What prompts the creation of a record in the data bank? e.g., hospital discharge, specialist encounter, dispensation of a medicinal product, recording of a congenital anomaly	Free text	Manual entry	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - prompt	C5.3	No	To Keep	Data bank prompt exception	Describe any exceptions when the data bank prompt would not lead to creation of a record.	Free text	Manual entry	C
Data bank - prompt	**C5.4**	**Yes**	Not a must	**Unit of observation**	Based on the prompt, what is the unit of observation of a record (e.g., person, prescription)?	Free text	Manual entry	C
Data bank - prompt	**C5.5**	**Yes**	Not a must	**Multiple entries**	Can there be multiple entries for a single person in the data bank? For example, may a person contribute multiple records to the data bank?	No/yes	Manual entry	C
Data bank - data model	C6.1	Yes	Not a must	**Data dictionary present**	Are a data dictionary and a data model available?	No/yes	Manual entry	C
Data bank - data model	**C6.1.1**	**Yes**	Not a must	**Data dictionary document**	If available, a list of data dictionary and data model documents is entered. The data dictionary contains the list of variables captured in the data bank, according to the data model, with descriptions.	Links to documents	Manual entry	C
Data bank - data model	C6.1.2	No	Not a must	Data bank number of tables	Number of separate tables within the data bank	Numeric	Manual entry	C
Data bank - data model	C6.1.3	No	Not a must	Data dictionary item names	Data dictionary: inclusion of data item names	Yes, for all variables; Yes, but not for all variables; No	Manual entry	C
Data bank - data model	C6.1.4	No	Not a must	Data dictionary values	Data dictionary: inclusion of codes, value ranges, & measurement units	Yes, for all variables; Yes, but not for all variables; No	Manual entry	C
Data bank - data model	C6.1.5	No	Not a must	Data dictionary algorithms	Data dictionary: inclusion of formulae & algorithms used for derived variables	Yes, for all derived variables; Yes, but not for all derived variables; No; Not required as source does not contain derived variables	Manual entry	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - data model	C6.1.6	No	Not a must	Data dictionary language	Languages in which that the data bank dictionary is recorded	List of languages linked to ISO 639 codes (https://www.iso.org/iso-639-language-codes.html)	Manual entry	C
Data bank - data model	C6.1.7	No	Not a must	Data dictionary summary	If available, a summary data dictionary or abridged version is entered	Links to documents	Manual entry	C
Data bank - data model	**C6.2**	**Yes**	To Keep	**Data bank language**	Languages in which that the records in the data bank are recorded	List of languages linked to ISO 639 codes (https://www.iso.org/iso-639-language-codes.html)	Manual entry	C
Data bank - data model	C6.3	No	Not a must	Processed tables	Does the data bank contain pre-processed tables (i.e., curated by an organisation other than the DAP)?	No/yes	Manual entry	C
Data bank - data model	C6.3.1	No	Not a must	Processed tables description	If yes to the above, a description of the tables that are pre-processed	Free text	Manual entry	C
Data bank - data model	**C6.4**	**Yes**	To Keep	**Unique identifier**	Is there a unique identifier for a person in the data bank?	No/yes	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - data model	C6.4.1	Yes	Not a must	Unique identifier description	Describe the variable that is used as a unique identifier for a person in the data bank? If the unique identifier is not at level of a person (for example hospital encounter), describe how this translated to an individual level.	Free text	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C
Data bank - data model	C6.5	Yes	To Keep	Geographic regions	Are data on geographic region available? Can geographical can be identify ?	No/yes	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - data model	**C6.6**	**Yes**	Manda-tory	**Family linkage**	Is any linkage available in the data bank between any of the following persons?	Household/ mother-child/ father-child/ sibling	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C
Data bank - data model	**C6.6.1**	**Yes**	Manda-tory	**Family linkage routine**	Is familial linkage created routinely or can linkage be performed on an ad-hoc basis?	Routine/ad hoc	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - data model	C6.6.2	No	Not a must	Family linkage availability	Is familial linkage available for all of the data bank population or a subset of individuals (or records)?	Full/subset	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C
Data bank - data model	C6.6.3	No	Not a must	Family linkage availability details	If a familial linkage is available for a subset of individuals (or records) only, provide a short description of the subset.	Free text	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - data model	**C6.7**	**Yes**	To Keep	**Sociodemographics included**	Which of the following sociodemographic information is included? Select all that apply.	None/ age/ sex/ gender/ ethnicity/ country of origin/ socioeconomic status/ marital status/ education level/ type of residency/ living in rural area/ health area/ deprivation index/ pharmaceutical copayment/ other	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C
Data bank - data model	**C6.8**	**Yes**	To Keep	**Lifestyle factors included**	Which of the following lifestyle information is included? Select all that apply.	None/ tobacco use/ alcohol use/ frequency of exercise/ other	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - data model	**C6.9**	**Yes**	Manda-tory	**Diagnosis codes included**	Are diagnostic codes included?	No/yes	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C
Data bank - data model	C6.9.1	No	Manda-tory	Diagnosis medical event code vocabulary	If yes, what vocabulary is used? Select from the following:	ICPC/ ICD9/ ICD10/ ICD1/ Read/ SNOMED/ SNOMED CT/ MedDRA/ OPCS/ CCS/ EDC/ Not coded (Free text)/ Other	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C
Data bank - data model	C6.9.2	No	To Keep	Diagnosis code not coded	If Not coded (Free text), approximately what percentage of records are not coded?	Unit: %, not applicable	Manual entry	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - data model	C6.9.3	No	Not a must	Diagnosis, other	Is other information on diagnoses of patient outcomes recorded, not in the form of diagnostic codes?	Free text	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C
Data bank - data model	**C6.10**	**Yes**	Manda-tory	**Hospital admission discharge**	Is information on hospital admission and/or hospital discharge available?	No/yes	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - data model	**C6.10.1**	**Yes**	To Keep	**ICU admission**	Is information on intensive care unit admission available?	No/yes	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C
Data bank - data model	**C6.11**	**Yes**	To Keep	**Cause of death**	Is cause of death available?	No/yes	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - data model	C6.11.1	No	To Keep	Cause of death vocabulary	If yes, what vocabulary is used? Select from the following:	ICPC/ ICD9/ ICD10/ ICD1/ Read/ SNOMED/ SNOMED CT/ MedDRA/ OPCS/ CCS/ EDC/ Not coded (Free text)/ Other	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C
Data bank - data model	C6.11.2	No	To Keep	Cause of death not coded	If Not coded (Free text), approximately what percentage of records are not coded?	Unit: %, not applicable	Manual entry	C
Data bank - data model	C6.11.3	No	To Keep	Cause of death code, other	If 'other,' what vocabulary is used?	Free text	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - data model	C6.12	Yes	Optional	Rare diseases	Are rare diseases (<1 in 2000 people) captured?	No/yes	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C
Data bank - data model	C6.13	Yes	To Keep	Prescriptions dispensing included	Is information on prescription of medications included?	No/yes	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - data model	C6.13.1	No	To Keep	Prescriptions dispensing vocabulary	If yes, what vocabulary is used? Select from the following:	ATC/ RxNorm/ EphMRA/ ALT/ DrugBank/ Not coded (Free text)/ Other	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C
Data bank - data model	C6.13.2	No	Optional	Prescriptions not coded	If Not coded (Free text), approximately what percentage of records are not coded?	Unit: %, not applicable	Manual entry	C
Data bank - data model	C6.13.3	No	Not a must	Prescriptions vocabulary, other	If 'other,' what vocabulary is used?	Free text	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - data model	C6.14	Yes	To Keep	**Dispensings included**	Is information on dispensing of medications included?	No/yes	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C
Data bank - data model	C6.14.1	No	Manda-tory	Dispensings vocabulary	If yes, what vocabulary is used? Select from the following:	ATC/ RxNorm/ EphMRA/ ALT/ DrugBank/ Not coded (Free text)/ Other	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C
Data bank - data model	C6.14.2	No	Not a must	Dispensings not coded	If Not coded (Free text), approximately what percentage of records are not coded?	Unit: %, not applicable	Manual entry	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - data model	C6.14.3	No	Not a must	Dispensings vocabulary, other	If 'other,' what vocabulary is used?	Free text	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C
Data bank - data model	C6.15	No	To Keep	Product information	What product information available? Select all items that apply.	Brand name/ batch number/ formulation/ strength/ package size Dose/ dosage regime/ route of administration/ product information	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - data model	C6.15.1	No	Not a must	Product information, product vocabulary	If product information is included, what vocabulary is used to describe the product?	ART 57/ IFA GmbH/ EQDM/ SPN/ MTHSPL/ Not coded (Free text)/ Other	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C
Data bank - data model	C6.15.2	No	Not a must	Product vocabulary, other	If 'other,' what vocabulary is used?	Free text	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - data model	C6.15.3	No	To Keep	Product information, substance vocabulary	If substance information is included, what vocabulary is used to describe the substance?	ART 57/ IFA GmbH/ EQDM/ SPN/ MTHSPL/ Not coded (Free text)/ Other	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C
Data bank - data model	C6.15.4	No	Not a must	Substance vocabulary, other	If 'other,' what vocabulary is used?	Free text ATC/ RxNorm/ EphMRA/ ALT/ DrugBank/ Not coded (Free text)/ Other	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - data model	C6.15.5	No	to keep	Other medicines information	Is other information on medicines or products included?	No/yes	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C
Data bank - data model	C6.15.6	No	Not a must	Other medicines information details	If yes to the above, select all the items that apply.	Dose/ dosage regime/ route of administration/ product information	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - data model	C6.16	No	To Keep	ATMP	Is information on advanced therapy medicinal products included?	No/yes	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C
Data bank - data model	**C6.17**	**Yes**	To Keep	**Contraception**	Is information on contraceptives included?	No/yes	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - data model	C6.17.1	No	Not a must	Emergency contraception	If yes, is information on emergency contraceptives included?	No/yes	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C
Data bank - data model	**C6.18**	**Yes**	To Keep	**Indication for use**	If medicines information is captured, is the indication for use recorded?	No/yes/not applicable	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - data model	C6.18.1	No	To Keep	Indication vocabulary	If yes, what vocabulary is used? Select from the following:	ICPC/ ICD9/ ICD10/ ICD1/ Read/ SNOMED/ SNOMED CT/ MedDRA/ OPCS/ CCS/ EDC/ Not coded (Free text)/ Other	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C
Data bank - data model	C6.18.2	No	Optional	Indications not coded	If Not coded (Free text), approximately what percentage of records are not coded?	Unit: %, not applicable	Manual entry	C
Data bank - data model	C6.18.3	No	Optional	Indication vocabulary, other	If 'other,' what vocabulary is used?	Free text	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - data model	C6.19	Yes	To Keep	Administration vaccines included	Is information on the administration of vaccines included?	No/yes	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C
Data bank - data model	C6.19.1	Yes	To keep	Administration other injectable medicines	Is information on the administration of other injectable medicines included?	No/yes	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - data model	**C6.20**	**Yes**	To Keep	**Medical devices**	Is information on medical devices included?	No/yes	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C
Data bank - data model	**C6.21**	**Yes**	To Keep	**Procedures included**	Is information on procedures (e.g.. diagnostic tests, therapeutic, surgical) included?	No/yes	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C
Data bank - data model	C6.22.1	No	To keep	Procedures not coded	If Not coded (Free text), approximately what percentage of records are not coded?	Unit: %, not applicable	Manual entry	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - data model	C6.22	No	To Keep	Procedures vocabulary	If yes, what vocabulary is used? Select from the following:	ICPC/ ICD9/ ICD10/ ICD1/ Read/ SNOMED/ SNOMED CT/ MedDRA/ OPCS/ CCS/ EDC/ Not coded (Free text)/ Other	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C
Data bank - data model	C6.22.2	No	Not a must	Procedure vocabulary, other	If 'other,' what vocabulary is used?	Free text	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - data model	C6.22.3	No	Not a must	Procedure results	Are results of procedures available?	No/yes	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C
Data bank - data model	C6.22.4	No	Not a must	Procedure results vocabulary	If yes, what vocabulary is used? Select from the following:	ICPC/ ICD9/ ICD10/ ICD1/ Read/ SNOMED/ SNOMED CT/ MedDRA/ OPCS/ CCS/ EDC/ Not coded (Free text)/ Other	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C
Data bank - data model	C6.22.5	No	Not a must	Procedure results not coded	If Not coded (Free text), approximately what percentage of records are not coded?	Unit: %, not applicable	Manual entry	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - data model	C6.22.6	No	Not a must	Procedure results vocabulary, other	If 'other,' what vocabulary is used?	Free text	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C
Data bank - data model	**C6.23**	**Yes**	To Keep	**Clinical measures included**	Is information on clinical measures included (e.g., BMI, blood pressure)?	No/yes	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - data model	C6.24	Yes	Not a must	Healthcare centre	Is information on the healthcare centre (e.g., pharmacy, general practice, hospital) available?	No/yes	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C
Data bank - data model	C6.24.1	Yes	Not a must	Healthcare provider	Is information on the person providing healthcare (e.g., physician, pharmacist, specialist) available?	No/yes	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - data model	**C6.25**	**Yes**	To Keep	**Genetic data**	Is genotyping information available?	No/yes	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C
Data bank - data model	C6.25.1	No	To Keep	Genetic data vocabulary	If yes, what vocabulary is used? Select from the following:	OGG/ FG/ GO/ EGO/ SOPHARM/ PHARE/ other	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - data model	C6.25.2	No	To Keep	Genetic data vocabulary, other	If 'other,' what vocabulary is used?	Free text	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C
Data bank - data model	C6.25.3	No	Not a must	Genetic data summary	A short summary of how genetic information is captured in the data source.	Free text	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - data model	**C6.26**	**Yes**	To Keep	**Biomarker data**	Is biomarker information available?	No/yes	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C
Data bank - data model	C6.26.1	No	To Keep	Biomarker data vocabulary	If yes, what vocabulary is used? Select from the following:	BMO/ SMASH/ FOBI/ other	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - data model	C6.26.2	No	Not a must	Biomarker data vocabulary, other	If 'other,' what vocabulary is used?	Free text	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C
Data bank - data model	C6.26.3	No	Not a must	Biomarker data summary	A short summary of how biomarker information is captured in the data source.	Free text	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - data model	**C6.27**	**Yes**	To Keep	**Patient-generated data**	Is patient-generated information (e.g., from wearable devices) available?	No/yes	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C
Data bank - data model	C6.27.1	No	To Keep	Patient-generated data summary	A short summary of how patient-generated information (e.g., from wearable devices) is captured in the data bank.	Free text	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - data model	**C6.28**	**Yes**	To Keep	**Quality of life included**	Are quality of life measures available? Select all that apply.	AQoL-8D/ QOLS/ MQOL/ MQOL-E/ HRQOL/ WHOQOL/ EQ5D/ 15D/ SF-36/ SF-6D/ HUI/ other	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C
Data bank - data model	**C6.28.1**	**Yes**	To Keep	**Quality of life, other**	If 'other,' specify additional quality of life measures.	Free text	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - data model	C6.29	Yes	To Keep	Costs	Is information on costs of healthcare available?	No/yes	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C
Data bank - data model	C6.30	Yes	Not a must	Free text	Is free text information available?	No/yes	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - data model	C6.30.1	No	Not a must	Use of NLP	Is natural language processing used to process free text information into standardised information?	No/yes/not applicable	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C
Data bank - data model	C6.30.2	No	Not a must	Details of NLP	If natural language processing is used, provide a summary description of how natural language processing is used.	Free text	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - data model	C6.30.3	No	Not a must	Deidentified text	Are procedures in place to de-identify or remove personal information from free text?	No/yes/not applicable	Manual entry; in future implementations of the catalogue, this information could be extracted automatically from the data dictionary, if an English-language version is uploaded, using scripts of functions	C
Data bank - quantitative descriptors	**C7.1**	**Yes**	To Keep	**Population size**	Total number of unique individuals with records captured in the data bank (most recent count). In the catalogue, this will accommodate counts per year.	Numeric	Manual entry; in future implementations scripts or functions could work together with a common data model to extract this information directly from a data bank	C
Data bank - quantitative descriptors	**C7.1.1**	**Yes**	To Keep	**Population size active**	Number of unique, active, or currently registered individuals with records captured in the data source (most recent count). In the catalogue, this will accommodate counts per year.	Numeric	Manual entry; in future implementations scripts or functions could work together with a common data model to extract this information directly from a data bank	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - quantitative descriptors	**C7.2**	**Yes**	Not a must	**Number of individuals with biological samples**	Number of unique individuals with records of biological samples (e.g., blood, urine) (most recent count). In the catalogue, this will accommodate counts per year.	Numeric	Manual entry; in future implementations scripts or functions could work together with a common data model to extract this information directly from a data bank	C
Data bank - quantitative descriptors	C7.3	No	Manda-tory	Population size by age	Total number of unique individuals with records captured in the data bank (most recent count), stratified by age. The time point at which age will be calculated will need to be specified	Numeric, for each of the following categories: newborn infants (0 to 27 days), infants and toddlers (28 days to 23 months), children (2 to 11 years), adolescents (12 to 17 years), adults (18 to 45 years), adults (46 to 64 years), adults (65 to 75 years), adults (76 years and over)	Manual entry; in future implementations scripts or functions could work together with a common data model to extract this information directly from a data bank	C
Data bank - quantitative descriptors	C7.3.1	No	Manda-tory	Population size active by age	Number of unique, active, or currently registered individuals with records captured in the data bank (most recent count), stratified by age	Numeric, for each of the following categories: newborn infants (0 to 27 days), infants and toddlers (28 days to 23 months), children (2 to 11 years), adolescents (12 to 17 years), adults (18 to 45 years), adults (46 to 64 years), adults (65 to 75 years), adults (76 years and over)	Manual entry; in future implementations scripts or functions could work together with a common data model to extract this information directly from a data bank	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - quantitative descriptors	C7.4	Yes	Not a must	Median age data bank	Median age of individuals within the data bank	Numeric	Manual entry; in future implementations scripts or functions could work together with a common data model to extract this information directly from a data bank	C
Data bank - quantitative descriptors	C7.5	Yes	Not a must	Proportion female data bank	Proportion of females in the data bank	Unit: %	Manual entry; in future implementations scripts or functions could work together with a common data model to extract this information directly from a data bank	C
Data bank - quantitative descriptors	C7.6	Yes	Not a must	Proportion ethnicity data bank	Proportion of people per ethnicity in the data bank	Unit: %, per ethnicity group	Manual entry; in future implementations scripts or functions could work together with a common data model to extract this information directly from a data bank	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - quantitative descriptors	**C8.1**	**Yes**	Not a must	**Data bank refresh**	Average number of days between refresh of data bank with new records	Unit: days	Manual entry	C
Data bank - updates and lag time	**C8.2**	**Yes**	To Keep	**Data bank refresh period**	If data are refreshed on fixed dates (e.g., every June and December), when are the refreshes scheduled? Select all that apply from the following:	January/ February/ March/ April/ May/ June/ July/ August/ September/ October/ November/ December	Manual entry	C
Data bank - quantitative descriptors	**C8.3**	**Yes**	To Keep	**Data bank last refresh**	Date of last update/refresh	DDMMYYY	Manual entry	C
Data bank - quantitative descriptors	**C8.4**	**Yes**	To Keep	**Data bank lag**	How long is the lag time after refresh before a record can be extracted? (e.g., a lag time may occur if the originator conducts quality checks)	Unit: days	Manual entry	C
Data bank - quantitative descriptors	C8.5	No	To Keep	Data bank preservation	Are records preserved in the data bank indefinitely?	No/yes	Manual entry	C
Data bank - quantitative descriptors	C8.5.1	No	To Keep	Data bank preservation length	If no to the above, for how long are records preserved in the data bank?	Unit: years	Manual entry	C
Data bank - quantitative descriptors	C8.6	No	Not a must	Data bank changes	Provide a description of any major changes to the data bank that have occurred over time, including updates in its data model, prompts, or underlying population.	Free text	Manual entry	C
Data bank - quality	C9.1	Yes	Not a must	**Number of records**	Total number of unique records captured in the data bank (most recent count)	Numeric	Manual entry	C
Data bank - quality	C9.2	No	Not a must	Number of excluded records	If records were excluded following quality checks or a quality assurance process, the total number of unique records that were excluded	Numeric	Manual entry	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - quality	C9.3	Yes	Optional	**Comment completeness missing values**	Describe the completeness of the data bank (e.g., variables with more or fewer missing values).	Free text	Manual entry	C
Data bank - quality	C9.3.1	Yes	Not a must	**Comment completeness missing values over time**	Describe any changes in completeness of the data bank (e.g., variables with more or fewer missing values) that have occurred *over time*.	Free text	Manual entry	C
Data bank - quality	C9.3.2	Yes	Not a must	**Comment completeness missing values subpopulations**	Describe any differences in completeness of the data bank (e.g., variables with more or fewer missing values) across subpopulations of people in the data bank.	Free text	Manual entry	C
Data bank - quality	C9.4	Yes	To Keep	**Comment quality**	Describe the quality of the data bank (e.g., variables with more or fewer missing values).	Free text	Manual entry	C
Data bank - quality	C9.4.1	Yes	Not a must	**Comment quality over time**	Describe any changes in quality of the data bank (e.g., variables with more or fewer missing values) that have occurred *over time*.	Free text	Manual entry	C
Data bank - quality	C9.4.2	Yes	Not a must	**Comment quality subpopulations**	Describe any differences in quality of the data bank (e.g., variables with more or fewer missing values) across subpopulations of people in the data bank.	Free text	Manual entry	C
Data bank - quality	C9.5	Yes	Optional	**Access validation**	Can validity of the data in the data bank be verified, e.g., by review of origin medical charts?	No/yes	Manual entry	C
Data bank - quality	C9.5.1	No	Optional	Data quality check and validation frequency	How often are data quality checks and validation steps conducted on the data bank?	Free text	Manual entry	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - quality	C9.5.2	No	Optional	Data quality check and validation methods	What methods or processes are applied for data quality checks and validation steps conducted on the data bank?	Free text	Manual entry	C
Data bank - quality	C9.5.3	No	Optional	Correction of illogical values	What methods or processes are applied to correct illogical values in the data bank?	Free text	Manual entry	C
Data bank - quality	C9.5.4	No	Optional	Data bank completeness results	What methods or processes are applied to check completeness of the data bank?	Free text	Manual entry	C
Data bank - quality	C9.5.5	Yes	Not a must	**Data quality check and validation results**	A table of summary measures of percentage completeness of variables in the data bank (per year) can be uploaded. The table variables should match the variables in the data dictionaries of the data bank. Upload in a standard form is envisioned, to facilitate visualisation.	Upload	Manual entry	C
Data bank - quality	C9.5.6	No	Optional	Data quality check and validation results publication	If available, provide a link to a publication of the data quality check and validation results.	Allow individual AMA format citations / URL to list of publications.	Manual entry	C
Data bank - underlying population	**C10.1**	**Yes**	To Keep	**Data bank studies**	A list of IDs of studies that have been conducted using the data bank that are recorded in the catalogue	Takes values from F1.1	Automatically populated when the data bank ID is entered in F3.4 - Study - data sources and data banks	C
Data bank - underlying population	C11.1	No	To Keep	Data bank website	If applicable, a dedicated website for the data bank	URL	Manual entry	C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Data bank - underlying population	C11.2	No	To Keep	Data bank publications	A list of peer-reviewed and other (as per data source) papers describing the data bank or link to a webpage with this list A list of URLs to other relevant resources describing the data bank or its use in pharmacoepidemiologic studies	Allow individual AMA format citations / URL to list of publications.	Automatically populated when an entry is made in F7.2-Study report, using the study IDs (F1.1) where the data source is selected in F3.3-Data sources and data banks	C
Data bank - underlying population	C11.3	No	Not a must	Data bank, other information	A list of URLs to other relevant resources describing the data bank or its use in pharmacoepidemiologic studies	URL	Manual entry	C
Common data model - general	**D1.1**	**Yes**	To Keep	**CDM ID**	A persistent identifier for a CDM. This may be created by the catalogue maintainer on first entry of a CDM. It provides links across tables.			D
Common data model - general	**D1.2**	**Yes**	To Keep	**CDM name**	Name by which the common data model is most frequently referred. Multiple from the list may be selected.	OMOP/ ConcepTION/ Nordic/ Sentinel/ PCORnet/ VSD/ i2b2/ CDISC SDTM/ PEDSnet/ other	Manual entry	D
Common data model - general	**D1.2.1**	**Yes**	To Keep	**CDM name, other**	If 'other,' provide the name of the CDM.	Free text	Manual entry	D
Common data model - general	D1.3	No	Not a must	CDM maintainer	Name of the organisation or institute that is responsible for developing and maintaining the CDM.	Free text	Manual entry	D
Common data model - general	**D1.4**	**Yes**	To Keep	**CDM website**	If there is a website for the CDM, provide it here.	URL	Manual entry	D

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Common data model - general	D1.4.1	No	Not a must	CDM model	If a website is not available for the CDM, upload documentation of the CDM model.	Document uploaded	Manual entry	D
Common data model - general	D1.5	No	Not a must	CDM tables	If there is a website for the CDM, provide a link to a description of the CDM tables.	URL	Manual entry	D
Common data model - general	D1.6	No	Not a must	CDM documentation	Provide any further resources describing the CDM and tools or software to utilise it.	URL	Manual entry	D
Common data model - vocabulary	**D2.1**	**Yes**	To Keep	**Standardised vocabularies**	Are data mapped to standardised vocabularies during ETL to the CDM?	No/yes	Manual entry	D
Common data model - vocabulary	D2.2	No	To Keep	CDM vocabulary events	If yes, what vocabularies are used for events, such as diagnoses?	ICPC/ ICD9/ ICD10/ ICD1/ Read/ SNOMED/ SNOMED CT/ MedDRA/ OPCS/ Not coded (Free text)/ Other	Manual entry	D
Common data model - vocabulary	D2.2.1	No	To Keep	CDM vocabulary events, other	If 'other,' what vocabulary is used?	Free text	Manual entry	D
Common data model - vocabulary	D2.3	No	To Keep	CDM vocabulary medicines	If yes, what vocabularies are used for medicines or vaccines?	ATC/ RxNorm/ ICPC/ ICD9/ ICD10/ ICD11/ Read/ SNOMED/ SNOMED CT/ MedDRA/ OPCS/ ART 57/ Not coded (Free text)/ Other	Manual entry	D
Common data model - vocabulary	D2.3.1	No	To Keep	CDM vocabulary medicines, other	If 'other,' what vocabulary is used?	Free text	Manual entry	D
Common data model - vocabulary	D2.4	No	To Keep	CDM vocabulary procedures	If yes, what vocabularies are used for procedures?	ICPC/ ICD9/ ICD10/ ICD1/ Read/ SNOMED/ SNOMED CT/ MedDRA/ OPCS/ Not coded (Free text)/ Other	Manual entry	D
Common data model - vocabulary	D2.4.1	No	To Keep	CDM vocabulary procedures, other	If 'other,' what vocabulary is used?	Free text	Manual entry	D

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Network - overview	E1.1	Yes	To Keep	Network ID	A persistent identifier for a network. This may be created by the catalogue maintainer on first entry of a network. It provides links across tables.			E
Network - overview	E1.2	Yes	To Keep	Network name	Official name of the network as used in EU projects	Free text	Manual entry	E
Network - overview	E1.3	Yes	Not a must	Network acronym	Official acronym of the network, if applicable	Free text	Manual entry	E
Network - overview	E1.4	No	Manda-tory	Network website	A link to a dedicated website for the network	URL	Manual entry	E
Network - overview	E1.5	No	To Keep	Network contact person	Name of contact for network	Free text	Manual entry	E
Network - overview	E1.6	No	To Keep	Network contact email	Contact email address for network contact	Free text	Manual entry	E
Network - overview	E1.7	Yes	To Keep	Network institutions	List of institutions that are within the network. Select from the following list of institutions.	Should take values within A1.1. If required value is not present in A1.1, an entry must be made in table A1-Institution-role	Manual entry	E
Network - overview	E1.8	Yes	To Keep	Network studies	List of studies that were conducted within the network.	Should take values within F1.1. If required value is not present in F1.1, an entry must be made in table F1-Study-institutions	Automated entry, populated with the corresponding value of F1.1 when an entry is made in F1.8.	E
Study - institutions	F1.1	Yes	To Keep	Study ID	A persistent identifier for a study. This may be created by the catalogue maintainer on first entry of a study. It provides links across tables.			F

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Study - institutions	**F1.2**	**Yes**	To Keep	**Study title (study acronym)**	Unique study title	Free text	Manual entry. If a persistent Study ID is maintained between this catalogue and the EUPAS register, this may be automatically populated on creation of an entry to the EUPAS register	F
Study - institutions	**F1.3**	**Yes**	To Keep	**Study institution**	Name of lead institution that conducted the study	Should take value within A1.1.	Manual entry. If a persistent Study ID is maintained between this catalogue and the EUPAS register, this may be automatically populated on creation of an entry to the EUPAS register	F
Study - institutions	**F1.4**	**Yes**	To Keep	**Study institution contact name**	Name of institution contact for study	Free text	Manual entry. If a persistent Study ID is maintained between this catalogue and the EUPAS register, this may be automatically populated on creation of an entry to the EUPAS register	F

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Study - institutions	**F1.5**	**Yes**	To Keep	**Study institution contact email**	Email of institution contact for study	Free text; must take form "prefix@domain"	Manual entry. If a persistent Study ID is maintained between this catalogue and the EUPAS register, this may be automatically populated on creation of an entry to the EUPAS register	F
Study - institutions	F1.6	No	To Keep	Study institution contact ORCID	If available, an ORCID number of the investigator named above	Numeric	Manual entry. If a persistent Study ID is maintained between this catalogue and the EUPAS register, this may be automatically populated on creation of an entry to the EUPAS register	F
Study - institutions	**F1.7**	**Yes**	To Keep	**Additional institutions**	List the names of any additional institutions that contributed to the study	Should take value within A1.1.	Manual entry. If a persistent Study ID is maintained between this catalogue and the EUPAS register, this may be automatically populated on creation of an entry to the EUPAS register	F

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Study - institutions	F1.8	No	To Keep	Study network	If relevant, name of the network associated with the study	Should take value within E1.1.	Manual entry. If a persistent Study ID is maintained between this catalogue and the EUPAS register, this may be automatically populated on creation of an entry to the EUPAS register	F
Study - protocols	**F2.1**	**Yes**	To Keep	**Protocol link**	A link to the latest version of the protocol, if published	URL	Manual entry	F
Study - protocols	**F2.2**	**Yes**	To Keep	**EU PAS number**	If registered to EU PAS registry, the EU PAS number of the study	Numeric	Manual entry	F
Study - protocols	**F2.3**	**Yes**	To Keep	**Protocol document**	An uploaded version of the protocol, if a URL or EU PAS entry is not available	Uploaded document	Manual entry	F
Study - protocols	**F2.4**	**Yes**	To Keep	**Type of study**	Select 1 of the following types of study	PASS/ PAES/ external control/ post-marketing surveillance/ other	Manual entry. If a persistent Study ID is maintained between this catalogue and the EUPAS register, this may be automatically populated on creation of an entry to the EUPAS register	F

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Study - protocols	F2.5	No	To Keep	Population	A short description of the study population	Free text	Manual entry. If a persistent Study ID is maintained between this catalogue and the EUPAS register, this may be automatically populated on creation of an entry to the EUPAS register	F
Study - protocols	F2.6	No	To Keep	Interventions	A short description of the study interventions	Free text	Manual entry. If a persistent Study ID is maintained between this catalogue and the EUPAS register, this may be automatically populated on creation of an entry to the EUPAS register	F
Study - protocols	F2.7	No	To Keep	Comparators	A short description of the study comparators	Free text	Manual entry. If a persistent Study ID is maintained between this catalogue and the EUPAS register, this may be automatically populated on creation of an entry to the EUPAS register	F

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Study - protocols	F2.8	No	To Keep	Outcomes	A short description of the study outcomes	Free text	Manual entry. If a persistent Study ID is maintained between this catalogue and the EUPAS register, this may be automatically populated on creation of an entry to the EUPAS register	F
Study - protocols	F2.9	No	To Keep	Timing start	The date of the start of the study period. For example, the date when patients are first included in the study.	DDMMYYY	Manual entry. If a persistent Study ID is maintained between this catalogue and the EUPAS register, this may be automatically populated on creation of an entry to the EUPAS register	F
Study - protocols	F2.10	No	To Keep	Timing end	The date of the end of the study period. For example, the last date of patient follow-up.	DDMMYYY	Manual entry. If a persistent Study ID is maintained between this catalogue and the EUPAS register, this may be automatically populated on creation of an entry to the EUPAS register	F

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Study - protocols	F2.11	No	To Keep	Setting	A short description of the study setting	Free text	Manual entry. If a persistent Study ID is maintained between this catalogue and the EUPAS register, this may be automatically populated on creation of an entry to the EUPAS register	F
Study - data sources and data banks	F3.1	No	To Keep	Number of data sources	Total number of data sources included in the study	Numeric	Manual entry	F
Study - data sources and data banks	F3.2	No	To Keep	Number of data banks	Total number of data banks included in the study	Numeric	Manual entry	F
Study - data sources and data banks	**F3.3**	**Yes**	To Keep	**Data source name**	Names of data sources used in the study. Select from the following list.	Should take values in B1.1	Manual entry	F
Study - data sources and data banks	**F3.4**	**Yes**	To Keep	**Data bank name**	Names of data banks used in the study. Select from the following list.	Should take values in C1.1	Manual entry	F
Study - data sources and data banks	**F3.5**	**Yes**	To Keep	**Data extraction date**	Date on which the study data was extracted	DDMMYYYY	Manual entry	F,C
Study - data sources and data banks	F3.6	No	To Keep	Institution performing extraction	Names of the institutes or organisations that extracted raw data for the study. Select from the following list	Should take values in A1.1	Manual entry	F,C
Study - data sources and data banks	F3.7	No	Not a must	Software data extraction	Name of software used for data extraction (i.e., retrieval of records from data banks)	Free text	Manual entry	F,C

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Study - data sources and data banks	F3.7.1	No	Not a must	Software data extraction version	Version of software used for data extraction (i.e., retrieval of records from data banks)	Free text	Manual entry	F,C
Study - data sources and data banks	F3.7.2	No	Not a must	Restrictions in the data bank extraction	Description of the limitations imposed to the records or values extracted from each data bank (e.g., only records of persons born after 1950; only diaganostic codes belonging to a specified codelist)	Free text	Manual entry	F,C
Study - data sources and data banks	F3.8	No	Not a must	Software data processing	Name of software used for data processing (i.e., creation of the analysis data set(s))	Free text	Manual entry	F,B
Study - data sources and data banks	F3.8.1	No	Not a must	Software data processing version	Version of software used for data processing (i.e., creation of the analysis data set(s))	Free text	Manual entry	F,B
Study - data sources and data banks	F3.8.2	No	Not a must	Characteristics data processing tailored to data source	Study variables that need to be defined with a specific procedure (e.g., using data banks that are specific to this data source)	Free text	Manual entry	F,B
Study - data sources and data banks	F3.9	No	Not a must	Software data analysis	Name of software used for data analysis	Free text	Manual entry	F,B
Study - data sources and data banks	F3.9.1	No	Not a must	Software data analysis version	Version of software used for data analysis	Free text	Manual entry	F,B
Study - data sources and data banks	F3.10	No	Not a must	Data source strengths and limitations for this study	Description *a priori* of the strengths and limitations of this data source to address this research question	Free text	Manual entry	F,B

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Study - data sources and data banks	F3.10.1	No	Not a must	Data characterisation results of this data source	Provide results of the data characterisation or quality checks, such as Sentinel Common Data Model level 1-4 checks. These can be provided in a single document or URL. Upload in a standard form is envisioned, to facilitate visualisation.	as in F5.8	Automatic copy from F5.8	F,B
Study - data sources and data banks	F3.10.2	No	Not a must	Fitness-for-pupose of the data source	Description of the strengths and limitations of this data source to address this research question *a posteriori*: based on the results of the quality checks	Free text	Manual entry	F,B
Study - data sources and data banks	F3.11	No	To Keep	Study results of the data source	Results for this data source	as in F6	Automatic copy from F6	F,B
Study - data sources and data banks	F3.11.1	No	Not a must	Comments on the study results of the data source	Comments on the results for this data source	Free text	Manual entry	F,B
Study - data sources and data banks	F3.12	Not applicable	Not applicable	List of key study variables of this study	List of the names of the key variables (outcome, exposure, selection criteria, covariates, ...) derived from the data source in this study	Free text	Manual entry	F,B
Study - data sources and data banks	F3.12.1	Not applicable	Not applicable	Operationalisation of the study variable	How the study variable was derived from this data source	Free text	Manual entry	F,B, study variable
Study - data sources and data banks	F3.12.2	Not applicable	Not applicable	Validity of the study variable	Comments on variable validity, for instance if validation studies were conducted	Free text	Manual entry	F,B, study variable

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Study - data sources and data banks	F3.12.3	Not applicable	Not applicable	Comments to support interpretation of the study variable	Comments on the interpretation of the variable (for instance, if an outcome is expected only to be captured in severe cases, or if the reimbursement of a medicine is restricted to some age categories, ...)	Free text	Manual entry	F,B, study variable
Study - mappings to CDM	**F4.1**	**Yes**	To Keep	**CDM mapping**	Were data sources/banks in the study (ETL-d) to a CDM?	No/yes	Manual entry	F,B
Study - mappings to CDM	**F4.2**	**Yes**	To Keep	**CDM name**	If yes, select the name of the CDM from the following list.	Should take values in D1.1	Manual entry	F,B
Study - mappings to CDM	**F4.3**	**Yes**	To Keep	**CDM mapping version**	The version of the CDM used	Free text	Manual entry	F,B
Study - mappings to CDM	**F4.4**	**Yes**	Not a must	**CDM ETL**	If available, list of ETL specification documents uploaded to the catalogue. These documents should describe the mapping of the data source to the CDM.	Link to document	Manual entry	F,B
Study - mappings to CDM	F4.4.1	No	Not a must	Data source ETL code	If available, the ETL code or scripts uploaded to the catalogue. These are the code or scripts that were used to transform data from the original source to the CDM.	URL/uploaded documents	Manual entry	F,B
Study - data characterisation	**F5.1**	**Yes**	To Keep	**Data characterisation conducted**	Was a data characterisation or quality check process completed?	No/yes	Manual entry	F,B
Study - data characterisation	**F5.2**	**Yes**	To Keep	**Data characterisation moment**	At what stages of the study were data characterisation steps or quality checks implemented? Select all that apply.	After data extraction/ after extract-transform-load to a common data model/ after creation of study variables	Manual entry	F,B

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Study - data characterisation	F5.3	No	To Keep	Check conformance	Was a check of the conformance of data (i.e., data are in the correct format/syntax) completed?	No/yes	Manual entry	F,B
Study - data characterisation	F5.4	No	To Keep	Check completeness	Was a check of the completeness of data completed?	No/yes	Manual entry	F,B
Study - data characterisation	F5.5	No	To Keep	Check stability	Was a check of the stability of data (e.g., codes) over time completed?	No/yes	Manual entry	F,B
Study - data characterisation	F5.6	No	To Keep	Check logical consistency	Was a check of the logical consistency of data completed?	No/yes	Manual entry	F,B
Study - data characterisation	F5.7	No	To Keep	Data characterisation details	Provide a summary description of the data characterisation or quality check process.	Free text/URL	Manual entry	F,B
Study - data characterisation	**F5.8**	**Yes**	To Keep	**Data characterisation results**	Provide results of the data characterisation or quality checks, such as Sentinel Common Data Model level 1-4 checks. These can be provided in a single document or URL. Upload in a standard form is envisioned, to facilitate visualisation.	Uploaded/URL	Manual entry	F,B
Study - results	F6.1	No	To Keep	Procedure data extraction	Upload or provide a link (e.g., GitHub) to the procedures, such as codes or scripts, used to extract data to create the data source instance used in the study.	Uploaded/URL	Manual entry	F,B
Study - results	F6.2	No	To Keep	Procedure results generation	Upload or provide a link (e.g., Github) to the procedures, such as study scripts, used to generate the study results	Uploaded/URL	Manual entry	F,B

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Study - results	F6.3	No	To Keep	Summary results	A brief summary of the results of the study on study completion (from abstract)	Free text	Manual entry	F,B
Study - results	F6.4	No	To Keep	Results tables	Results tables from the study uploaded to the catalogue. Upload in a standard form is envisioned, to facilitate visualisation.	Uploaded	Manual entry	F,B
Study - publications	**F7.1**	**Yes**	To Keep	**Study publications**	Peer-reviewed papers reporting the study	AMA citation format	Manual entry. If a persistent Study ID is maintained between this catalogue and the EUPAS register, this may be automatically populated on creation of an entry to the EUPAS register	F
Study - publications	**F7.2**	**Yes**	To Keep	**Study report**	A link to the study report, if publicly available online	URL	Manual entry. If a persistent Study ID is maintained between this catalogue and the EUPAS register, this may be automatically populated on creation of an entry to the EUPAS register	F

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Study - publications	F7.3	Yes	To Keep	Study, other information	A list of URLs to other relevant resources describing the study	URL	Manual entry. If a persistent Study ID is maintained between this catalogue and the EUPAS register, this may be automatically populated on creation of an entry to the EUPAS register	F
Study - Data sources	F8	Not applicable	Not applicable	Multiple (see sheet F8)	Multiple (see sheet F8)	Multiple (see sheet F8)	Multiple (see sheet F8)	B,F
Study - Data banks	F9	Not applicable	Not applicable	Multiple (see sheet F9)	Multiple (see sheet F9)	Multiple (see sheet F9)	Multiple (see sheet F9)	C,F
Catalogue - entry	M1.1	Yes	To Keep	Creation date	Date of creation of first entry	DDMMYYYY	Automated entry, only on first entry	M
Catalogue - entry	M1.2	Yes	To Keep	Date of last update	Date on which the entry to the catalogue was made	DDMMYYYY	Automated entry	M
Catalogue - entry	M1.3	Yes	To Keep	Name of contact person	Person responsible for the individual entry	Free text	Manual entry or automated entry using information from user account	M
Catalogue - entry	M1.4	Yes	To Keep	Institution ID	A persistent identifier for an institution. This may be created by the catalogue maintainer on first entry of an institution. It provides links across tables.	Should take values within A1.1	Manual entry or automated entry using information from user account	M

Full metadata list and prioritisation

Table	Var ID	Priority (Deliverable 5)	EMA Revised priority (19 Nov 2021)	Var name	Description	Standards	Entry	Unit of observation (A = institution; B = data source, C = data bank, D = CDM, E = network, F = study)
Catalogue - entry	**M1.5**	**Yes**	To Keep	**Type of contact**	Role within the institution	Administrative/ scientific/ other	Manual entry or automated entry using information from user account	M
Catalogue - entry	**M1.5.1**	**Yes**	To Keep	**Type of contact, other**	If other, free text description of type of contact	Free text	Manual entry or automated entry using information from user account	M
Catalogue - entry	**M1.6**	**Yes**	To Keep	**Email contact**	Email address of person responsible for the entry	Free text; must take form "prefix@domain"	Manual entry or automated entry using information from user account	M
Catalogue - entry	**M1.7**	**Yes**	To Keep	**Tables updated**	List of the data tables that were changed in the update. Select one or more of the following	A1/ A2/ A3/ A4/ B1/ B2/ B3/ B4/ B5/ B6/ B7/ B8/ B9/ C1/ C2/ C3/ C4/ C5/ C6/ C7/ C8/ C9/ C10/ C11/ D1/ D2/ E1/ F1/ F2/ F3/ F4/ F5/ F6/ F7	Manual entry or automated entry depending on catalogue software	M
Catalogue - entry	M1.8	No	Not a must	Reason for update	List of the reasons for the update. Select one or more of the following	new entry/ metadata needed update following changes/ correction of prior entry/ other	Manual entry	M
Catalogue - entry	M1.8.1	No	Not a must	Reason for update other	If other, free text description of the reason for the update	Free text	Manual entry	M
Catalogue - entry	M1.8.2	No	Not a must	Update description	Description of changes made in the update	Free text	Manual entry	M