Analysis of article exposing databases content and their ability to answer the three following questions:

1- Does data of the CDW match the ones used by the MD ?
2- How much data will be availabe ?
3- How data were generated ?

Databases	Reference	Description	Positive Aspects	Negative Aspects
MIMIC-IV	Johnson, A. E., Bulgarelli, L., Shen, L., Gayles, A., Shammout, A., Horng, S., ... \& Mark, R. G. (2023). MIMIC-IV, a freely accessible electronic health record dataset. Scientific data, 10(1), 1.	MIMIC-IV emerges from a partnership between Beth Israel Deaconess Medical Center (BIDMC) and the Massachusetts Institute of Technology (MIT). Data from BIDMC encompassing patients admitted to the emergency department or an ICU from 2008 to 2019 are collected, including information from the ICU's MetaVision system and external sources ("acquisition" from CDW).	The data within the database is meticulously structured based on OMOP standard and comprehensively described both in the article and on their online website. It allow us to primarily address the generated data, as well as its alignment with the MD target utility.	However, to explore cohort demographics and metadata, you would need to download the database. Additionally, it only pertains to ICU data, with no information available on medical devices.
eICU	Pollard, T. J., Johnson, A. E., Raffa, J. D., Celi, L. A., Mark, R. G., \& Badawi, O. (2018). The eICU Collaborative Research Database, a freely available multi-center database for critical care research. Scientific data, 5(1), 1-13.	The eICU Collaborative Research Database is a vast ICU dataset from multiple centers, covering over 200,000 admissions across eICU Programs in the United States. This anonymized repository contains extensive data including vital signs, care plans, severity of illness assessments, diagnoses, treatments, and more.	Given that the eICU database builds upon MIMIC and encompasses numerous ICU hospitals, the cohort should be more comprehensive. The metadata structure is clearly defined within the article and appears to be beneficial for addressing questions regarding data generation and potential MD positioning.	To examine cohort demographics, downloading the database is necessary. Additionally, it solely covers ICU data, lacking information on medical devices.

AmsterdamUMCcd	Thoral, P. J., Peppink, J. M., Driessen, R. H., Sijbrands, E. J., Kompanje, E. J., Kaplan, L., ... \& Elbers, P. W. (2021). Sharing ICU patient data responsibly under the society of critical care medicine/European society of intensive care medicine joint data science collaboration: the Amsterdam university medical centers database (AmsterdamUMCdb) example. Critical care medicine, 49(6), e563-e577.	The department sourcing AmsterdamUMCdb is a mixed surgical-medical tertiary referral center for critical care medicine at an academic medical center in Amsterdam. AmsterdamUMCdb contains approximately 1 billion clinical data points from 23,106 admissions of 20,109 patients. Data from multiple clinical information systems were combined and includes : demographics, adminssion, diagnosis, vitalsigns, lab values, drugs administration, procedures, photos, notes, reports.	After downloading, the database appears to be sufficiently structured to facilitate data extraction and answer questions about the quantity of available data and their generation process.	There is limited information available on data format, as the provided link is currently inaccessible. This lack of access hinders the ability to address any questions effectively. Similarly, this database encounters the same issue regarding understanding demographic cohorts.
HiRID	Faltys, M., Zimmermann, M., Lyu, X., Hüser, M., Hyland, S., Rätsch, G., \& Merz, T. (2020). HiRID, a high time-resolution ICU dataset (version 1.0). PhysioNet. 2020. DOI: https://doi. org/10.13026/hz5m-md48.	HiRID is a freely accessible critical care dataset containing data relating to almost 34 thousand patient admissions to the Department of Intensive Care Medicine of the Bern University Hospital, Switzerland (ICU). It contains demographic data, measurements from bedside monitoring, observations by health care providers e.g.: GCS, RASS, urine and other fluid output, lab values and dministered drugs, fluids and nutrition.	The structuring of the catalog appears to be as comprehensive as the other databases listed above. Answering questions regarding the quantity of available data and their generation process should be feasible without excessive precision.	To analyze cohort demographics, downloading the database is required. Furthermore, it exclusively encompasses ICU data, with no information provided on medical devices.

PIC	Zeng, X., Yu, G., Lu, Y., Tan, L., Wu, X., Shi, S., ... \& Li, H. (2020). PIC, a paediatric-specific intensive care database. Scientific data, 7(1), 14.	PIC (Paediatric Intensive Care) is a large paediatric-specific, single-centre, bilingual database comprising information relating to children admitted to critical care units at a large children's hospital in China. The database is deidentified and includes vital sign measurements, medications, laboratory measurements, fluid balance, diagnostic codes, length of hospital stays, survival data, and more.	As the PIC database can be seen as an extension of the MIMIC database, it shares the same characteristics, with the data focusing specifically on the pediatric domain.	To delve into cohort demographics and metadata, downloading the database is necessary, followed by subsequent analysis. Moreover, it exclusively covers pediatric ICU data and specific medical diagnoses. Notably, there is a lack of information regarding the medical devices employed.
HDH	Catalogue des données (2024)| Health Data Hub. (s. d.). https://www.health-data-hub.fr/catalogue-de-donnees	The Health Data Hub is a French platform dedicated to centralizing, analyzing, and sharing health data from various sources while ensuring data protection in compliance with regulations. It aims to facilitate medical research, innovation, and healthcare improvement. Their catalog comprises a collection of multiple studies, each with brief descriptions on the dedicated datamart.	Summary of diverse studies conducted, providing insights into general aspects such as medical domains, number of patients, and so forth.	Catalog of databases specific to particular subjects, accompanied by brief data descriptions, but without direct access. Since the creation of this database appears to be incomplete, it fails to provide the necessary answers for the catalogue.
G4	Notre catalogue de métadonnées (2024)| G4. (s. d.). https://www.gcs-g4.fr/notre-catalogue-de-metadonnees	G4 University Hospitals, encompassing Amiens, Lille, Rouen, and Caen, fosters a collaborative vision among these four university hospital centers. The goal is to leverage each institution's strengths and expertise, along with their academic partners, for mutual benefit. Their "metadata" catalog is list of databases available across the four institutions, containing data usable for research purposes, including data from Electronic Health Records (EHRs) as well as registries, cohorts, clinical-biological databases, and more. It gather information such as data type, coordinating organization, scientific lead, medical domain, brief description, start and end dates, number of patients, etc.	Overview of various studies conducted, offering details on general aspects (stakeholders involved, medical domain, number of patients, etc.).	Catalog of databases specific to particular subjects, accompanied by brief data descriptions, but without direct access. It fails to address any of the questions essential for the catalogue.

Data visualization tool MIMIC	Lee, J., Ribey, E., \& Wallace, J. R. (2015). A web-based data visualization tool for the MIMIC-II database. BMC medical informatics and decision making, 16, 1-8.	The tool provides two primary functions: Explore and Compare. With the Explore feature, users can choose a patient group within MIMIC-II and view the distributions of different administrative, demographic, and clinical variables within that group. The Compare feature allows users to select two patient groups and visually analyze the differences between them across various variables. This tool proves valuable to seasoned MIMIC-II researchers by streamlining the laborious process of writing SQL queries and manually visualizing extracted data, thereby significantly expediting their research efforts.	The tool features an Explore option, which enables users to select a patient cohort based on factors such as admission ICU service type, gender, age, and diagnosis. This functionality aligns with the objectives of the catalog. Additionally, it allows for visualization of the selected cohorts and potentially facilitates comparison. Thus, it offers a preview of the quantity of available data.	However, it does not allow for visualizing the data journey or the data source.
Interactive data visualization tool to assess the impact of decision support on clinical operations	Huber, T. C., Krishnaraj, A., ... , \& Gaskin, C. M. (2018). Developing an interactive data visualization tool to assess the impact of decision support on clinical operations. Journal of digital imaging, 31, 640-645.	Radiology practices nationwide are adopting clinical decision support (CDS) software. This software aids referring providers with imaging study decisions at the point of order entry, generating a significant volume of data ripe for research and quality improvement. To better understand and analyze trends within this data, an interactive data visualization dashboard was developed using a commercially available platform. By integrating a clinical decision support product into the electronic health record and exporting its data to the visualization platform (Tableau, Seattle, WA), real-time visualization of CDS-generated data became possible. This dashboard enhances the analysis of CDS output, facilitating hypothesis generation and streamlining research and quality improvement endeavors. Integrating data visualization tools with clinical decision support systems simplifies data analysis and enhances the efficiency of research and quality improvement processes.	This article describes a tool very similar to what we would like to develop, but without the online aspect. It give some details on the journey of the data and it enables high-level visualization of a set of parameters relevant to the tool's objective.	This tool lacks filtering capabilities, despite having elements to modify data display on the dashboard. However, it doesn't address our specific question due to its divergent objectives.

Data visualization tool Ophtalmology	Kortüm, K. U., Müller, M., ... \& Hirneiss, C. (2017). Using electronic health records to build an ophthalmologic data warehouse and visualize patients' data. American journal of ophthalmology, 178, 84-93.	An academic ophthalmologic center established a near-real-time data warehouse (DW) to leverage the growing digital data from electronic medical records (EMR) and diagnostic devices. They integrated specific macular clinic interfaces into the hospital information system, enabling seamless ordering of imaging modalities. The DW, powered by an SQL database, compiled data from over 325,767 patients since 2002 and included a data discovery tool. Notably, a search for patients with age-related macular degeneration who underwent cataract surgery and received at least 10 intravitreal injections yielded 450 patients meeting the criteria. A web-based browsing tool facilitated data visualization and filtering based on various criteria, streamlining analysis and enhancing insights into ophthalmologic data.	This tool aligns with our vision for developing our catalog from the clinical data center's database. It enables targeting patient cohorts, diagnoses, and procedures. Additionally, a dashboard-style visualization appears to be available. It seems that the origin of the images (imagers) is also provided.	This tool facilitates the extraction of a significant amount of data but is limited to the field of ophthalmology. Moreover, its objectives differ as it aims to simplify the identification of eligible patients for studies.