

Design and architecture of the CARA infrastructure for visualising and benchmarking patient data from general practice

Nathaly Garzón-Orjuela ^{1,2}, Agustin Garcia Pereira ³, Heike Vornhagen ³, Katarzyna Stasiewicz³, Sana Parveen ^{1,2}, Doaa Amin ^{1,2}, Lukasz Porwol ³, Mathieu d'Aquin ⁴, Claire Collins ^{5,6}, Fintan Stanley⁵, Mike O'Callaghan ⁵, Akke Vellinga ^{1,2}

To cite: Garzón-Orjuela N, Garcia Pereira A, Vornhagen H, *et al.* Design and architecture of the CARA infrastructure for visualising and benchmarking patient data from general practice. *BMJ Health Care Inform* 2024;**31**:e101059. doi:10.1136/bmjhci-2024-101059

► Additional supplemental material is published online only. To view, please visit the journal online (<https://doi.org/10.1136/bmjhci-2024-101059>).

This study was previously presented as an abstract (poster display) at the 16th European Public Health Conference, Digital health and communication, Dublin, Ireland, from 8 to 11 November 2023.

Received 21 February 2024
Accepted 27 July 2024



© Author(s) (or their employer(s)) 2024. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to
Professor Akke Vellinga;
akke.vellinga@ucd.ie

ABSTRACT

Objective Collaborate, Analyse, Research and Audit (CARA) project set out to provide an infrastructure to enable Irish general practitioners (GPs) to use their routinely collected patient management software (PMS) data to better understand their patient population, disease management and prescribing through data dashboards. This paper explains the design and development of the CARA infrastructure.

Methods The first exemplar dashboard was developed with GPs and focused on antibiotic prescribing to develop and showcase the proposed infrastructure. The data integration process involved extracting, loading and transforming de-identified patient data into data models which connect to the interactive dashboards for GPs to visualise, compare and audit their data.

Results The architecture of the CARA infrastructure includes two main sections: extract, load and transform process (ELT, de-identified patient data into data models) and a Representational State Transfer Application Programming Interface (REST API) (which provides the security barrier between the data models and their visualisation on the CARA dashboard). CARAconnect was created to facilitate the extraction and de-identification of patient data from the practice database.

Discussion The CARA infrastructure allows seamless connectivity with and compatibility with the main PMS in Irish general practice and provides a reproducible template to access and visualise patient data. CARA includes two dashboards, a practice overview and a topic-specific dashboard (example focused on antibiotic prescribing), which includes an audit tool, filters (within practice) and between-practice comparisons.

Conclusion CARA supports evidence-based decision-making by providing GPs with valuable insights through interactive data dashboards to optimise patient care, identify potential areas for improvement and benchmark their performance against other practices.

Supplementary file 1. Graphical abstract

INTRODUCTION

Patient management software (PMS) systems were introduced in the 1990s in Ireland and

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Health data is generally stored in isolated systems (data silos), making it difficult for healthcare professionals to access data in aggregated form or to compare across different systems. This issue is prominent in primary care, where patient data is commonly recorded using patient management software (PMS) systems, which do not allow analysing the aggregated data stored in the backend. No publications have been identified that presented a data extraction process, any interactive elements to compare or filter or the potential to sustain the infrastructure or dashboard beyond the aim of the project.

WHAT THIS STUDY ADDS

⇒ Health research funders and governments developed data-sharing policies. However, most data sharing fails due to a lack of common data models and infrastructures supporting such models. We developed a comprehensive infrastructure, including a data extraction tool (CARAconnect) to combine data from different general practices with different PMS, circumventing the need to streamline and link data at the source. The tools can easily be adapted to include/exclude specific data fields to adhere to ethical and General Data Protection Regulation requirements. Additionally, the infrastructure supports visualising within-practice and between-practice comparisons (anonymised). The CARA infrastructure offers a blueprint for extracting, transforming and sharing data from data silos and presenting data in interactive visualisations.

have a client-server architecture allowing general practitioners (GPs) to access patient information from within the practice.^{1,2} Data is stored centrally on the practice server in multiple separate tables, adding tables during a continuing data modelling process to include more or new information. This has

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Visualising health data allows comparisons within and between general practices, countries and institutions. However, the challenge with data stored in data silos remains and prevents the possibility of aggregated data visualisation to understand patterns and provide better treatments and interventions. The CARA infrastructure described in this paper supports data extraction, transformation and sharing across data silos.

resulted in a complex database with an expansive list of tables interlinked through various relationships defined in consultations, patients, prescriptions and other codes. Despite efforts to achieve data interoperability, the data remains sequestered within the general practice, which acts as a data silo.

PMS is organised around consultation to support the management of individual patients and the organisation of general practices. However, PMS does not allow straightforward practice-level data aggregation to visualise, compare or benchmark patient and consultation data. This hampers overall care improvement as detecting trends, revealing occurrences, reducing the incidence of medical errors and identifying quality improvement opportunities require within-practice and between-practice comparisons.³

Four accredited PMS systems are used in Ireland but two main PMS systems account for over 90% of the practices.⁴ The general practice is that the data controller and GPs can, as part of training, teaching, auditing and research, analyse their practice data.⁵ An annual clinical audit is a requirement for the renewal of their clinical competence certification by the Medical Council, and its provision as part of a (research) project has been shown to increase participation and commitment from GPs.⁵

The Irish health system is characterised by a complex structure involving a mix of public and private financing and service delivery.⁶ Ireland does not have universal health coverage. Approximately 31% of the Irish population are entitled to a general medical services card, while a further 11% have a doctor-visit card and both secure free access to general practice care.⁷ Medical cards are allocated based on age, economic factors or long-term illnesses. Patients without a medical card pay per GP visit (up to €70) and often take out additional private health insurance.⁸ Detailed primary care information is only available to the Health Service Executive (HSE) from medical card holders, as part of the HSE's payment to GPs for their service. The HSE directs the public healthcare system, including acute care, primary care and community care services.⁹

In 2019, the Irish Medical Organisation, the HSE and the Department of Health reached an agreement to support and maintain general practice services and ensure the effective interoperability between PMS systems and the HSE information technology systems to integrate some information in eHealth Ireland.⁷ However, records

are not accessible by GPs once submitted and no overview of practice-level or individual feedback is provided. Another improvement initiative was developed for antibiotic prescribing in which quarterly feedback is provided on antibiotic prescribing for medical card patients.¹⁰ The design of this feedback has been reported to be complex and without an interactive option to query practice data, while it only includes information from public patients, excluding antibiotic prescribing for private patients.¹¹

The COVID-19 pandemic highlighted the importance of data accessibility and sharing across countries and institutions, and the initial fragmented public health response to the pandemic demonstrated the value of timely, publicly available data.¹² The pandemic also led to some global bodies collecting datasets at the country level and illustrating the situation in interactive dashboards.¹³ Visualising COVID-19 data allowed comparisons within and between countries; however, the issue with data stored in isolated data silos remained. This challenge persists in general practice where reusing and sharing accurate and detailed health data recorded by GPs and stored in PMS systems is not possible.¹⁴

CARA set out to provide a sustainable infrastructure to help GPs understand their patient population and their disease management in addition to monitoring their prescribing through the use of dashboards. The first exemplar dashboard focused on antibiotic prescribing to develop and showcase the proposed infrastructure, including an audit tool as well as filters (within-practice) and between-practice comparisons. This paper explains the design and development of the CARA infrastructure, which consists of CARAconnect, a data extraction tool, a data model (to combine data from different PMS systems) and the CARA dashboard for use in Irish general practice.

METHODS

Definition of requirements

The development of a valuable integrated infrastructure for healthcare data is a complex task and demands a very close interaction with the key stakeholders—GPs. Therefore, an agile development process was applied, in which a user-centred design approach was taken to ensure the prioritisation of GPs' needs at every project stage.^{15 16} A number of user stories were created in cooperation with seven GPs to describe possible data exploration scenarios which were used to create data requirements. These requirements informed the data modelling process and a set of variables for the initial data extraction.

For example, one user story focused on a GP's need to visualise their practice population, which formed the basis for a set of 'patient characteristics' variables. This process was iteratively reviewed and adapted to ensure that the extracted data still fulfilled the requirements of the user story. The user story further informed a 'concept dashboard', used as an example when involving GPs and creating new user stories and thereby actively involving

Table 1 Main requirements

Requirement description	Challenges	Proposed solution
General practices should be able to use the platform independently of their patient management software (PMS)	General practices use different PMS with different complex database schemas and are owned by private companies	Data will be extracted, integrated and transformed into common data models Developing an external platform (CARA dashboards) to provide a more comprehensive and flexible solution for General practitioners (GPs) that enables more accessible updates and additions of new features without affecting the core PMS functionality.
Practice data to allow for between-practice comparisons and research	Data from general practice stored in independent databases located on practice servers Comparison with other practices requires high computational power to be aggregated	Independent and secure servers Using extract, load and transform (ELT) approach
De-identified patient data has to be stored securely and trusted	GP/practice is a data controller. Data sharing has to be restricted and regulated	Servers provided by the Irish College of General Practitioners Research ethical approval (including Data Protection Impact Assessment)
Patients and practices cannot be identified	GPs should be able to see their own data but not be able to identify other practices	Data will be extracted from a static time point and will overwrite previous versions for a one-way de-identification
	Potential identification of patients for rare diseases or prescriptions	k-anonymisation will be applied with k=5
	Data anonymisation will require high computational power. Trade-off between security and usability	Anonymisation during extraction, using hashing and salting
GPs should be able to perform audits for quality improvement and yearly clinical professional competence reporting	Data should allow and support comparisons over time or before/after events	Multiple year data extraction and integration of automated audits
Patients should be able to opt out of their data being extracted (General Data Protection Regulation)	PMS does not currently offer an option to facilitate this requirement	A data entry field was identified in the PMS to indicate the exclusion of a record
Data visualisations should be available on demand and interactive	GPs' practice computers do not have the computing power or resources to run complex queries or transformations	Data models stored on the servers include the transformations necessary for the dashboards

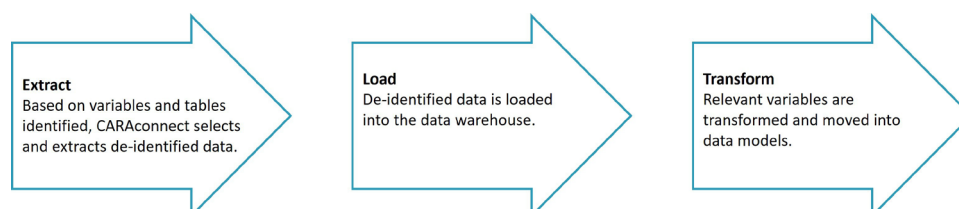
ELT, extract, load and transform; GPs, general practitioners; PMS, patient management software.

GPs in proposing, developing and identifying visualisations of interest and consequent extraction requirements. An overview of the requirements, its challenges and the solution proposed is shown in [table 1](#).

Data integration

A set of variables for the initial data extraction was identified, and a data transformation pipeline was created to transform the extracted data into bespoke data models. To this end, practice data was extracted from various

tables stored on the practice server and processed before being moved into new data models. Data integration uses the extract, load and transform process (ELT) ([figure 1](#)), where data from general practices is 'Extracted', 'Loaded' into a data warehouse and 'Transformed'.¹⁷ This ensures that data necessary for any transformation is always loaded and does not need to be re-extracted.


Figure 1 Extract, load and transform approach.

Coding systems

The data extracted includes the two main coding systems for classifying disease and therapeutic prescriptions, which are used by general practices in Ireland.¹⁸ The International Classification of Primary Care (ICPC) was designed to capture the interaction between the GP and the patient and is structured around the consultation. ICPC has fewer diagnosis codes than other systems, such as the International Classification of Diseases (ICD-10).¹⁸ PMS includes both standard coding systems for coding consultations and diseases.¹⁸ However, GPs are not incentivised to code and few GPs code consultations, making the prevalence of common diseases difficult to measure accurately.² Coding in Irish general practice has improved for a few selected conditions since the introduction of the chronic disease management (CDM) programme in 2020. CDM was integrated into the four accredited PMS systems resulting in the automatic recording of CDM disease codes.¹⁹

Therapeutic prescriptions are coded within Irish PMS using the Anatomical Therapeutic Chemical (ATC) code, a unique code assigned to any medicine according to the organ or system it works on and its action.²⁰ The classification system is maintained by the WHO. ATC codes have five levels, the highest level is a letter 'X' for the main group (eg, J), the second level is two numbers '##' for the therapeutic group (eg, 01), the third level is a letter 'X' for the pharmacological subgroup (eg, C), the fourth is a letter 'X' for the chemical subgroup (eg, A) and the final, fifth level is the active substance '##' (eg, 01). This results in ATC codes written as X##XX## (ie, J01CA01 for the antibiotic ampicillin). This allows easy classification of medicines to specific or larger therapeutic groups.²⁰

RESULTS

The CARA data processing infrastructure encompasses the general practice database, secure data extractor and uploader, data integration component and a database that connects to the relevant dashboard interfaces. **Figure 2** shows the architecture of the CARA infrastructure, which

includes the ELT process to extract, load and transform de-identified practice data into data models, the Representational State Transfer Application Programming Interface (REST API), which provides the security barrier between the data models and their visualisation on the CARA dashboard.

Data integration

Data extraction and loading

CARAconnect was created to facilitate the data extraction process. The data extraction and loading process encompasses selecting relevant data from the practice database, data de-identification and de-identified data upload to the CARA remote servers. In order to conduct this process in an automatic, structured and secured way, a desktop application was developed to streamline this task and assist the GP. CARAconnect was envisaged as an easy-to-use application once the practice was registered. The link to download CARAconnect is sent in an email to the practice secure email account (see CARA registration process below) and can easily be downloaded and initiated by a double click. On activation, CARAconnect identifies the practice server(s). On the practice server, the database is identified and selected fields from different tables are automatically extracted and securely uploaded to the infrastructure and processed before saving into the new data models during the data transformation steps. At each stage, confirmation by the GP is requested to start extraction and to finalise and upload/send the data.

Through extensive exploration and elimination, a basic understanding of practice databases was developed, which was tested with a reference (anonymous) dataset to finalise specific tables needed to fulfil the data model requirements. CARAconnect facilitates the secure data extraction process based on the variables and tables identified.

Data security and confidentiality

CARAconnect extracts de-identified data from the practice database and ensures data security through a practice-specific login-based access with two-factor authentication.

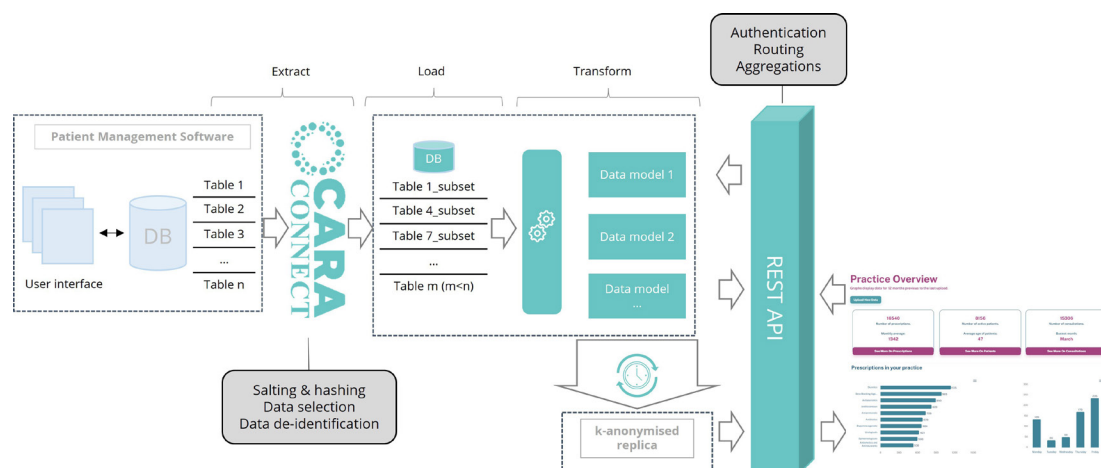


Figure 2 The architecture of the CARA infrastructure.

GPs can view their practice data but are only allowed to view aggregated practice data from all other participating practices to avoid possible identification of another practice. Uploads are for a static period to irrevocably de-identify the practice data. Any new data upload overwrites previously uploaded data. Data extraction does not include any patient identifiers or free text, nonetheless, specific technical identifiers needed to link together data in different tables are hashed and extracted. As combinations of specific variables with other externally identifiable data sources may potentially lead to identification, two additional processes were applied to facilitate the use of specific technical identifiers to link data in different tables²¹:

- ▶ Salted hashing is unidirectional encryption and decryption is almost impossible. Salting introduces an additional random part (a set of strings of fixed length) to a hash function to create a one-way function. The random part remains the same during extraction but is unique for every extraction.²¹
- ▶ k-anonymisation is applied to the data from all other practices, with k set at 5. This guarantees that a minimum of five similar patients are included in any comparison (between practice) visualisation. A k-anonymised dataset implies that each record is de-identified from at least k - 1 other.

To accommodate the patients' right to object to the processing of personal data, which provide the option to exclude their data from data processing for research purposes in accordance with General Data Protection

Regulation (GDPR) requirements,⁵ a data entry field was identified in the PMS to indicate the exclusion of a record.

Data transformation

The data extracted and loaded during the previous steps is not fully compatible with the desired output format for analysis and dashboard creation, and data quality, compatibility and usability must be ensured. In order to convert the data into the desired format, bespoke target data models were created based on the data requirements gathered (figure 3) and a number of data transformations were written to populate these models. The bespoke data models 'CARA Consultations' and 'CARA Prescriptions' represent the main sources of data for the dashboard visualisations and constitute the final product of the data transformation process.

Data transformations include date formatting, field calculation (such as age), mapping to different standards (such as ATC, ICPC or ICD-10), mapping to created taxonomies (such as green and red antibiotics) and data aggregations (such as antibiotics aggregations per consultation).

For the purpose of visualisations, the ICD-10 codes were mapped to the ICPC codes, taking a pragmatic approach and considering their occurrence in general practice. For the exemplar dashboard for antibiotics, the antibiotic ATC code J01 was used and subdivided into classes J01C. Additionally, a second categorisation implemented as part of a national antimicrobial stewardship initiative that

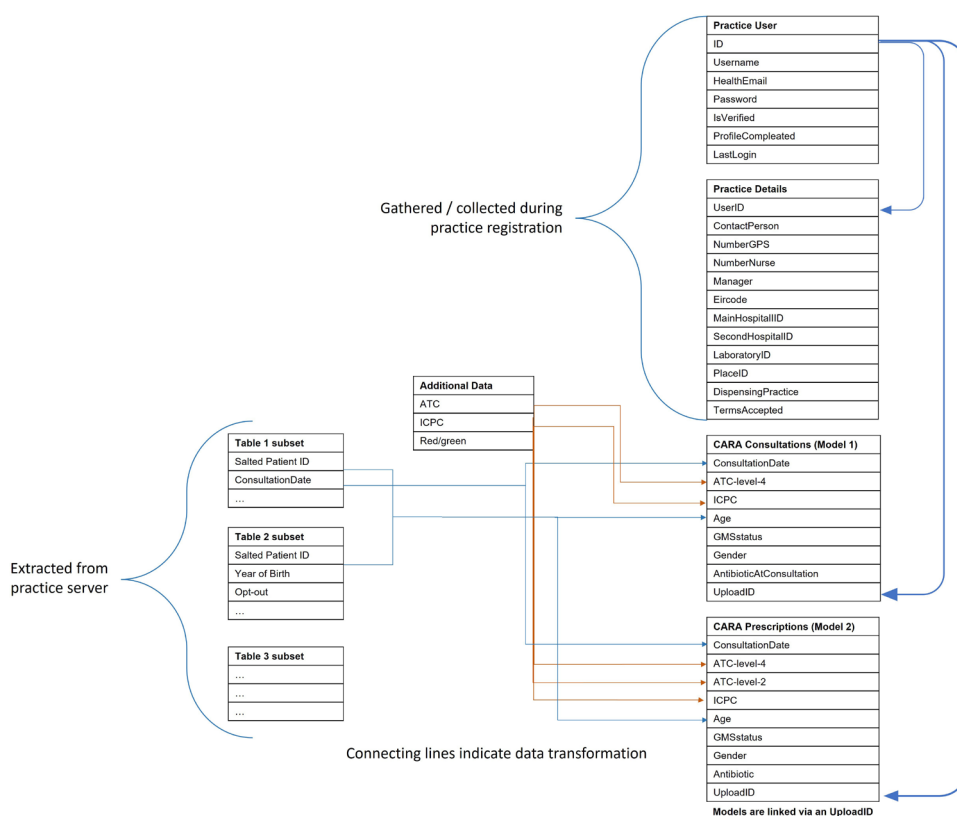


Figure 3 Data requirements gathered to data model.

divides antibiotics into green (preferred) and red (non-preferred) antibiotics guidelines was used.^{10 22}

CARA Consultations and CARA Prescriptions are not the only data models created. Two other data models were created to support the dashboard functionality. 'Practice User' represents the general practice and conceptualises its main attributes, including the credentials for login. 'Practice Details' depicts the practice in-detail aspects, such as location, type of practice (rural/urban/mixed) and laboratory they send their samples to.

CARA dashboard

The development and testing of the CARA dashboard have been described in detail elsewhere.²³ CARA dashboard was developed as a web application for GPs to visualise and explore their data and metrics visually and interactively. The data models created in data transformation connect to the CARA data dashboard and enable the analytical features. To create a visualisation from the data models, a REST API was set up. It allows applications or devices to communicate with each other using Hypertext Transfer Protocol (HTTP) methods. In this case, a REST API is used for communication between the dashboard running in the GP internet browser and the CARA infrastructure. The REST API handles user authentication and registration and provides a secure interface to query different models using aggregations and filters.

The CARA dashboard customises open-source tools and different software frameworks and libraries for data manipulation, transformation and visualisation, using Python, Javascript and a Postgres database. The dashboard is provided through charts (pie, bar, scatter and line) and allows GPs to compare their data with data from other unidentifiable practices. The visualisations can be used to generate automatic audit reports. The dashboard is based on GP requirements as elicited from a series of interviews. This led to the creation of a prototype that was tested with users before being included in the overall data infrastructure.

CARA registration process

The CARA network registration process for general practices follows a well-defined workflow to ensure secure and efficient onboarding (figure 4). Initiating the registration, new practices provide necessary details, and the system validates that the email is associated with the closed and secure email service adopted in Irish general practice and hospitals as the primary mechanism for secure communication between health systems (hospitals, laboratories, general practices, pharmacies).⁷ On entering the registration details, a one-time password (OTP) is generated and sent via email to the provided secure email address. Users are required to enter the OTP within a specified time limit. In case of delay, an option to regenerate a new OTP is available. Following successful OTP validation, users proceed to fill out the registration form, where confirmation of terms and conditions (GP agreement) is a prerequisite before finalising the registration. The GP

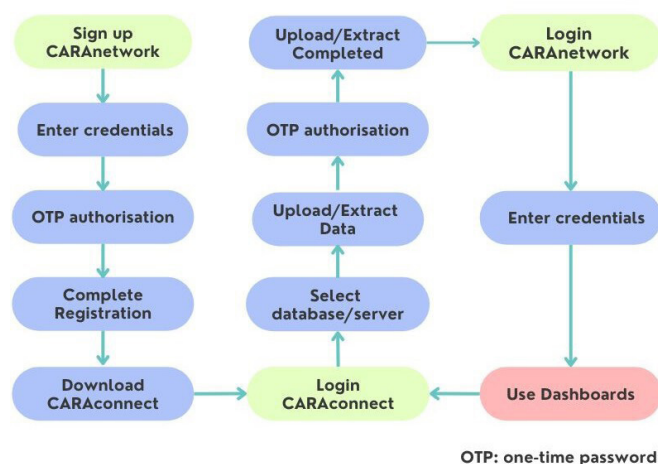


Figure 4 CARA registration process for general practices.

agreement includes a detailed list of the data that will be extracted and an explanation of the aggregated use of this data for practice comparisons and for research purposes. Afterwards, an email is sent to the secure email address, facilitating the download of the CARAconnect application. The users can subsequently login using their registered credentials, ensuring a seamless and secure experience throughout the CARA network registration and login process. At every subsequent data upload, the GP agreement has to be re-confirmed.

DISCUSSION

This paper describes the design and development of the CARA infrastructure to help GPs explore and visualise their practice data and compare this to the combined data from all other practices using interactive data dashboards. CARAconnect allows seamless connectivity for and compatibility with the main PMS, as well as making it possible to include other PMS in the future by specific identification and mapping of the relevant data fields.

A unique feature of CARA infrastructure is using ELT instead of 'Extract, Transform, Load' (ETL). ELT includes faster loading times, more flexibility, the possibility of cloud computing to deliver data in real-time and the separation of potential errors between the loading and the transformation stages.¹⁷ While the ETL approach has been implemented successfully in healthcare systems,²⁴ using ELT is more effective in ensuring successful data integration.²⁵ This implementation should consider factors from the system quality (availability, responsiveness, reliability, usability, capability, compatibility, safety and maintenance) and data quality (perfection, validity, consistency, accuracy, integrity and timeliness).²⁵ The CARA infrastructure will continue to integrate other data sources, and using ELT will avoid potential data problems (missing data, mismatched data, etc).

The CARA infrastructure ensures data security and confidentiality by extracting de-identified data, thereby addressing GDPR considerations concerning privacy and

security.²⁶ Data minimisation is applied as only limited data on prescribing and consultations is extracted at a practice level, avoiding the identification of individual GPs. Any personal patient data is de-identified and anonymised, and individual patients can request their data to be excluded from any extraction.

The CARA infrastructure is unique, in that it combines primary data sources (practice databases), a user-centred approach (involving GPs in the design and testing of the CARA dashboard throughout) and an extensive architecture. Other similar projects have included only administrative health databases as data sources,²⁷ do not offer a description of either the architecture or the data integration used or omit a description of user involvement.^{28 29} This has allowed the CARA network to ensure the project addresses users' needs and will be scalable and sustainable beyond the life of the project.

A limitation of the CARA infrastructure is that it was developed based on two PMS in Ireland. Other systems, such as different PMS or electronic health records, may require adapting the current mapping system, which could involve challenges around data integration, breadth of use, data complexity and statistical rigour.³⁰ Another limitation is that the CARA infrastructure has yet to be implemented in real-world practice, posing a challenge in optimising its functionality. However, the innovation of the CARA infrastructure lies in its ability to offer a novel approach to overcoming data silos while providing a framework that can be universally applied across diverse systems. The CARA infrastructure is the first step in tackling data fragmentation, visualising aggregated data and exploring treatment strategies and interventions. However, the innovative implication of the CARA infrastructure lies in its ability to offer a novel approach to overcoming data silos while providing a framework that can be universally applied across diverse systems. Patient consultations are typically documented in various PMS in general practice. While these systems facilitate the efficient administration of consultations, PMS often lack the capability for in-depth analysis of aggregated patient data or the ability to compare and benchmark with other general practices. Therefore, the CARA infrastructure serves as a pivotal initial step in tackling the issue of data fragmentation, enabling the visualisation of aggregated data to discern patterns and enhance treatment strategies and interventions.

CONCLUSION

CARA set out to provide a sustainable infrastructure to help GPs understand their patient population and management in addition to monitoring their prescribing through the use of data dashboards. Currently, two dashboards are presented, a practice overview and an exemplar dashboard focused on antibiotic prescribing, which includes an audit tool in addition to filters (within-practice) and between-practice comparisons. This provides a reproducible template to access and visualise general practice

data in a setting where a centralised database does not exist. CARA supports evidence-based decision-making by providing GPs with valuable insights through interactive data dashboards to optimise patient care, identify potential areas for improvement, and benchmark their performance against other practices.

Author affiliations

¹CARA Network, School of Public Health, Physiotherapy and Sports Science, University College Dublin, Dublin, Ireland

²School of Public Health, Physiotherapy and Sports Science, University College Dublin, Dublin, Ireland

³Insight Centre for Data Analytics, University of Galway, Galway, Ireland

⁴LORIA - Laboratoire Lorrain de Recherche en Informatique et Applications, Université de Lorraine, Nancy, France

⁵Irish College of General Practitioners, Dublin, Ireland

⁶Public Health and Primary Care, Ghent University, Ghent, Belgium

X Nathaly Garzón-Orjuela @nathalyGarzonO

Contributors AV conceived the CARA project and is its guarantor; AV, AGP and KS developed and tested CARAconnect and developed the infrastructure; AV, AGP, HV and NG-O organised, synthesised the information and wrote the first draft; NG-O, LP, CC, Md'A, MO'C, FS, SP and DA were involved in the development of the CARA infrastructure and dashboard and critically revised drafts. All authors participated in the manuscript revision and approved the final version. NG-O and AGP are co-primary authors

Funding This work was funded by grant number RL-20200-03 from Research Leader Awards (RL) 2020, Health Research Board, Ireland. NG-O and DA were funded by this grant and are a SPHeRE scholar.

Competing interests None declared.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data sharing not applicable as no datasets generated and/or analysed for this study.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Nathaly Garzón-Orjuela <http://orcid.org/0000-0001-6181-8154>

Agustín García Pereira <http://orcid.org/0000-0003-1874-3454>

Heike Vornhagen <http://orcid.org/0000-0002-7254-9786>

Sana Parveen <http://orcid.org/0000-0003-2382-1066>

Doaa Amin <http://orcid.org/0000-0001-7114-1693>

Lukasz Porwol <http://orcid.org/0000-0003-2417-3654>

Mathieu d'Aquin <http://orcid.org/0000-0001-7276-4702>

Claire Collins <http://orcid.org/0000-0001-8967-5159>

Mike O'Callaghan <http://orcid.org/0000-0002-8450-186X>

Akke Vellinga <http://orcid.org/0000-0002-6583-4300>

REFERENCES

- Licken E. Software claims to improve health of GPs' records Business. *The Irish Times* 1998. Available: <https://www.irishtimes.com>

- com/business/software-claims-to-improve-health-of-gps-records-1-153139
- 2 Pericin I, Larkin J, Collins C. Diagnostic coding of chronic physical conditions in Irish general practice. *Ir J Med Sci* 2022;191:1693–9.
 - 3 Rostamzadeh N, Abdullah SS, Sedig K. Data-driven activities involving electronic health records: an activity and task analysis framework for interactive visualization tools. *MTI* 2020;4:7.
 - 4 Irish College of General Practitioners. Accredited GP practice software management systems. 2022. Available: https://www.icgp.ie/go/in_the_practice/it_in_the_practice/software_companies
 - 5 ICGP Data Protection Working Group. *Processing of patient personal data: a guideline for general practitioners*. Irish College of General Practitioners, 2019.
 - 6 Cullinan J, Connolly S, Whyte R. The sustainability of Ireland's health care system. In: Baltagi BH, Moscone F, eds. *The sustainability of health care systems in Europe*. Emerald Publishing Limited, 2021: 61–79.
 - 7 Health Service Executive. Terms of Agreement between the Department of Health, the HSE and the IMO regarding GP contractual reform and service development. 2019.
 - 8 Sean M. Private expenditure on general practice Irish government economic and evaluation service. 2020. Available: <https://igees.gov.ie/wp-content/uploads/2020/11/Private-Expenditure-on-GP-Care-in-Ireland.pdf>
 - 9 Health Service Executive. HSE organisational structure. 2023. Available: <https://www.hse.ie/eng/about/who/>
 - 10 O'Connor N, Breen R, Carton M, et al. Improving the quality of antibiotic prescribing through an educational intervention delivered through the out-of-hours general practice service in Ireland. *Eur J Gen Pract* 2020;26:119–24.
 - 11 Roche KF, Morrissey EC, Cunningham J, et al. The use of postal audit and feedback among Irish General Practitioners for the self-management of antimicrobial prescribing: a qualitative study. *BMC Prim Care* 2022;23:86.
 - 12 Gardner L, Ratcliff J, Dong E, et al. A need for open public data standards and sharing in light of COVID-19. *Lancet Infect Dis* 2021;21:e80.
 - 13 European Centre for Disease Prevention and Control. Download COVID-19 data sets. 2023. Available: <https://www.ecdc.europa.eu/en/covid-19/data>
 - 14 Dagliati A, Malovini A, Tibollo V, et al. Health informatics and EHR to support clinical research in the COVID-19 pandemic: an overview. *Brief Bioinformatics* 2021;22:812–22.
 - 15 Hartson R, Pyla P. Chapter 10 - UX design requirements: user stories and requirements. In: Hartson R, Pyla P, eds. *The UX book*. 2nd edn. Boston: Morgan Kaufmann, 2019: 227–50.
 - 16 The Interaction Design Foundation. User centered design. n.d. Available: <https://www.interaction-design.org/literature/topics/user-centered-design>
 - 17 Singhal B, Aggarwal A. ETL, ELT and reverse ETL: a business case study. 2022 Second International Conference on Advanced Technologies in Intelligent Control, Environment, Computing & Communication Engineering (ICATIECE); 2022
 - 18 Meade B. *Clinical disease coding and classification: an overview for general practitioners*. Irish College of General Practitioners, 2011.
 - 19 Irish College of General Practitioners. Q. Are there any tips to improve the efficiency of my practice with the new chronic disease management programme? 2020. Available: https://www.icgp.ie/go/in_the_practice/it_faqs/managing_information/13719E9B-A6B7-4E58-A127A3516A647676.html
 - 20 World Health. Anatomical therapeutic chemical (ATC) classification. n.d. Available: <https://www.who.int/tools/atc-ddd-toolkit/atc-classification>
 - 21 Ghosh A, Singh R, Pandey G, et al. A novel solution to cloud data security issues. 2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN); 2020
 - 22 Government of Ireland. *Ireland's second one health national action plan on antimicrobial resistance 2021-2025*. 2021.
 - 23 Garzon-Orjuela N, Vornhagen H, Stasiewicz K, et al. Interactive dashboard development for Irish general practices to visualise and compare patient data. *Eur J Public Health* 2023;33.
 - 24 Kozielski S, Mrozek D, Kasprowski P, et al. E-LT concept in a light of new features of oracle data integrator 12c based on data migration within a hospital information system. In: *Beyond databases, architectures and structures*. Cham: Springer International Publishing, 2015.
 - 25 Ahmad K, Itmin SZS. The success factors of extract-load-transform process in data integration implementation. *JISTM* 2022;7:243–56.
 - 26 The General Data Protection Regulation (GDPR). What is GDPR, the EU's new data protection law? 2023. Available: <https://gdpr.eu/what-is-gdpr/>
 - 27 Health Quality Ontario. MyPractice: primary care. 2022. Available: <https://www.hqontario.ca/Quality-Improvement/Practice-Reports/Primary-Care>
 - 28 The University of Manchester. Building rapid interventions to reduce antibiotic resistance (BRIT). 2022. Available: <https://www.britanalytics.uk/>
 - 29 Palin V, Tempest E, Mistry C, et al. Developing the infrastructure to support the optimisation of antibiotic prescribing using the learning healthcare system to improve healthcare services in the provision of primary care in England. *BMJ Health Care Inform* 2020;27:e100147.
 - 30 Gotz D, Borland D. Data-Driven Healthcare: Challenges and Opportunities for Interactive Visualization. *IEEE Comput Graph Appl* 2016;36:90–6.