

East London Database

Data Usage and Reporting Policy

1 Background and Legal Compliance

- 1.1 The East London Database (ELDB) is a resource maintained by the Clinical Effectiveness Group (CEG), which is an academic unit within Queen Mary University of London (QMUL). It is a series of annual databases containing a curated snapshot of the data for patients registered with GP practices within north east London on the 1st April for each year. The data is derived from the QMUL CEG Compass database, which is a subscriber database of the Discovery Data Service (DDS), covering all practices (EMIS and TPP systems) within the North East London ICB area (City, Hackney, Newham, Tower Hamlets, Waltham Forest, Redbridge, Barking, Dagenham, and Havering).
- 1.2 The data processing for Compass and ELDB is approved and defined in the Data Processing Agreement between the Discovery Board, representing North East London Integrated Care Board (NELICB), and CEG. ELDB is approved for providing information on primary services to NELICB commissioners, public health departments, NHS provider organisations and approved researchers.
- 1.3 CEG operates in accordance with the relevant UK and NHS legislation, most specifically:
 - General Data Protection Regulation (GDPR)
 - Data Protection Act 2018 (DPA2018),
 - NHS Act 2006 Section 251 (S251)
 - Health Service (Control of Patient Information) Regulations 2002
 - NHS Data Security Protection Toolkit
- 1.4 QMUL is required by law to comply with GDPR and DPA2018. It requires, therefore, CEG to process and ensure data is store in accordance with its policies:
<https://www.qmul.ac.uk/its/about/governance--committees/governance/policies/>
- 1.5 The ELDB contains de-identified patient data for the purpose of creating aggregated reports. The National Data Opt Out (NDOO) and the previous Type 1 and 2 Data Opt Out scheme only applied only to confidential patient information (CPI) and patient identifiable data (PID). On this basis, no patient data has been excluded from the ELDB. The NDOO Policy however does require appropriate organisational and technical control measures to be in place for managing de-identified data, as covered in this document. For further details, see <https://digital.nhs.uk/services/national-data-opt-out/operational-policy-guidance-document>, in particular 2.4: Type of data.
- 1.6 This document provides guidance and requirements for the use of the East London Database for users external to CEG, in order to ensure CEG meets its statutory and organisational requirements. As part of those obligations, it also sets out reporting requirements designed to maintain the confidentiality of individual patients and to guard against the inadvertent creation of unsafe or misleading analyses.
- 1.7 The development and use of the ELDB data is managed through the ELDB Steering Group, which meets at regular intervals throughout the year.

- 1.8 The ELDB contains tables of patients with defined conditions; the results of specific tests or examinations; some status indicators; and key demographics. The latest and/or earliest date is provided, depending on the most practical application, and other fields are provided to assist the analyst. The full dataset for a patient or a condition is not available in the ELDB. CEG may revise the data and/or table structures, in order to improve the accuracy or useability of the database. The full details of the ELDB contents and further information is available on the CEG website:
- 1.9 <https://www.qmul.ac.uk/ceg/data-resources/east-london-database/> Users can request changes or additional data tables via the ELDB Steering Group. Requests for data extracts outside the remit of the ELDB should be made to CEG separately.

2 Data Connection and Access

- 2.1 Anyone who works with, or may have access to, the ELDB data must be aware of their requirements and responsibilities under GDPR/DPA2018 and the requirements for NHS safe data handling. All ELDB users and data handlers must therefore:
- Complete, and maintain on an annual basis, information governance training that complies with the NHS Core Skills Training Framework (CSTF).
 - Provide evidence of that training for CEG records.
- 2.2 Users are expected to be fully cognisant of the policies of their own organisation, governing the accessing, processing, storing and sharing of data, and complete all required associated training.
- 2.3 Connection to the ELDB and analysis of the ELDB data must be via the designated CEG VPN and a secure, non-public network within the UK.
- 2.4 Connection to the ELDB and analysis of the ELDB data must be via a device that uses an Operating System (OS) actively supported by the supplier with regular security updates and fixes.
- The device must be maintained with up-to-date security patches and updates.
 - The device must have active, up-to-date anti-virus/malware software application(s) as appropriate for the OS.
 - The device must utilise available chip-based security features: TPM and BitLocker for Windows, T2 and Startup Security Utility for macOS.
- 2.5 Connections from automated devices and scripts running on a Linux server or other such methods, should be discussed with CEG.
- 2.6 Connection to the ELDB and analysis of the ELDB data must be conducted in a secure, non-public, environment within the UK.
- 2.7 ELDB users must not attempt to access any other CEG data system, unless approved by CEG.
- 2.8 CEG audits the activity of all data systems on a regular basis using logs and records specific to each system. This may include reviewing or monitoring individual user actions, scripts or connections.
- 2.9 All users must maintain awareness of the risk of data loss or exposure. In particular:
- Data files and systems should not be left open when users are not at their desk.
 - Passwords and PIN numbers should not be stored on or with devices or PCs.

- Logins must not be shared between users.
 - Devices used for authentication should be kept safe and secure
- 2.10 Any identified or suspected data loss, data exposure or breach of data security must be reported to CEG immediately.
- 2.11 CEG may place a temporary hold on a user's access into ELDB whilst it investigates any apparent breach of these requirements or whilst a user is taking action to ensure their compliance. CEG maintains the right to refuse access for any user that is deemed to not be meeting these required standards.

3 Data Extraction, Storage and Linkage

- 3.1 ELDB users are not permitted to extract and download line level data (ie patient level) from the ELDB database. All extractions and reports must aggregate counts or calculations (for example by age band, locality or practice - see 4.2 below).
- 3.2 Processed data extracts and reports from ELDB must be stored in a secure location that meets the information governance requirements stipulated in section 1. CEG may require a user or organisation to provide the details of the data storage environment that is in use and confirmation that the appropriate policies and safeguards are in place.
- 3.3 The linkage of other datasets with the ELDB patient data is not permitted. Linkage to non-patient information, such as GP Practice or LSOA, should be undertaken only with discussion and agreement with CEG.
- 3.4 Users are permitted to query, extract and process data only be for the purposes of Local Authority planning, clinical research and supporting NHS activities and organisations.

4 Data Reporting and Statistical Disclosure Control

- 4.1 Clinical data is not always recorded consistently, complete or as expected. It is strongly recommended that ELDB users discuss their proposed data analyses with CEG, either directly or within ELDB User Group meetings, in order to ensure their reports are accurate.
- 4.2 To safeguard against unsafe or misinterpretations of the data, CEG requires ELDB data to be aggregated on the following categories:
- Organisation (GP Practice, PCN etc)
 - Geographical categories (Borough, LSOA etc)
 - Age Bands with 5+ year banding
 - Gender (Male Female, Other & Unknown)
 - Ethnicity with the standard 18+1, 16+1 or 5+1 categories
 - Time Periods of a month or greater

Any other aggregation must be discussed and agreed with CEG

- 4.3 All users should be aware of statistical disclosure control (SDC) methods and apply these measures where appropriate to ensure the non-disclosure of identifiable, individual information. Whilst CEG primarily applies a Principle Based approach to

SDC, users should be prepared to discuss and justify any SDC methods used in their reporting.

- 4.4 As aggregated counts form a large part of ELDB analysis, CEG is specifically concerned with small cell counts. Users should pay particular attention to table counts and the application of SDC suppression techniques.
 - Cell counts should only as small as required for reporting purposes.
 - Counts <5 should generally be suppressed, unless they can be clearly shown to non-disclosive.
 - For clarity, suppressed counts should be appropriately indicated ('-', '<5', 'SUPP' etc).
 - Care should be taken not to allow secondary disclosure, where it is possible to calculate a suppressed count using accompanying unsuppressed counts.
- 4.5 ELDB does not provide data on patient sexual activity, sexual orientation, religion or ascribed racial or ethnic origin. Users must not attempt to infer or report on these categories using the ELDB data and must take care that no such inference can be taken from their report. Users should discuss with CEG any reporting that may contain counts or information that could be understood as sensitive.

5 Data Publication and Acknowledgment

- 5.1 Any reports or products created using ELDB data remain under the purview and aegis of CEG.
- 5.2 CEG should be informed at an early stage of any project involving the large-scale reporting of ELDB data. Where it is felt necessary, CEG will refer back to the GP Practices, either individually or through the Local Medical Committee (LMC), as the data controllers, to seek agreement for proposed ELDB data reporting.
- 5.3 The use of ELDB data for any large-scale reporting (eg online dashboards) should be brought to the ELDB Steering Group for discussion and agreement.
- 5.4 CEG retain the right to request copies of any data reports created from ELDB data and must be provided with appropriate accounts and access to allow the review of any online dashboards created from ELDB data.
- 5.5 All outputs created from ELDB data (including but not limited to presentations, reports, dashboards, published research) must contain the following acknowledgement of the ELDB data source and of CEG:

Data source: CEG East London Database, courtesy of the Clinical Effectiveness Group at Queen Mary University of London.

Where appropriate, the Queen Mary and CEG logos should also be displayed to indicate us as a collaborative organisation – these can be provided on request.

6 SDC References and Resources

NHSBSA Statistical Disclosure Control Protocol (2020)

<https://www.nhsbsa.nhs.uk/policies-and-procedures>

Ritchie, F. et al (2023) *The SACRO Guide to Statistical Output Checking*. Zenodo

<https://doi.org/10.5281/zenodo.10282526>

Welpton, R. (2019) *SDC Handbook*. Figshare
<https://doi.org/10.6084/m9.figshare.9958520.v1>

ONS *Protecting Personal Data in Census 2021 Results* [webpage]
<https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/methodologies/protectingpersonaldataincensus2021results>

7 CEG Contact

All queries and discussion concerning the ELDB should be directed to Martin Sharp (m.a.sharp@qmul.ac.uk), who will action or pass on the information as required.

Version Control

Version	Date	Summary of Changes
1.0	14/11/2024	New document developed from CEG Data Policy v1.2

Authorisation

Created by	Date	Approved by
Kelvin Smith	14/11/2024	CEG Senior Management Team