COVID-19 RESEARCH DATABASE

GUIDE FOR RESEARCHERS

VERSION: 4.0

Document Control	
Version	2020 05 12 COVID-19 Research Database – Guide for Researchers
Date Issued	12/05/20
Author(s)	eDRIS
Other Related Documents	
Comments to	michael.sibley@nhs.net

Document History				
Version	**Date**	**Comment**		**Author(s)**
1.0	12/05/20	1st published version		M Sibley
2.0	11/06/20	Additional information to reflect the changes in SICSAG approval process		J Bruce D Bailey
3.0	02/02/2021	Updated dataset list. Added details about availability of familial linkage and CURL. Information added about additional permission requirements for some datasets		
4.0	12/03/2021	Addition of vaccination data tables Addition of COGUK data		

TABLE of CONTENTS

1. INTRODUCTION

1.1 Purpose

The Chief Medical Officer has established a Scientific Advisory Group, chaired by Prof Andrew Morris and a COVID Data Task Force under this group to oversee the need for rapid access to data to support this urgent research. The COVID Data Task Force is being established with representatives (PHS, SG, HDRUK, ADR and Academia) and its remit will be to work with data controllers to identify and agree the datasets that are key to supporting rapid research and analysis.

To this end, data to respond to critical COVID-19 related research and analysis has been brought together within the National Safe Haven (NSH) from:

- Public Health Scotland
- National Services Scotland
- Scottish Government
- Open Data
- Publicly available data

This document is a guide for researchers that performs the following functions:

- Provides links to the information governance approvals for the database
- Links to appropriate websites and supporting information
- The source of the data, frequency and mechanism of transfer
- How to make a COVID-19 research request

Please send any queries to phs.edris@nhs.net .

For the duration of the COVID-19 outbreak, regular updates will be provided to the COVID Data Task Force Group, on which Phil Couser, Director of Data Driven Innovation for Public Health Scotland sits. This group will also provide a steer on what data should be included in this data platform to meet COVID research needs.

1.2 Glossary of Terms

GLOSSARY of TERMS	
ADR	Administrative Data Research
ECOSS	Electronic Communication of Surveillance in Scotland
eDRIS	electronic Data Research and Innovation Service
EPCC	Edinburgh Parallel Computing Centre
HDRUK	Health Data Research UK
NSH	National Safe Haven
PBPP	Public Benefit and Privacy Panel
PHS	Public Health Scotland
PIS	Prescribing Information System
RC	Research Coordinator
SERV-U	Managed File Transfer for the National Safe Haven
SICSAG	Scottish Intensive Care Audit Group
SMR	Scottish Morbidity Record

2. INFORMATION GOVERNANCE BACKGROUND

A PBPP application was approved by the Public Benefit and Privacy Panel for Health and Social Care (https://www.informationgovernance.scot.nhs.uk/pbpphsc/) on 23rd April 2020.

The eDRIS website (https://www.isdscotland.org/Products-and-Services/eDRIS/COVID-19/) has copies of the following documents showing the legal basis for establishing the database:

- PBPP application for the Scottish Collaborative COVID Research Data Platform
- PBPP approval letter for the Scottish Collaborative COVID Research Data Platform
- Data Protection Impact Assessment for the Research Database (to follow)

3. SPECIFICATION

3.1 Overview
The database works on the following broad mechanism:

1. eDRIS gather datasets from data controllers and transfer them securely to a SQL database in the national Safe Haven
2. Researchers review what data is available alongside supporting dataset definitions to assist in the formulation of their research questions
3. Researchers work with eDRIS to prepare PBPP applications. Further details on this process, including pricing, can be found at https://www.isdscotland.org/Products-and-Services/eDRIS/Research-Proposal/
4. If approved, eDRIS will extract the approved data from the database to the researcher's folder in the Safe Haven as per existing processes. It might also be possible for researchers to receive data outwith the NSH if PBPP deem it necessary and sufficiently robust security and processes are in place.
5. However, the research database itself persists longer term.

3.2 Datasets
The COVID-19 research database consists of the following datasets and their date ranges as at **06/05/20 and are subject to change**:

Dataset	Data Range	Frequency of Database Update
NRS Deaths	01/01/15 to present	Weekly
Infant deaths	01/01/15 to present	Monthly
NRS Stillbirths	All stillbirths	Monthly
NRS Births	All births	Weekly
SBR	All SBR	Monthly
Unscheduled Care Deaths		Weekly
SMR (all excluding 11 and 12)	01/04/97 to present	Monthly
DCVP Dispensed	Dispensing, which was paid on or after 01/01/15 to present	Monthly
DCVP Prescribed	Prescriptions, which were paid on or after 01/01/15 to present	Monthly
DCVP Paid	Paid on or after 01/01/15 to present	Monthly
SICSAG Daily	2019 to present	Weekly
SICSAG Episodes	[TBC]	Weekly
Diabetes Covariates		

ECOSS (-/+VE tested patients only)	All data	Daily
A&E		Weekly
NHS24	2011 to present	Weekly
GP OOH	2018 to present	Weekly
SAS	2011 to present	Weekly
RAPID (System Watch feed)		Weekly
ISARIC		Weekly
Scottish Stroke Care Audit (SSCA)		Monthly
Census	2011 Census	One-off
GP cluster tables		One-off
Vaccination tables		Weekly
COGUK		Daily

This guide will be updated as new datasets are added. For further details, please see section 5.

3.3 Linkage and Limitations
eDRIS replaces the CHI Number on the files with an Encrypted CHI Number (for which only a restricted group in NSS IT have the key).

We have the ability to perform familial linkage, please let your research coordinator know if this is a requirement.

We can access the CURL lookup file which links the CHI Unique Patient Identifier and the UPRN (unique property reference number) for current GP registered patients in Scotland but also includes any person who has died since January 2020. The CURL file held in PHS contains a number of processing fields including the full address, but the fields that can be made available to researchers, subject to requirement and approval. Please let your research coordinator know if this is a requirement.

4. TRANSFER MECHANISM AND TIMINGS

4.1 Mechanism

All data specific to a project will be extracted by eDRIS or EPCC NSH staff. It will then be linked, de-identified and transferred into a study folder with restricted access to only those working on that project in line with the project approvals. The Encrypted CHI Number will be replaced with a study specific ID number as per existing processes. Each project must have data access approvals in place before data will be provisioned. eDRIS has a sign out process to check the data provided is per approvals prior to granting access to the researchers. All outputs following analyses will be checked for disclosure by the eDRIS team as per existing processes, or by other data controller representatives if stipulated in their data access approvals.

4.2 Transfer Frequency
The frequency will vary between datasets as per 3.2. When data is extracted from the database for researchers following approval, it will be done using the most recent data available at the time of extraction. Depending on the nature of the project, the data provided to the researcher might be refreshed as updates from the source datasets become available. In other instances, the data provided will be a snapshot at time of extraction.

5. MAKING A COVID-19 RESEARCH REQUEST

eDRIS has well established process to facilitate researchers in making applications to access health and social care data. Details of this process can be found at https://www.isdscotland.org/Products-and-Services/eDRIS/Research-Proposal/ .

For any potential research projects utilising the database, the first step is to complete eDRIS enquiry form that can be downloaded from the webpage above. **Please state in the 'Data requirements' section of the form that access to the COVID-19 database is being requested and also in the form who has commissioned the work.** This is important as eDRIS are giving a higher priority to COVID-19 related requests as are PBPP (see https://www.informationgovernance.scot.nhs.uk/pbpphsc/). When complete, forms should be emailed to phs.edris@nhs.net .

The SICSAG steering committee have resumed more of their business as usual processes. Priority level 1 and level 0 requests for SICSAG data will now need to be reviewed by the SICSAG steering group in addition to the PBPP application. Please see their website for details and to complete their Information Request Form - https://www.sicsag.scot.nhs.uk/index.html.

NRS require a stats-PBPP application to be completed and submitted for Census requests.

ISARIC requests require approval by the consortium. More details available at https://isaric4c.net/sample_access/..

The Research Subgroup of the SSCA Steering Committee continues to oversee the use of SSCA data in research - a information request form must be completed and submitted to the committee.

If GP data is requested, PBPP are required to gain approval from a GP committee prior to the linkage and provision of any data.

Once your form has been received, an eDRIS Research Coordinator will be assigned and get in contact regarding your request. The RC will guide you through the process as quickly as possible.

To expedite the process, the eDRIS website (https://www.isdscotland.org/Products-and-Services/eDRIS/COVID-19/) contains a document showing the datasets and their data items that form the database. This document can be downloaded and the data items required selected. This can then be included at an attachment with the PBPP application.

Upon successful approval of a PBPP application, an eDRIS analyst will extract the approved data from the research database to a secure folder in the National Safe Haven that will be setup for your research project. Because the database has been established, this extraction can be expedited compared to the standard process where each data provider must then be contacted to provide the data.

COVID-19 research requests requiring data outwith the database, are still given a high priority, but will go through the standard process for extraction process and so will be slower. This is due to competing pressures at the current time on those data providers. Please also note that for administrative data, the Scottish Government is not accepting any new bespoke data access requests, unless directly COVID-19 related for Scottish Government datasets or NRS census data. Please discuss with your RC if administrative data is required.