Digital Health and Informatics Innovations for Sustainable Health Care Systems
J. Mantas et al. (Eds.)

doi:10.3233/SHTI240595

Implementation of an Atrioventricular Valve Intervention Registry: Comparative Study of REDCap vs. CDR-Based openEHR Registry

Benjamin KINAST [a,b,1], Henrik ROHDE [a,b], Michael ANYWAR [a,b], Tobias, BRONSCH [a,b], Jakob VORAN [c,d], Felix KREIDL [c], Derk FRANK [c,d] and Björn SCHREIWEIS [a,b]

[a] *Institute for Medical Informatics and Statistics, Kiel University and University Hospital Schleswig-Holstein, Campus Kiel, Germany*

[b] *Medical Data Integration Center, University Hospital Schleswig-Holstein, Germany*

[c] *Department of Internal Medicine III, Cardiology and critical care, University Hospital Schleswig-Holstein, Campus Kiel, Kiel, Germany*

[d] *German Centre for Cardiovascular Research, partner site Hamburg/Kiel/Lübeck, Kiel, Germany*

ORCiD ID: Benjamin Kinast https://orcid.org/0000-0003-2554-4381, Henrik Rohde https://orcid.org/0009-0004-8021-0151, Michael Anywar https://orcid.org/0000-0001-8028-803X, Tobias Bronsch https://orcid.org/0000-0002-3622-0625 Jakob Christoph Voran https://orcid.org/0009-0000-4547-9260 , Felix Kreidl https://orcid.org/0000-0003-0884-6572, Derk Frank https://orcid.org/0000-0001-7561-075X, Björn Schreiweis https://orcid.org/0000-0002-1748-1563

Abstract. This comparative study examines the transition from isolated registries to a consolidated data-centric approach at University Hospital Schleswig-Holstein, focusing on migrating the Atrioventricular Valve Intervention Registry (AVIR) from REDCap to a Medical Data Integration Center based openEHR registry. Through qualitative analysis, we identify key disparities and strategic decisions guiding this transition. While REDCap has historical utility, its limitations in automated data integration and traceability highlight the advantages of a data-centric approach, which include streamlined data (integration) management at a single-point-of-truth based on e.g., centralized consent management. Our findings lay the groundwork for the AVIR project and a proof-of-concept data-centric registry, reflecting a broader industry trend towards data-centric healthcare initiatives.

Keywords. Clinical registries, openEHR, Data integration, Requirements, Clinical data management, electronic data capture

[1]Corresponding Author: Benjamin Kinast, Institute for Medical Informatics and Statistics, Kiel University and University Hospital Schleswig-Holstein, Kiel, Germany; E-mail: benjamin.kinast@uksh.de.

1. Introduction

The University Hospital Schleswig-Holstein (UKSH) operates a Medical Data Integration Center (MeDIC), systematically structuring clinical routine and research data [1,2]. These data are stored in the platform's central openEHR clinical data repository (CDR) for further use in both the patient care and the research context. In addition to this incrementally growing central platform, a historically grown ecosystem of isolated study databases and registries thrives across various institutes and departments within UKSH. Those registries and databases are data silos which contain extensive amounts of valuable subject-specific data meticulously entered with significant effort [3]. Due to the lack of integration with other UKSH IT systems, the possibilities regarding data access, linkage and analysis are severely limited. UKSH is on a transformative journey, steering away from stand-alone research silos towards a consolidated data-centric approach. As part of this paradigm shift, we are currently in the decision making process of migrating a Atrioventricular Valve Intervention Registry (AVIR) from an isolated REDCap system [4] to our MeDIC. In this paper, we demonstrate the requirements behind this transition, using the AVIR as an example.

2. Methods

In this study, we qualitatively compare the requirements of our envisioned next-generation AVIR with the current state of the Electronic Data Capture (EDC) system REDCap in our setting qualitatively. To strategically navigate UKSH's paradigm shift towards a data-centric approach, we adopted the V-Model framework (see **Figure 1**), a well-established system development and validation methodology closely mirroring the classic Software Development Life Cycle [5]. Initiating the *Project Definition* phase within this framework, we engaged in extensive discussions with key stakeholders from the clinic of cardiology. Subsequently, the identified requirements underwent a rigorous refinement process during *Requirements Analysis*. Recognizing the unique requirements of clinical registries, we further tailored these requirements using an existing requirements catalog [6], placing emphasis on criteria specifically aligned with our data-centric approach.

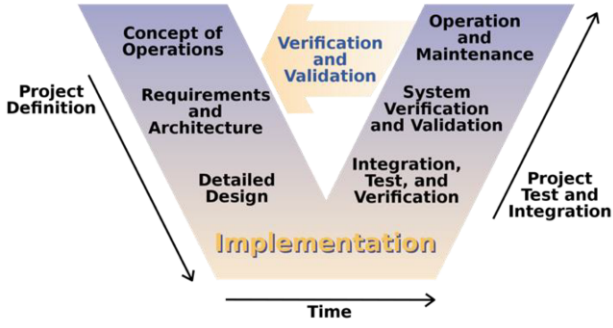

Figure 1. V-Model of Systems Engineering Process by L. Osborn et al. [7].

We compared these refined requirements with the functionalities of the REDCap EDC system and the MeDIC's data-centric CDR approach. This comparative analysis aims to lay the foundation of a detailed and strategic design. We excluded general

requirements for basic functions, such as GCP-conformance, user-interface provision, user authentication, logging, or scalability during this customization process. As per requirements engineering theory, these subconscious functions are considered *Dissatisfiers* [8]. The exclusion aimed to streamline the requirements elicitation process, focusing solely on essential needs for the envisioned next-generation AVIR:

R1: The registry must (re-)use data from different heterogeneous source systems/ clinical routine IT systems.
R2: The system must acquire data from cross-institutional sources (multi-center setting).
R3: The registry must provide customizable reporting capabilities.
R4: The registry must provide traceability to the original data.
R5: The registry must process data based on interoperability standards.
R6: The registry must provide updateable data integration if at least one of the following changes: * source data, * data sources, * clinical context, * analysis requirements.
R7: The registry must only allow for the analysis of (patient-) data for which consent has been obtained.

3. Results

The comparative analysis between the REDCap Registry and a MeDIC-based Registry highlights key differences with implications for the development of a next-generation AVIR. REDCap, primarily designed for manual data entry, is widely used in various clinical research settings. Despite its popularity, manual entry is time-consuming and error-prone. To address this, REDCap offers a Restful API for automated data integration, including importing medical record information. It provides Clinical Data Interoperability Services, such as Clinical Data Pull and Clinical Data Mart, allowing real-time or bulk data import from Electronic Health Records (EHRs). The MeDIC adopts an automated approach facilitated by Extract, Transform, Load (ETL) pipelines linked to clinical routine IT systems. In the MeDIC each automatically filled datapoint will contain a link to the source, which would be a composition in the care context. Manually entered datapoints would be flagged as such. This allows automatic correction in case data in the care context are corrected. This does refer to values which have been newly acquired in the care context (i.e., a new body-weight, diagnosis billed). Technically those two scenarios can be distinguished due to the difference between updating an existing data point and creating a new one. The MeDIC -based registry puts a strong emphasis on re-use of established data from other scenarios (R1) [9]. This means it seamlessly integrates data processed in content-specific ETL routines, allowing for the reuse of existing routines. Beyond the cross-institutional data, registries often depend on multiple sites to contribute to a sufficiently large cohort. For acquiring data from cross-institutional sources in a multi-center setting, the REDCap registry allows to connect projects across institutions. However, currently, this capability is underutilized as the AVIR projects are not connected, and external sites provide manual exports, which again must be merged with UKSH AVIR data manually. The MeDIC offers external accessibility without necessitating a separate installation and data storage, offering a common frontend to the eCRFs via its interface layer which is also true for REDCap. Furthermore, the MeDIC integration is coupled with a centralized data repository, ensuring streamlined management and accessibility (R2). This unified access ensures that all data are integrated and stored in a central system, facilitating streamlined management and accessibility. Moreover, external data are required to be labeled to

maintain transparency and traceability throughout the registry. Regarding customizable reporting capabilities, the REDCap registry enables users to generate customizable reports based on its data dictionary. In contrast, the CDR-based registry not only offers customizable reporting but also allows the blending or supplementation of registry data with the entirety of data available in the CDR when exporting data (R3). Another important capability of registry management is data traceability and transparency which differs between the two systems. As stated, the REDCap registry involves manual data transfer from e.g., the EMR, with the drawback of not automatically providing metadata on the origin of data (R4). While REDCap generally provided the capability to handle metadata, the manual addition process is labor-intense. Conversely, the MeDIC addresses this by incorporating provenance data throughout the ETL process. In terms of processing data based on interoperability standards, REDCap adopts FHIR as an exchange standard for data exchange for EHR integration from selected EHR systems via a FHIR Adapter [10]. The MeDIC-based registry adheres to the openEHR community standard integrating terminologies and allowing for SNOMED Expression Constraint Language (R5). Addressing the changing nature of medical data and its impact on integrating e.g., changes in source data, updates in the REDcap Registry rely on manual processes (R6). In contrast, the MeDIC automates updates through real-time ETL processes, ensuring seamless data updates to adapt to changes in source data, clinical context, or evolving analysis requirements (R6).

Finally, the REDCap registry relies on manual processes for consent management. In contrast, the MeDIC seamlessly integrates with a dedicated consent management system at UKSH [1]. This system receives and manages consent information from the EMR, reducing the workload associated with consent management (R7).

4. Discussion

The comparative study shows disparities in core requirements, establishing a foundation for the decision-making process. While REDCap has demonstrated historical efficiency in the scientific community, it only partially aligns with the identified stakeholder requirements. Despite the technical feasibility of automated data integration into REDCap, the sustainable advantages offered by integration into a MeDIC-based registry are more substantial. It allows for versatile reuse in different contexts, such as other registries, patient care, and AI scenarios. It's imperative to recognize the potential noise in EMR data, where crucial information may reside in free text, such as notes or discharge letters [11]. Consequently, occasional manual data entry may be necessary in the MeDIC-based registry. Furthermore, not all relevant information may be in the EMR, requiring integration of data from subsystems, with already productive ETL pipelines, like laboratory or medication data, offering viable solution. Concerning data integration, the likelihood of manual data entry at external study centers remains high, but the advantage lies in the ability of study centers to use a webservice of the MeDIC for direct registry data input, eliminating the need for a separate installation and manual exports. By embracing a shift to a data-centric approach, we align with an industry trend observed in initiatives like the next-generation study database [12]. They pursued a similar path, collecting requirements for a study database connected to their EMR [12]. However, while setting up a REDCap project may be accomplished swiftly, transitioning to a data-centric platform demands proper strategic planning and implementation efforts. We started our transition of developing, refining, and mapping heterogeneous data to

openEHR back in 2020, experiencing a steep learning curve. Finally, this comparative study forms the foundation for the AVIR project and the implementation of a proof-of-concept MeDIC-based registry. Following the pilot phase, the next step involves exploring the potential reuse of AVIR's CRFs for other cardiological registries as a long-term perspective. We need to acknowledge the study's focus on local conditions at UKSH, making it not universally representative.

5. Conclusions

Today, automatic EMR data integration into clinical registries is uncommon. Our evaluation shows long-term advantages of the data-centric approach for an AV registry, particularly in terms of sustainable utilization. The requirements support the decision process for the registry architecture. Data integrated into the AVIR and the existing ETL pipelines can be reused for additional registries. If commonalities are identified at the data level between the registries, a core dataset for cardiological registries will be derived.

References

[1] Haarbrandt B, Schreiweis B, et al. HiGHmed–an open platform approach to enhance care and research across institutional boundaries. Methods of information in medicine. 2018 May;57(S 01):e66-81. doi:10.3414/ME18-02-0002.

[2] Kock-Schoppenhauer AK, et al. Medical Data Engineering–Theory and Practice. InAdvances in Model and Data Engineering in the Digitalization Era: MEDI 2021 International Workshops: DETECT, SIAS, CSMML, BIOC, HEDA, Tallinn, Estonia, June 21–23, 2021, Proceedings 10 2021 (pp. 269-284). Springer International Publishing. doi:10.1007/978-3-030-87657-9_21.

[3] Pronker E, Geerts BF, Cohen A, Pieterse H. Improving the quality of drug research or simply increasing its cost? An evidence-based study of the cost for data monitoring in clinical trials. British journal of clinical pharmacology. 2011 Mar;71(3):467-70. doi:10.1111/j.1365-2125.2010.03839.x.

[4] Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, McLeod L, Delacqua G, Delacqua F, Kirby J, Duda SN. The REDCap consortium: building an international community of software platform partners. Journal of biomedical informatics. 2019 Jul 1;95:103208. doi:10.1016/j.jbi.2019.103208.

[5] Müller-Ettrich G, System Development with V-Model and UML, in: M. Schader, and A. Korthaus (Eds.), The Unified Modeling Language, Physica-Verlag HD, Heidelberg, 1998: pp. 238–249. doi:10.1007/978-3-642-48673-9_16.

[6] Kinast B, Ulrich H, Bergh B, and Schreiweis B, Functional Requirements for Medical Data Integration into Knowledge Management Environments: Requirements Elicitation Approach Based on Systematic Literature Analysis, Journal of Medical Internet Research. 2023;25:e41344. doi:10.2196/41344.

[7] Osborn LF, Brummond J, Hart R, et al., Clarus: Concept of Operations, Clarus: Concept of Operations. (2005). https://rosap.ntl.bts.gov/view/dot/3710 (accessed March 14, 2024).

[8] Frühauf K, Fuchs E, Glinz M, et al., IREB Certified Professional for Requirements Engineering, Foundation Level. (2020). https://www.ireb.org/en/downloads/#cpre-foundation-level-syllabus-3-0.

[9] Sommer KK, Amr A, Bavendiek U, et al., Structured, Harmonized, and Interoperable Integration of Clinical Routine Data to Compute Heart Failure Risk Scores, Life. 2022;12:749. doi:10.3390/life12050749.

[10] Metke-Jimenez A, and Hansen D, FHIRCap: Transforming REDCap forms into FHIR resources, AMIA Joint Summits on Translational Science Proceedings AMIA Joint Summits on Translational Science. 2019;54–63. doi:PMID: 31258956.

[11] Assale M, Dui LG, Cina A, et al., The Revival of the Notes Field: Leveraging the Unstructured Content in Electronic Health Records, Front. Med. 2019;6:66. doi:10.3389/fmed.2019.00066.

[12] Dugas M, Blumenstock M, Dittrich T, et al., Next-generation study databases require FAIR, EHR-integrated, and scalable Electronic Data Capture for medical documentation and decision support, Npj Digit. Med. 2024;7:1–8. doi:10.1038/s41746-023-00994-6.