



Medicines & Healthcare products
Regulatory Agency



Intensive Care National Audit & Research Centre (ICNARC) COVID-19 Dataset

2022-01 CPRD ICNARC Documentation

Version 1.0

Date: 31 January 2022

Documentation Control Sheet

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

Version	Affected Areas Summary of Change	Prepared By	Reviewed By
1.0	Initial	Mia Harley & Mike Lonergan	Susan Hodgson

Summary of Changes

Intensive Care National Audit & Research Centre COVID-19 dataset linked to CPRD primary care data

This document provides an overview of the Intensive Care National Audit & Research Centre (ICNARC) Case Mix Programme (CMP) records for patients identified as COVID-19 cases, and the available subset that is linked to CPRD GOLD and CPRD Aurum.

What is the dataset?

ICNARC¹ is a dataset of adult patients (aged 18 years and over) in adult, general critical care units (intensive care and combined intensive care/high dependency units) that had a laboratory confirmed COVID-19 case in England since 12th December 2019.

What data has ICNARC been linked to?

ICNARC data has been linked to CPRD GOLD and CPRD Aurum primary care databases. Linkage is available for patients who are registered at primary care practices in England and have consented to participate in the linkage scheme. Patients who have opted out are not included in the linked data. Access to linked CPRD ICNARC data is dependant on approval of study protocol, and all required linked data sources must be requested on the application form. These linked data are only provided by CPRD when part of a data extract is linked to CPRD primary care data.

Linkage algorithm and the match rank variable

Linkage between the ICNARC Covid-19 data and CPRD primary care data uses an eight-step deterministic linkage algorithm based on four identifiers, shown in Table 1 below. The linkage is undertaken by NHS Digital, acting as a trusted-third party, on behalf of CPRD. No personal identifiers are held by CPRD, or included in the CPRD GOLD, CPRD Aurum, or linked I data.

Table 1: NHS Digital 8 step linkage algorithm

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

The matching steps are applied sequentially. If a CPRD GOLD or CPRD Aurum patient record is matched in one step, it is no longer available for matching in subsequent steps.

CPRD provides users with a match_rank variable which corresponds to the step at which the match was established. In general, a lower value for the match_rank is considered stronger evidence for a positive match. Note that only patients with a match_rank of 5 or less are considered definitive matches and are included in the linked ICNARC data.

¹ <https://www.icnarc.org/Our-Audit/Audits/Cmp/About>

Linkage method

Linkage was based on identifiers including NHS number, full patient postcode, date of birth, gender and patient pseudonyms. NHS digital linked the ICNARC identifiers to the CPRD identifiers using the set 22 source file. The set 22 source file is the most appropriate to use for understanding the subset of the population potentially eligible for linkage. No personal identifiers are held by CPRD, and none are included in the linked ICNARC data.

Linkage eligibility

The linkage was carried out by NHS Digital using April 2021 EMIS and Vision identifiers, which is also the basis for the set 22 source file (enhanced_eligibility_January 2022). The set 22 source file is the most appropriate to use for understanding the subset of the population potentially eligible for linkage. The ICNARC eligibility flag (icnarc_e=1) in this source file can be used to identify a denominator population potentially eligible for linkage to ICNARC, and indicates that the patient is registered at a practice participating in the linkage scheme, and has a valid NHS Number.

Linkage coverage period

The ICNARC CPRD linkage dataset contains data collected from 12th December 2019 to 18th March 2021 inclusive.

Data structure and formatting

ICNARC data is supplied “as is”, with minimal modification or cleaning during processing by CPRD. Where CPRD has modified the data, this is detailed below.

DOI

Please cite in any publications using these data:

CPRD GOLD ICNARC January 2022 - <https://doi.org/10.48329/9mhe-bm42>

CPRD Aurum ICNARC January 2022 - <https://doi.org/10.48329/vpxm-8q69>

Data quality – known issues

These data are being made available for research purposes on accelerated timelines. Minimal changes have been made to the raw data and there is likely to be some requirement for data cleaning. We therefore encourage our user community to share with us any issues or insights gained from using these data, which can be disseminated to other users via the documentation supporting the next update.

- **Trends:** Because ICNARC data is reported for laboratory confirmed COVID-19 cases only, trends in hospital and critical care admission rates need to be interpreted in the context of testing recommendations, which have changed over time. For example:
 - Since 2 March 2020, all patients in ICU/HDU with influenza like illness (ILI) and/or lower respiratory tract infections (LRTI) and/or pneumonia have been tested for SARS-CoV-2
 - Since 15 March 2020, SARS-CoV-2 testing was extended to all patients hospitalised with ILI, LRTI or pneumonia
 - Since 27 April 2020, trusts were advised to test all non-elective admissions, including asymptomatic patients for SARS-CoV-2.
- **Multiple CPRD GOLD or CPRD Aurum patient records linked to a single ICNARC record (n_patid_icnarc):** In CPRD GOLD and CPRD Aurum, individuals are assigned a new patient identifier (patid) each time they move from one contributing practice to another. A single individual can therefore be represented by two or more different patient identifiers, and each can be linked to the same ICNARC

record (recordid). This information is captured by the variable n_patid_icnarc which indicates how many different patient identifiers are linked to the same specimen number. Data for patients where n_patid_icnarc is large (>20) may not be reliable. Users should consider excluding these patients from analysis.

ICNARC: Data dictionary

File names:

CPRD_GOLD_ICNARC_January_2022.txt; CPRD_Aurum_ICNARC_January_2022.txt

Column name	Description	Type	Format
patid	Encrypted unique key given to a patient in CPRD GOLD or CPRD Aurum [primary key]	TEXT	6-19 numeric characters
pracid	Encrypted unique key given to a practice in CPRD GOLD or CPRD Aurum	INTEGER	5
n_patid_icnarc	Generated by CPRD. Indicates the number of distinct patids associated with the same recordid	INTEGER	3
match_rank ²	Indicates the quality of matching between a record in ICNARC and the CPRD primary care data and gives the level of confidence that a record has been correctly matched to a patient in CPRD GOLD or CPRD Aurum.	INTEGER	1
recordid	Unique record identifier (anonymised, project-specific)	INTEGER	5
unitid	Unique critical care unit identifier (anonymised, project-specific)	INTEGER	3
acsd	Advanced cardiovascular support days	INTEGER	3
aids_v3	HIV/AIDS	STRING	1
amlallmm	Acute myelogenous/lymphocytic leukaemia or multiple myeloma	INTEGER	1 ³
arsd	Advanced respiratory support days	INTEGER	3
bcsd	Basic cardiovascular support days	INTEGER	3
bpc	Biopsy proven cirrhosis	INTEGER	1 ³
brsd	Basic respiratory support days	INTEGER	3
chemox	Chemotherapy	INTEGER	1 ³
cicids	Congenital immunohumoral or cellular immune deficiency state	INTEGER	1 ³
cmlcll	Chronic myelogenous/lymphocytic leukaemia	INTEGER	1 ³
cpr_v3	Cardiopulmonary resuscitation (CPR) within 24 hours prior to admission to your unit	STRING	1
crrx	Chronic renal replacement therapy	INTEGER	1 ³
dah	Date of admission to your hospital	DATE	dd/mm/yyyy
daicu	Date of admission to your unit	DATE	dd/mm/yyyy
ddh	Date of discharge from your hospital	DATE	dd/mm/yyyy

² An eight-step process is used to match patients in CPRD primary care data (CPRD GOLD or CPRD Aurum) and ICNARC using some or all of the following: NHS number, date of birth, sex and postcode. Only data for patients matched using steps 1-5 has been provided (see section 'Linkage algorithm and the match rank variable' above for more details).

³ 0 means FALSE; 1 means TRUE (the condition or event was recorded as having occurred).

ddicu	Date of discharge from your unit	DATE	dd/mm/yyyy
dep	Dependency prior to admission to acute hospital	STRING	1
dis	Status at discharge from your unit	STRING	1
doah	Date of original admission to acute hospital	DATE	dd/mm/yyyy
doaicu	Date of original admission to critical care	DATE	dd/mm/yyyy
dod	Date of death	DATE	dd/mm/yyyy
dudh	Date of ultimate discharge from acute hospital	DATE	dd/mm/yyyy
dudicu	Date of ultimate discharge from critical care	DATE	dd/mm/yyyy
filpo	Associated FIO ₂ (<i>from arterial blood gas with lowest PaO₂</i>)	DECIMAL	1.2
hbl	Highest blood lactate	DECIMAL	2.1
hcm	Height	INTEGER	3
hcmet	Height estimated	INTEGER	1 ³
hcreat	Highest serum creatinine	INTEGER	4
hctemp	Highest central temperature	DECIMAL	2.1
hdis	Status at discharge from your hospital	STRING	1
he	Hepatic encephalopathy	INTEGER	1 ³
hhr	Highest heart rate	INTEGER	3
hnctemp	Highest non-central temperature	DECIMAL	2.1
hnvrr	Highest non-ventilated respiratory rate	INTEGER	3
hu	Highest serum urea	DECIMAL	3.1
hv	Home ventilation	INTEGER	1 ³
hvrr	Highest ventilated respiratory rate	INTEGER	3
ilpo	PaO ₂ from arterial blood gas with lowest PaO ₂	DECIMAL	3.1
lhb	Lowest haemoglobin	DECIMAL	2.1
lpc	Lowest platelet count	INTEGER	4
lsys	Lowest systolic BP	INTEGER	3
ltot	Lowest total Glasgow Coma Score	INTEGER	2
lym	Lymphoma	INTEGER	1 ³
meta	Metastatic disease	INTEGER	1 ³
pcilpo	Associated PaCO ₂ (<i>from arterial blood gas with lowest PaO₂</i>)	DECIMAL	3.1
pdial	Paired diastolic BP for lowest systolic BP	INTEGER	3
ph	Portal hypertension	INTEGER	1 ³
philpo	Associated pH (<i>from arterial blood gas with lowest PaO₂</i>)	DECIMAL	1.2
radiox	Radiotherapy	INTEGER	1 ³
rdis_v3	Reason for discharge from your unit	STRING	1
resa	Residence prior to admission to acute hospital	STRING	1
sedpar	Sedated or paralysed and sedated for whole of first 24 hours in your unit	STRING	1

sex	Sex	STRING	1
srd	Severe respiratory disease	INTEGER	1 ³
sterx	Steroid treatment	INTEGER	1 ³
taicu	Time of admission to your unit	TIME	hh:mm
tdicu	Time of discharge from your unit	TIME	hh:mm
tod	Time of death	TIME	hh:mm
udis	Status at ultimate discharge from critical care	STRING	1
uhdis	Status at ultimate discharge from acute hospital	STRING	1
vscd	Very severe cardiovascular disease	INTEGER	
wkg	Weight	INTEGER	3
wkgest	Weight estimated	INTEGER	1 ³
age	Age in whole years at admission to your unit	INTEGER	3
ap2aps	APACHE II acute physiology score	INTEGER	2
ap2score	APACHE II score	INTEGER	2
bmi	Body mass index	DECIMAL	2.2
imscore	ICNARC physiology score	INTEGER	2
pfratio	PaO ₂ /FiO ₂ ratio from arterial blood gas with lowest PaO ₂	DECIMAL	2.2
vent	Mechanical ventilation during the first 24 hours in your unit	INTEGER	1 ³