A national evaluation analysis and expert interview study of real-world data sources for research and healthcare decision-making

Veronika Mikl
Gesellschaft für Pharmazeutische Medizin e.V. (GPMed)
Dejan Baltic
Gesellschaft für Pharmazeutische Medizin e.V. (GPMed)
Thomas Czypionka
Institut für Höhere Studien – Institute for Advanced Studies (IHS)
Alexander Degelsegger-Márquez
Gesundheit Österreich GmbH (GÖG)
Nikolaus Forgó
University of Vienna
Ghazaleh Gouya-Lechner
Gesellschaft für Pharmazeutische Medizin e.V. (GPMed)
Arnold Herzog
Austrian Medicines and Medical Devices Agency (AGES Medizinmarktaufsicht)
Peter Klimek
Supply Chain Intelligence Institute Austria (ASCII)
David Benjamin Lumenta
Medical University of Graz
Bernhard Mraz
Gesellschaft für Pharmazeutische Medizin e.V. (GPMed)
Herwig Ostermann
Gesundheit Österreich GmbH (GÖG)
Robert Scharinger
Federal Ministry of Social Affairs, Health, Care and Consumer Protection
Tanja Stamm
Gesellschaft für Pharmazeutische Medizin e.V. (GPMed)
Michael Strassnig
Wiener Wissenschafts-, Forschungs- und Technologiefonds (Vienna Science and Technology Fund)
Markus Zeitlinger
Gesellschaft für Pharmazeutische Medizin e.V. (GPMed)
Johannes Pleiner-Duxneuner (johannes.pleiner@gpmed.at)

Gesellschaft für Pharmazeutische Medizin e.V. (GPMed)

Article

Keywords: real-world data, real-world evidence, data quality, data quality criteria, health data use, secondary use of health data, health data strategy, FAIR data principles, data quality recommendations, pharmaceutical research, healthcare decision-making, quality criteria for RWD in health care, Gesellschaft für Pharmazeutische Medizin, GPMed

Posted Date: October 17th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3423573/v1

Additional Declarations: Competing interest reported. All authors declare no financial support or funding for this project. DB is an employee of Amgen GmbH, Vienna, Austria. BM is an employee of Novartis Pharma GmbH, Vienna, Austria. VM. and JPD are employees of Roche Austria GmbH, Vienna, Austria. All other authors declare no other conflicts of interest.

Abstract

Real-world data (RWD) can provide intel (real-world evidence, RWE) for research and development, as well as policy and regulatory decision-making along the full spectrum of health care. Despite calls from global regulators for international collaborations to integrate RWE into regulatory decision-making and to bridge knowledge gaps, some challenges remain. In this work, we performed an evaluation of Austrian RWD sources using a multilateral query approach, crosschecked against previously published RWD criteria and conducted direct interviews with representative RWD source samples. This article provides an overview of 73 out of 104 RWD sources in a national legislative setting with favourable RWD incentives, which can be used to extrapolate to other EU data regions under the General Data Protection Regulation (GDPR) and upcoming legislation such as the European Health Data Space Act (EHDS). We were able to detect omnipresent challenges associated with data silos, variable standardisation efforts and governance issues. Our findings suggest a strong need for a national health data strategy and governance framework, which should inform researchers, as well as policy- and decision-makers to improve RWD-based research in the healthcare sector to ultimately support actual regulatory decision-making and provide strategic information for governmental health data policies.

Introduction

Real-world data (RWD) generates evidence for various research, development, policy and regulatory decision-making purposes along the product lifecycles of pharmaceuticals and medical devices. The increasing use[1,2,3,4] of RWD also provides significant opportunities beyond the aforementioned across the full spectrum of health care, ranging from clinical trial design to the study of medical (mal-)practice[5], to public health and health policy[6]. To account for the transformative potential of RWD, the European Union has recently passed in addition to existing legislations such as the GDPR, the European Data Governance Act[7]. Furthermore the European Commission proposed a regulation for the European Health Data Space (EHDS)[9] to facilitate, among other aims, the safe and secure use and re-use of health data for better healthcare delivery, research and policy-making. Yet, progress in the digitalisation of health care systems is unevenly distributed across Europe[9], casting doubts on achieving the ambitious aims of the EHDS.

Despite calls from global regulators for international collaboration to integrate real-world evidence (RWE) into regulatory decision-making[10] and to bridge knowledge gaps, some challenges like heterogeneity of data sources, linkability / sharing of data, variable quality of data and differing approaches for data access require more and appropriate attention.

Research objectives

In this work, a multi-stakeholder group coordinated by the Gesellschaft für Pharmazeutische Medizin (GPMed, Austrian Society for Pharmaceutical Medicine) compiled and classified already used national RWD sources in Austria and made an in-depth assessment of the research readiness of selected datasets. The group reviewed a previously proposed quality checklist for RWD in pharmaceutical research

and regulatory decision-making[11]. Results and findings intend to emphasise the relevance of RWD and to inform researchers, health care regulators as well as decision-makers and strategic governmental health data policy working groups on national and international level about their availability and currently identified limitations. The objectives are in detail:

- to provide an initial overview of available Austrian healthcare RWD sources and their research and decision-making readiness, data locations and data custodians,
- to analyse and improve the already published checklist[11] (Table 5),
- to discuss and conclude which data quality aspects should be applied to improve the use of RWD for scientific and regulatory purposes.

Results and Discussion

We identified 73 out of 104 RWD sources, which met the defined criteria and objectives (Table 1a and Table 1b). 31 out of 104 RWD sources mentioned in publications were not findable or accessible online any more (Table 2). Table 3 provides a matrix indicating the main purpose of the RWD and the type of institutions, which hold and manage the data.

Table 2
List of unverifiable lost RWD sources.

ID	Name of RWD	Description and further information of de-selection
74	AGMT Neck Cancer Registry	Registry not findable on https://www.agmt.at/register/ any longer
75	Bone and soft tissue tumor registry	Other than the entry on orpha, there is no evidence that the register exists http://www.orpha.net/data/prj/AT/ID127487ger.pdf;
76	CEDATA-GPGE® Registry	Registry held and managed in Germany
77	ECFS Patient Registry	Registry held and managed not in Austria
78	European LeukemiaNet	Registry held and managed not in Austria
79	European Myelodysplastic Syndromes (EUMDS) Registry	Registry held and managed not in Austria
80	European Registry for Endocrine Surgery (Eurocrine®)	No Austrian data captured, Eurocrine is registered as a not for profit organisation organized and duly registered under the laws of Austria for societies. The owner of the platform is Region Skåne, the County Council of Scania Region in Sweden.
81	NF-10 - Prospective collection of potentially prognostically relevant data in patients with indolent non-follicular B-cell lymphoma	Not verifiable via 2nd source
82	Austrian registry for BRCA-1 and BRCA-2 mutation	Not verifiable via 2nd source
83	Medical claims database Austrian Ministry of Health	unspecific information about the real data source
84	Österreichisches Gesundheitsinformationssystem	No RWD for research purposes, only for monitoring and reporting.
85	Paediatric Congenital adrenal hyperplasia (CAH) registry	Registry held and managed not in Austria
86	REGIS – Regionales Gesundheitsinformationssystem	No RWD for research purposes, only for monitoring and reporting.
87	NSCLC Stadium III - Zentrales Datenregister für das Management von Patienten mit nichtkleinzelligem Lungenkrebs in Stadium III	unspecific information about the real data source. No further information is available apart from information on submission to the Ethics Committee of the Medical University of Vienna.
88	Österreichisches Register für fortgeschrittenes Prostata Karzinom	Not findable via online search

ID	Name of RWD	Description and further information of de-selection
89	Peritonealkarzinose- Register der Medizinischen Universität Wien	Not findable via online search
90	Prospektives Register, mit Errichtung einer Biobank, und Genanalysen von Patienten mit Pulmonaler Hypertension	Not findable via online search
91	Register über die Behandlung von PatientInnen mit hirneigenen Tumoren an der KIM1	Not findable via online search
92	AGO R01 Breast Cancer in Pregnancy Register Study (BCP) Registerstudie	Registry study, unclear whether own real-world data source
93	AGO R03 - ROC Register to Describe the Treatment Pattern of Platinum-sensitive Relapsed Epithelial Ovarian Cancer Patients in Austria	Registry study, unclear whether own real-world data source
94	AGO R05 AXillary Surgery After NeoAdjuvant Treatment (AXSANA)	Registry study, unclear whether own real-world data source
95	Observational study of pediatric thrombotic disease: the Throm-PED registry	Observational study, out of scope of RWD definition described in methods
96	Registry study in NSCLC patients with EGFR, ALK, or ROS1 mutations	Registry study, unclear whether own real-world data source
97	Covid-19 Datenplattform	No own data source
98	Styrian registry of congenital anomalies - contributes to the EUROCAT network	European Registry with Austrian Data https://eu-rd-platform.jrc.ec.europa.eu/sites/default/files/eurocat-pub-docs/JA%20EUROCAT%20Final%20Report.pdf
99	Non-Interventional, web-based Registry for Histiocytic Disorders	Such a register is not specified on the OeGHO homepage.
100	Akut PTCA (Perkutane Transluminale Coronare Angioplasie) Register der ÖKG	Website can no longer be accessed https://ptca.i-med.ac.at/
101	Registry of the NHL-BFM study group for all subtypes of Non-Hodgkin lymphoma in children and adolescents	Not findable via online search
102	Registry for relapsing acute lymphoblastic leukemia in childhood and adolescence	Not findable via online search

ID	Name of RWD	Description and further information of de-selection
103	Registry for Philadelphia chromosome-positive acute lymphoblastic leukemia in childhood and adolescence	Not findable via online search
104	Austrian Breast Implant Registry	2009–2022: discontinued by founding professional society (Österreichische Gesellschaft für Plastische, Ästhetische und Rekonstruktive Chirurgie (ÖGPÄRC)

Table 3
RWD main purpose and type of data holder matrix.

	Admini-strative	Clinical	Epidemi-ological	Quality assurance	Regulatory	Research	Total
Expert community		7	6	3		1	**17**
Government Organisation	6	2	7	6	6		**27**
Hospital (Association)		2	3	4			**9**
Other				1			**1**
Professional Society		1	5	4			**10**
Social Insurance Institution				1			**1**
University		3	3			2	**8**
Total	**6**	**15**	**24**	**19**	**6**	**3**	**73**

We identified 30 different organisations holding and managing RWD sources (Table 4), which we further grouped in seven institutional types of RWD holders (Fig. 1). Expert communities and professional academic societies owned 27 verified RWD sources in Austria. All Austrian medical universities hold at least one RWD source. For the Austrian governmental organisations, all of the main institutions appeared as data holders (e.g. Federal Ministry of Social Affairs, Health, Care and Consumer Protection (BMSGPK), Federal Office for Safety in Health Care / Austrian Medicines and Medical Devices Agency (BASG/AGES), Austrian National Public Health Institute (GÖG), etc.) and this group holds 27 RWD sources. The Austrian social insurance is one of the key holders of a major RWD source. The selected interview-sample reflects the overall distribution of institutional types of RWD holders as shown in Fig. 2.

Table 4
Identified RWD holder.

No.	Name
1	Agentur für Gesundheit und Ernährungssicherheit (AGES)
2	AGO Austria Arbeitsgemeinschaft für Gynäkologische Onkologie der OEGGG
3	Arbeitsgemeinschaft medikamentöse Tumortherapie (AGMT)
4	Arbeitskreis für Vorsorge- und Sozialmedizin
5	Austrian Mesothelioma Interest Group
6	BBMRI.at
7	Bundesamt für Sicherheit im Gesundheitswesen
8	Bundesministerium für Soziales, Gesundheit, Pflege und Konsumentenschutz
9	ELGA GmbH
10	Gesundheit Österreich GmbH
11	Klinikum Klagenfurt
12	LIV - Landesinstitut für Integrierte Versorgung Tirol
13	Medizinische Universität Graz
14	Medizinische Universität Innsbruck
15	Medizinische Universität Wien
16	Niederösterreichischen Landeskliniken-Holding
17	Österreichische Arbeitsgemeinschaft für Urogynäkologie & Rekonstruktive Beckenbodenchirurgie
18	Österreichische Gesellschaft für Chirurgie - OEGCH
19	Österreichische Gesellschaft für Hämatologie & Medizinische Onkologie
20	Österreichische Gesellschaft für Neurologie
21	Österreichische Hämophilie Gesellschaft
22	Österreichische Parkinsongesellschaft
23	Österreichische Sozialversicherung
24	Österreichischen Gesellschaft für Nephrologie
25	Österreichischen Kardiologischen Gesellschaft
26	Österreichischen Kardiologischen Gesellschaft Österreichische Gesellschaft für Thorax- und Herzchirurgie (ÖGTHC)

No.	Name
27	Statistik Austria
28	Trägerverein für das Österreichische Register für Biologica, Biosimilars und tsDMARDs bei der Behandlung von entzündlichen rheumatischen Erkrankungen (BioReg)
29	Tumorzentrum Oberösterreich
30	Uniklinikum Salzburg

The majority of identified and verified RWD sources are registries followed by administrative data collections, biobanks, health care databases or observational collections. 39 RWD sources belonged to the category “disease registry” (Fig. 3). The distribution of the main purpose mainly follows a functional differentiation: governmental organisations and social insurance carriers hold RWD sources with an administrative and quality assurance purpose. Governmental organisations are also central for RWD with an epidemiological as well as regulatory purpose (Fig. 4). Medical universities as well as professional organisations often run clinical RWD sources. More strikingly, there are only a few RWD sources whose main purpose lies in research (beyond clinical questions). The selected interview-sample of 11 RWD sources corresponds well to the general overall picture for what purposes RWD are collected (Fig. 5).

Disease or topic wise, cancer diseases dominated the field of RWD sources (26 out of the 73) in the clinical and/or epidemiological domain (Fig. 6), as cardiovascular disease RWD sources did in quality assurance. Due to the strict regulation of the pharmaceutical domain, a high number of RWD for regulatory, administrative and quality assurance purposes exist. Only a few remaining RWD sources focus on other specific diseases.

The 11 RWD sources examined already met many quality criteria of the checklist. The parameters "Infrastructure”, "Data Elements", "Data Provider" and “Quality requirements” were among the most commonly fulfilled criteria (Fig. 7). Of the four FAIR Data Principles, “Findable” was the least met (vs. “Accessible”, “Interoperable”, “Reusable”), suggesting that this principle requires attention from data owners. Interview results correlated well with the issues encountered during our own research, where "finding" the relevant RWD sources was subjectively perceived as cumbersome and time-consuming. The quality criterion "data privacy and transparency" produced low ratings due to the ambiguous interpretation resulting from the type of used regulations: e.g. informed consent processes and GDPR for research vs. national regulations implemented by law. The same applied to the low rating of “Research objectives”, since RWD sources set-up by law do not necessarily follow a research question or protocol like approach inherent to classic clinical research projects. This also concerned the parameter “Patient population covered” due to the heterogeneity and disease-specification not applying to the general population.

Discussion

The evaluation of the identified RWD sources used in the publications highlights a number of issues regarding the availability and accessibility of the national RWD landscape. The effort required in this work to identify RWD resources underscored the importance for providing a central directory for RWD sources aligned with DGA and EHDS requirements (e.g. data catalogues) to facilitate research with high quality data sets, which could serve as a valuable resource for all stakeholders. The time and resources required to search for and locate each of the identified RWD sources were a major obstacle to utilizing the available data sets in a more efficient manner.

Several RWD sources identified in the search process were not findable online (31 of 104 RWD sources, Table 2). It remains unclear if adequate metadata descriptions of these RWD sources were just unavailable or if they have been deleted since. This, however, puts the research integrity of these sources, notably data transparency and reproducibility into question. This highlights the importance of data holders ensuring the long-term accessibility of collected RWD, enabling their reuse for (secondary) research purposes. Without such accessibility features, the potential benefits of using RWD for research, public health policy, and society in general cannot be reached.

RWD with a dedicated research purpose used in the analysed articles were rather a national exception. Predominantly, publications on RWD data sets are characterised by the secondary use of quality assurance data or epidemiological RWD, indicating a gap in the integration of academic research into public health policy-making in Austria. This suggests that research with RWD seems to be secondary thought if at all considered following the the esablishment of such registries. The limited availability of RWD collected for research purposes hinders the potential to develop evidence-based policies and strategies that could positively impact public health outcomes in the country.

Expert communities and professional societies hold a substantial number of RWD sources. However, these organizations are often characterized by lacking adequate resources to maintain robust data management practices, e.g. up-to-date content and long-term availability. Due to missing directories, lacking online meta data descriptions and undefined rules for third party access, these RWD sources appear to be data silos or “club good” for “insiders” and cannot provide any benefit for healthcare research or policymaking.

The population of RWD data holders in Austria is quite diverse ranging from small professional societies to large public authorities. While this diversity could prove beneficial, this is also a source of the silo-ization of health data in Austria as demonstrated by the fact that barely any article in our sample used more than one data set in each publication due to legal and technical restrictions.

These findings prompt a critical discussion regarding the current state of working with or setting up RWD sources that do not adhere to FAIR data principles. It raises the question of whether such practices can still be considered state-of-the-art demonstrating a striking contrast to the initiatives on the European level as stated in the introduction. A substantial share of the RWD sources was not findable (Table 2). Accessibility was another major issue, either based on the lacking “findability”, or if findable on undefined rules for third party access. This concerns also public RWD where some institutions could use

administrative datasets based on contracts, but given the transaction costs, this impedes smaller research groups and individual researchers to use these data. Therefore, the prevalence of data silos and the lack of data interoperability and standardization[12] continues to pose challenges in this fragmented RWD landscape impeding the potential of RWD in general. The shortcomings of the RWD landscape in Austria have shown that the previously published RWD quality checklist[11] and the feedback from the interviewees were valuable resources to inform future RWD efforts to consider multifunctional use of the data in the long term. A response was: "We would have needed this checklist before we built the registry."

Furthermore, the findings of the interviews confirmed our initial assumption that research readiness for secondary purposes and broader applicability were albeit often forgotten during the inauguration of RWD sources. In the assessment of the checklist by the interviewees, registers/cohorts dedicated to specific purposes tended to receive high scores in terms of research readiness. However, their usefulness was limited due to the prevailing data silo-ization. This lack of data integration and interoperability prevents researchers from harnessing the full benefits of these "research-ready" datasets, leading to their underutilization. Interestingly, some of the most comprehensive and interesting RWD datasets obtain the lowest scores, putting their as RWE source into question. This highlights the prevailing marginal status of RWD utilization, as these valuable datasets remain underutilized and underappreciated in the research community. We also received valuable and constructive suggestions on how to further improve or adapt the criteria listed in the checklist, so that it can be used more broadly. We adapted the checklist accordingly and provided in this publication (Table 5).

Table 5

Checklist on Quality criteria for RWD revised version 2.0.

Data management and stewardship	· The "FAIR Data Principles" formulate principles that sustainable, reusable research data and research data infrastructures must meet. Definitions see here: https://www.go-fair.org/fair-principles/
Governance framework	· Available policy for collaborations with external organizations · Governance structure for decision-making on requests for collaboration · Available templates for research/data-sharing contracts · Involvement of Patient Organizations
Quality requirements	· High RWD quality standards are implemented – such as: completeness – accuracy – timeliness – comparability · Process in place for ongoing data quality assessments · Processes in place for quality planning, control, assurance and improvement · Data verification (method and frequency of verification) · Auditing practice
Data privacy & transparency	· Informed consent form and its validity for research purposes according to GDPR, EHDS and relevant national regulations.
Research objectives	*Note – Only applicable if the primary purpose of the RWD is research* · Well defined research question outlined in a research plan · Available documentation, protocol or proposal which describes purpose of RWD use and rational that the RWD data sources adequately addresses the research questions (e.g. study protocol) · Approval of RWD use of independent review board/ethics committee · Protocol should follow the Declaration of Helsinki and furthermore the Declaration of Taipei [26] on Research on Health Databases, Big Data and Biobanks should be taken into account.
Data providers	· Description of data providers, such as patients, carers or health care professionals, their geographical area and any selection process (inclusion and exclusion criteria) that may be applied for their acceptance as data providers
Patient population covered	· Description of the type of patient population (disease, condition, time period covered, procedure), which defines the criteria for patient eligibility · Relevance of setting and catchment area · Clarity on patients' inclusion and exclusion criteria · Methods applied to minimise selection bias and loss to follow-up
Data	· Definition, dictionary and format of data elements

elements	· Standards and terminologies applied · Capabilities and plans for amendments of data elements
Infrastructure	· High quality systems for RWD collection, recording and reporting, including timelines · Capability (and experience) for expedited reporting and evaluation of severe suspected adverse reactions in RWD collection · Capability (and experience) for periodic reporting of clinical outcomes – ideally patient reported outcomes - and adverse events reported by physicians, at individual-patient level and aggregated data level · Capability (and experience) for data cleaning, extraction, transformation and analysis · Capability (and experience) for data transfer to external organisations · Capabilities for amendment of safety reporting processes

Conclusion

The health data landscape changes constantly due to new data collection points, cheaper and faster availability of omics data, digital health and digital care pathways, imaging technology and artificial intelligence. This evolution creates opportunities not only for healthcare research & development, but also for public health and health policy[6]. This necessitates increased coordination, the creation of common (meta)data standards and interoperability to avoid silo-ization and to maximise the benefits of RWD through data exploration in linked datasets, which are able to represent the complexities of public and individual health issues.

However, the legislative environment is yet not ready to support RWD within the boundaries of fundamental rights. This has several reasons, not all of them being purely of legal nature. Strictly legally speaking, Austria made already a major attempt to increase access to secondary use of data via several reforms of the federal law on the organisations of research ("Forschungsorganisationsgesetz") and of the law on statistics ("Bundestatistikgesetz") in 2018[13] and in 2021[14], respectively. The aim of these reforms was in particular to increase accessibility of existing (personal) data for research purposes. However, due to several reasons, including the lack of secondary legislation on a ministerial level that would have been needed and due to legal complexity, these attempts have not yet sufficiently reached their goals. The already complex national situation faces new challenges by the planned European legislative initiatives, in particular the Data Governance Act (DGA[7]) and the European Health Data Space Act (EHDS[15]). The DGA aims to improve data sharing and data reuse within the European Union (EU) by introducing, inter alia, competent bodies (art. 7), single information points (art. 8), data intermediation services (art. 10) and public registers of recognised data altruism organisations (art. 17). The EHDS will likely introduce a whole chapter on secondary use of electronic health data (Chapter IV), introducing health data access bodies (art. 36), rules on data altruism in health (art. 40), a cross-border infrastructure

for secondary use of electronic health data (HealthData@EU) (art. 52) and new governance bodies such as the European Health Data Space Board (art. 64). Whereas these European attempts have the potential to improve the accessibility of RWD, there exists at the same time a significant risk of even more legal complexity by legal inconsistency, national deviations and unclarity as an unwanted offspring of these initiatives.

High quality criteria for RWD are key for improved data utilization in research and healthcare decision-making[4]. The herein provided improved checklist (Table 5) may also support authorities and government institutions in their attempt to ensure data quality for the whole sector, in particular with regard to the implementation of the DGA and the coming EHDS as well as national and European activities of open science. RWD sources can foster a more open culture of data sharing and reuse, which is unfortunately almost absent in the currently reviewed health data sector.

We also call for a critical, scientifically driven analysis of the regulatory environment, together with an attempt to simplify the legal landscape, and more ambitious and structured governance activities regarding health data, in particular for a more comprehensive approach to data collection, considering the potential for future research and wider utilization. Multipurpose datasets may increase efficiency and may act as boost for research on topics that are often neglected due to the lack of data. A significant improvement in data utilization could be achieved through better linking of data from both public and private sources. Our findings emphasize the creation of a comprehensive data strategy in the healthcare domain, especially in the reviewed national framework in Austria. Despite the introduction of the Digital Austria Act[16] by the Government in mid-2023 and the immediate health care reform package from the Council of Ministers[17], the current efforts fall short of establishing a robust health data strategy.

On the upside, Austria employs already sector-specific personal identifiers to link data across data sets without compromising privacy and data protection (the so-called "bereichsspezifische Personenkennzeichen (bPK)") and the recently established Austria Microdata Center (AMDC) at Statistics Austria can serve as a role model for the use of administrative and statistics data for research (legally, technically, organisational, ...).

In conclusion, the findings underscore the need for

- a central directory of RWD which also helps to enact quality standards on data sets,
- raising awareness and compliance with data standards, in particular the “Findable”-“Accessible”-“Interoperable”-“Reusable” (FAIR) data principles given that a substantial share of RWD is neither findable nor accessible,
- a more strategic approach to think about the roles and features of existing and future data sets, in particular by including the research purpose in RWD,
- resolving issues to warrant sustainable data management by providing adequate resources,
- a fundamental legal work and willingness to simplify the existing national legislation as well as to adapt it in an RWD-supportive manner to the (reformed) EU-layer of relevant secondary law and to

- leave data silo-ization behind and start creating interoperable data sets.

Methods

To meet the objectives, we tapped into expert knowledge within and outside the group of authors, conducted interviews, and common desktop research using search engines and employing snowballing techniques, i.e. searching research articles on Austrian healthcare and extracting the RWD source used. We applied of the following research strategies:

- Initially, based on a past survey we identified health data registers established by Austrian law.
- In addition, we searched the PubMed database for publications based on Austrian RWD sources (articles in the period from February 2017 to February 2022 including the criteria ((Austria[Affiliation]) AND (Austrian[Title/Abstract])) AND (data[Title/Abstract]).
- We then performed a targeted search for RWD at professional societies' and universities' websites.
- Last but not least, we searched in international RWD directories (e.g., OrphaNet) for Austrian RWD.
- Fifth and finally, the authors of this paper used their practitioners' knowledge to identify additional RWD sources in Austria.

Based on this search strategy, we extracted only healthcare-related RWD sources as described in the articles and listed those who fit the RWD definitions as published previously[11] (Fig. 8). We categorized results according to institutional data holder and category of the RWD source:

- For data holders, we differentiated between types of institutions that hold the data, of which include (1) expert communities (loose networks of experts without any formal organization), (2) professional societies (formally organized associations), (3) universities (organization under public law), (4) government institutions (ministries and public authorities including organization under direct state control based on private law), (5) hospitals, (6) social insurance organizations.
- We categorized the RWD sources based on collection main purpose derived from information available on the web and verified in interviews. "Main purpose" does not mean that the data cannot be used for other purposes, however it was defined based on the intended use during RWD establishment (= database setup / inauguration). We identified seven main purposes: (1) clinical, (2) epidemiological, (3) quality assurance, (4) regulatory, (5) administrative, (6) research, (7) informational.
- Finally, we categorized the subject of the RWD: (1) administrative data are data that are generated in administrative activities, (2) administrative registry also follow administrative purposes but have a legal basis, (3) biobanks store biological samples, (4) disease registries: the main data unit is a disease, (5) patient registries: the main data unit are human subjects, (6) product registries: the main data unit are products, (7) intervention registries: the main unit is an intervention (8) health care data bases includes various health care data, (9) observational study.

Following our objectives, we also conducted interviews with data holders of 11 RWD sources according to previously published criteria[11]. The sampling strategy was agreed upon by the author consortium and was used to create a representative RWD sample based on (1) purpose as well as (2) institutional type of data holder.

Declarations

Data Availability Statement

The datasets generated during the PubMed research approach described in the methods section are available from the corresponding author on reasonable request. All data analysed during this study are included in this published article (and its Supplementary Information files).

Author information

Contributions

B.M., V.M. and M.S. and conceived and conducted the research strategies and methodologies. D.B., T.C., P.K., DB.L., B.M., V.M., J.P-D. and M.S. conceived and conducted the deep dive interview(s). D.B., T.C., G.G-L., P.K., DB.L., B.M., V.M., J.P-D., M.S. and T.S. and analysed the results. All authors derived conclusion and discussion point(s). All authors reviewed the manuscript.

Ethics declarations

Competing interests

All authors declare no financial support or funding for this project. D.B. is an employee of Amgen GmbH, Vienna, Austria. B.M. is an employee of Novartis Pharma GmbH, Vienna, Austria. V.M. and J.P-D. are employees of Roche Austria GmbH, Vienna, Austria. All other authors declare no other conflicts of interest.

References

1. Burns, L., Le Roux, N., Kalesnik-Orszulak, R., Christian, J., Hukkelhoven, M., Rockhold, F., O'Donnell, J. Real-World Evidence for Regulatory Decision-Making: Guidance From Around the World. *Clin. Ther.* 44(3), 420-437. ISSN 0149-2918. DOI: https://doi.org/10.1016/j.clinthera.2022.01.012 (2022)
2. Valla, V., Tzelepi, K., Charitou, P., Lewis, A., Polatidis, B., Koukoura, A., Karapatsia, A., Antonopoulou, K., Prodromidou, K., Papadaki, E., Vassiliadis, E. Use of Real-World Evidence for International Regulatory

Decision Making in Medical Devices. *Int. J. Digital Health.* 3, 1. DOI: http://doi.org/10.29337/ijdh.50 (2023)

3. Baumfeld, A.E., Reynolds, R., Caubel, P., et al. Trial designs using real-world data: The changing landscape of the regulatory approval process. *Pharmacoepidemiol Drug Saf.* 29, 1201-1212. DOI: https://doi.org/10.1002/pds.4932 (2020)
4. European Medicines Agency. High-quality data to empower data-driven medicines regulation in the European Union. News 10/10. URL: https://www.ema.europa.eu/en/news/high-quality-data-empower-data-driven-medicines-regulation-european-union (2022)
5. Rudrapatna, V.A., Butte, A.J. Opportunities and challenges in using real-world data for health care. *J Clin Invest.* 130(2), 565-574. DOI: https://doi.org/10.1172/JCI129197 (2020)
6. Degelsegger-Márquez, A. The future of European health (data) systems. *PHIRI Spring School on Health Information.* URL: https://jasmin.goeg.at/2768/1/PHIRI%20Spring%20School%202023%20-ADM.pdf (2023)
7. Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act). EUR-Lex: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022R0868 (2022)
8. European Commission. Public Health. eHealth : Digital health and care. European Health Data Space, URL: https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space_en
9. Thiel, R., Deimel, L., Schmidtmann, D., Piesche, K., Hüsing, T., Rennoch, J., Stroetmann, V., Stroetmann, K., Kostera, T. #SmartHealthSystems. Digitalisierungsstrategien im internationalen Vergleich. BERTELSMANN STIFTUNG (HRSG.). URL: https://www.bertelsmann-stiftung.de/de/publikationen/publikation/did/smarthealthsystems (2018)
10. European Medicines Agency. Global regulators call for international collaboration to integrate real-world evidence into regulatory decision-making. News 22/07. URL: https://www.ema.europa.eu/en/news/global-regulators-call-international-collaboration-integrate-real-world-evidence-regulatory-decision (2022)
11. Klimek, P., Baltic, D., Brunner, M., Degelsegger-Marquez, A., Garhöfer, G., Gouya-Lechner, G., Herzog, A., Jilma, B., Kähler, S., Mikl, V., Mraz, B., Ostermann, H., Röhl, C., Scharinger, R., Stamm, T., Strassnig, M., Wirthumer-Hoche, C., Pleiner-Duxneuner, J. Quality Criteria for Real-world Data in Pharmaceutical Research and Health Care Decision-making: Austrian Expert Consensus. *JMIR Med Inform.* 10(6), e34204. https://medinform.jmir.org/2022/6/e34204 DOI: 10.2196/34204 (2022)
12. Feldman, K., Johnson, R.A., Chawla, N.V. The State of Data in Healthcare: Path Towards Standardization. *J Healthc Inform Res*. 2, 248–271. DOI: https://doi.org/10.1007/s41666-018-0019-8 (2018)
13. Bundesgesetzblatt. Datenschutz-Anpassungsgesetz 2018 – Wissenschaft und Forschung – WFDSAG 2018. BGBl. I Nr. 31/2018, https://www.ris.bka.gv.at/eli/bgbl/I/2018/31 (2018)
14. Bundesgesetzblatt. Änderung des Bundesstatistikgesetzes 2000 und des Forschungsorganisationsgesetzes. BGBl. I Nr. 205/2021,

https://www.ris.bka.gv.at/eli/bgbl/I/2021/205 (2021)

15. Proposal for a Regulation of the European Parliament and of the Council on the European Health Data Space. COM/2022/197 final. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52022PC0197 (2022)
16. Staatssekretariat für Digitalisierung, Bundesministerium für Finanzen. Digital Austria Act. Das Digitale Arbeitsprogramm der Bundesregierung. Digitales Gesundheitswesen. URL: https://www.digitalaustria.gv.at/Strategien/Digital-Austria-Act—das-digitale-Arbeitsprogramm-der-Bundesregierung.html (2023)
17. Bundeskanzleramt Österreich. Bundesministerium für Soziales, Gesundheit, Pflege und Konsumentenschutz. Geschäftszahlen BKA: 2023-0.004.651. BMSGPK: 2023-0.546.529. 67/18 MRV: Sofortmaßnahmen Gesundheitsreformpaket, URL: https://www.bundeskanzleramt.gv.at/dam/jcr:0006ab80-11f1-4716-877d-61e5c0a2f60a/67_18_mrv.pdf (2023)

Table

Table 1 is available in the Supplementary Files section.

Figures

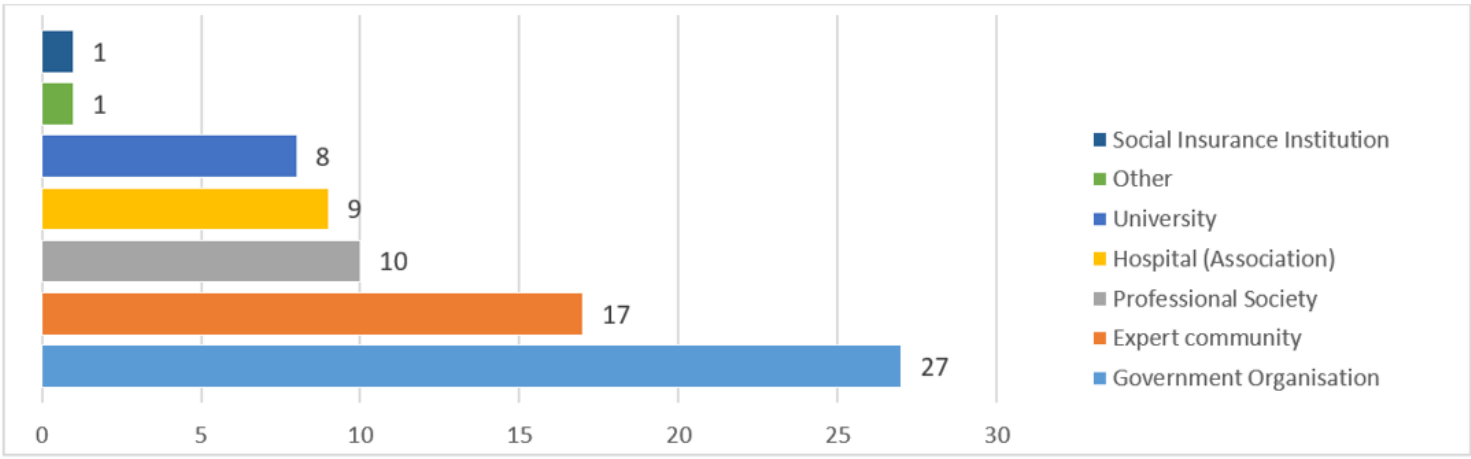

Figure 1

Number of RWD Sources per Institutional Type.

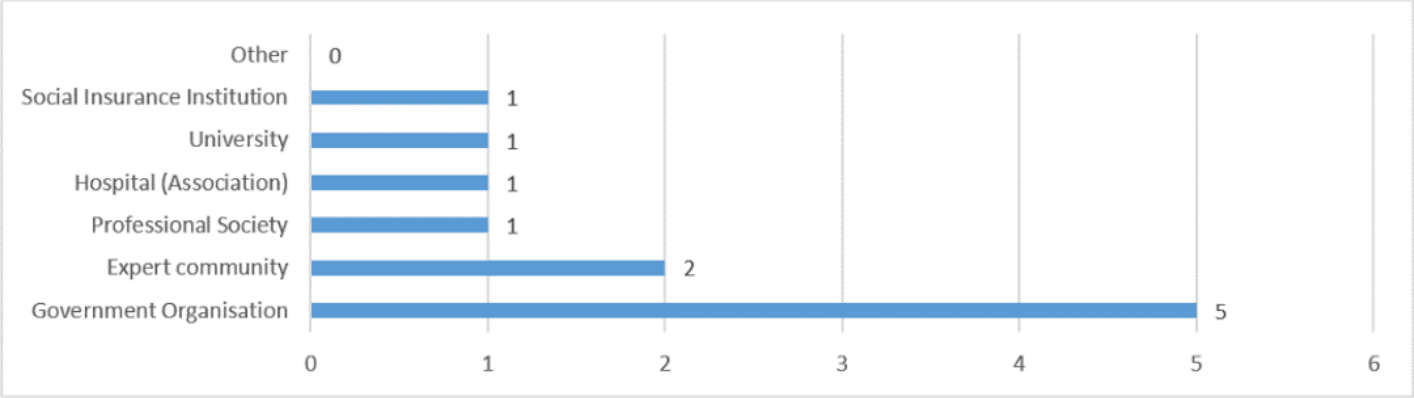

Figure 2

Distribution of institutional types of RWD holders amongst interview sample.

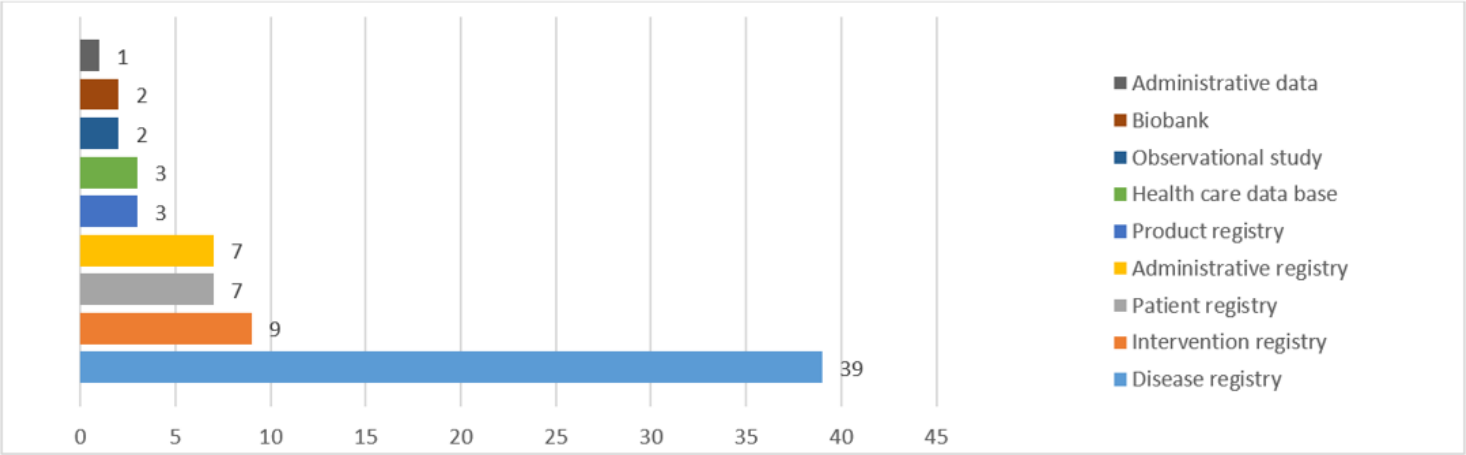

Figure 3

Main Category of RWD Sources.

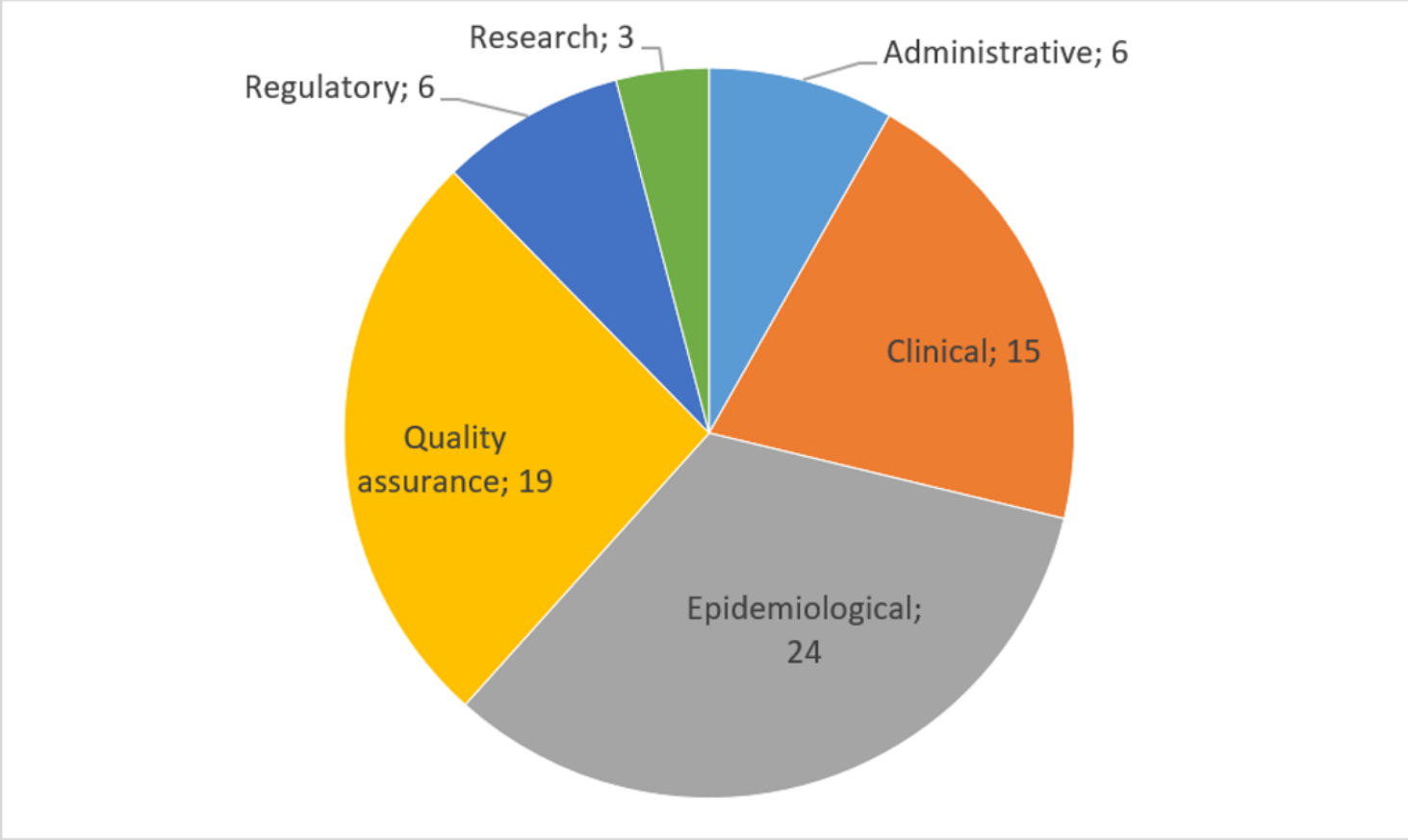

Figure 4

Distribution of main collection purpose of RWD Sources overall.

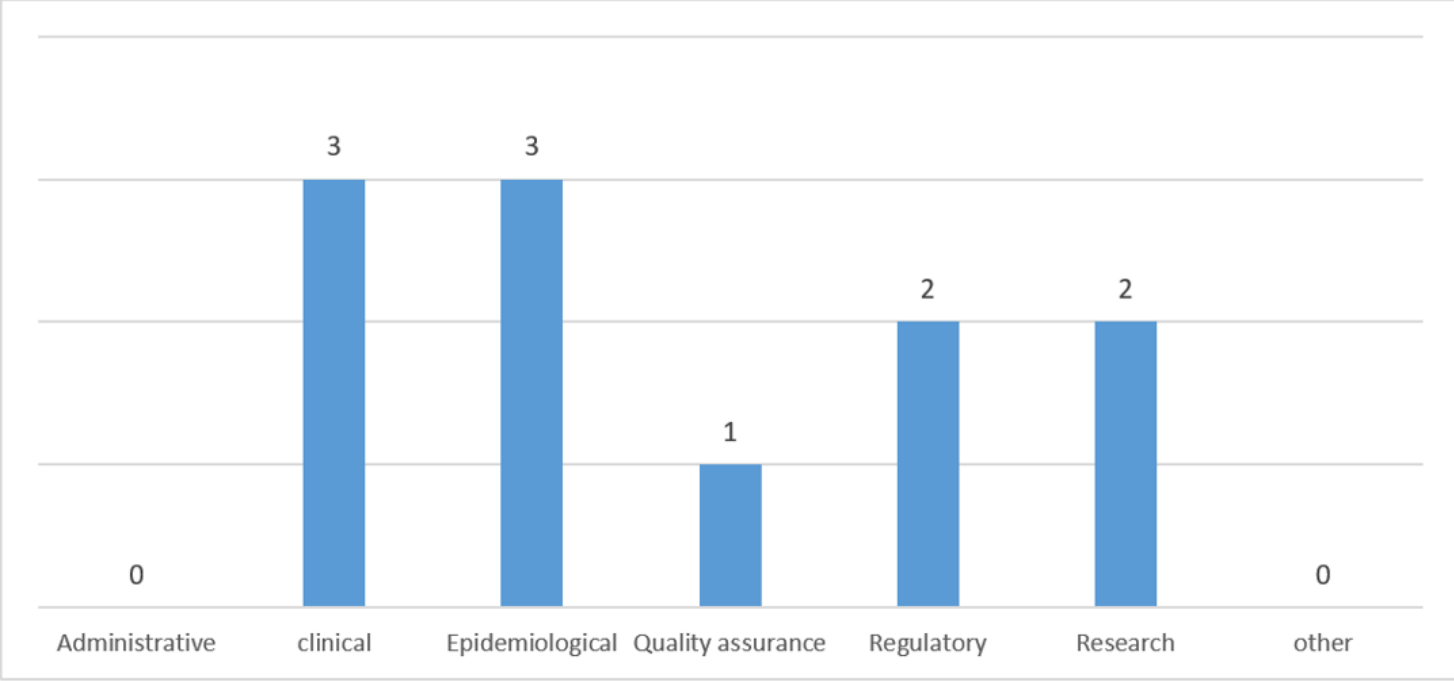

Figure 5

Distribution of main purpose of RWD source amongst Interview-Sample.

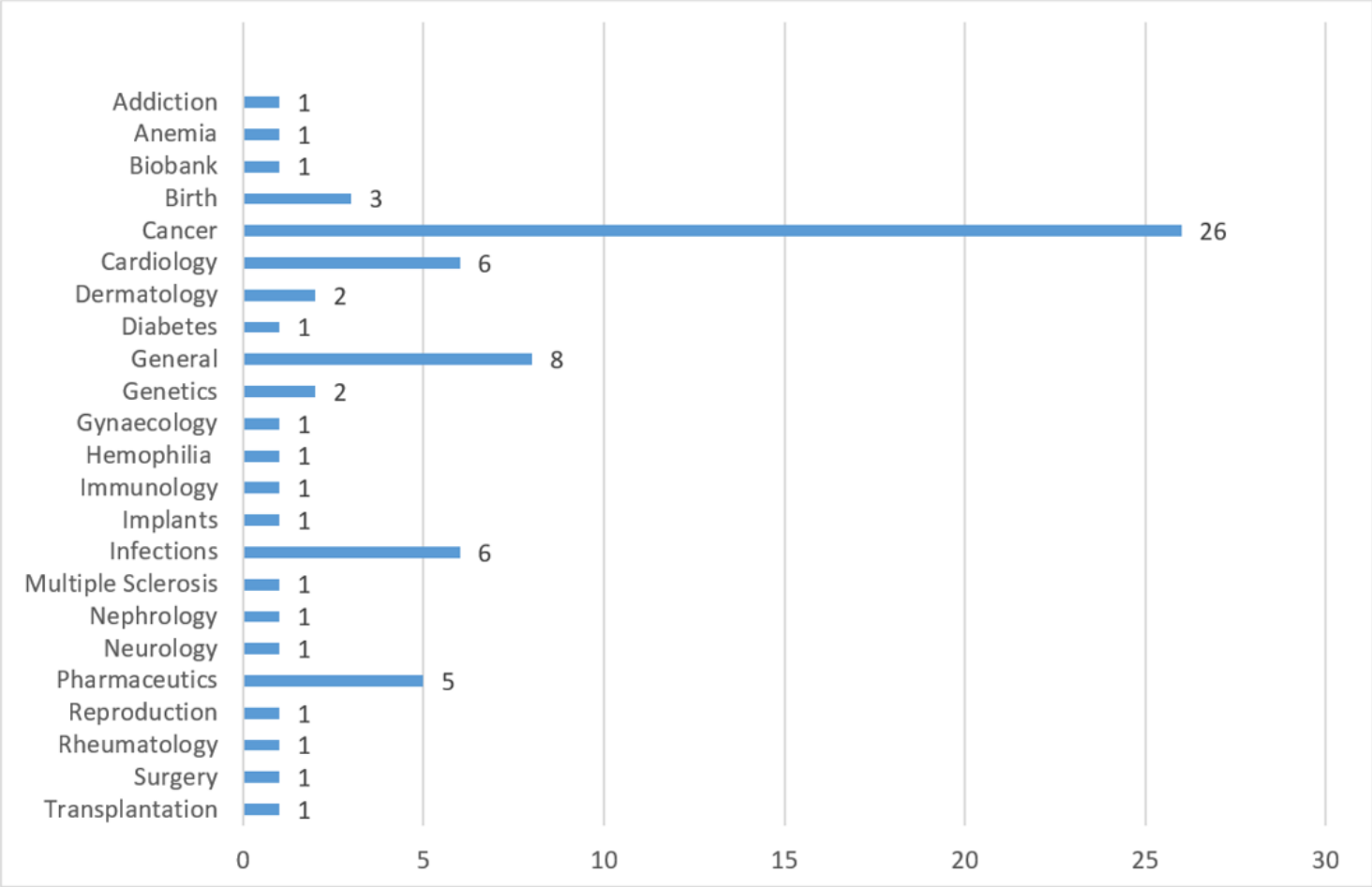

Figure 6

Disease or topic wise areas of RWD sources in Austria.

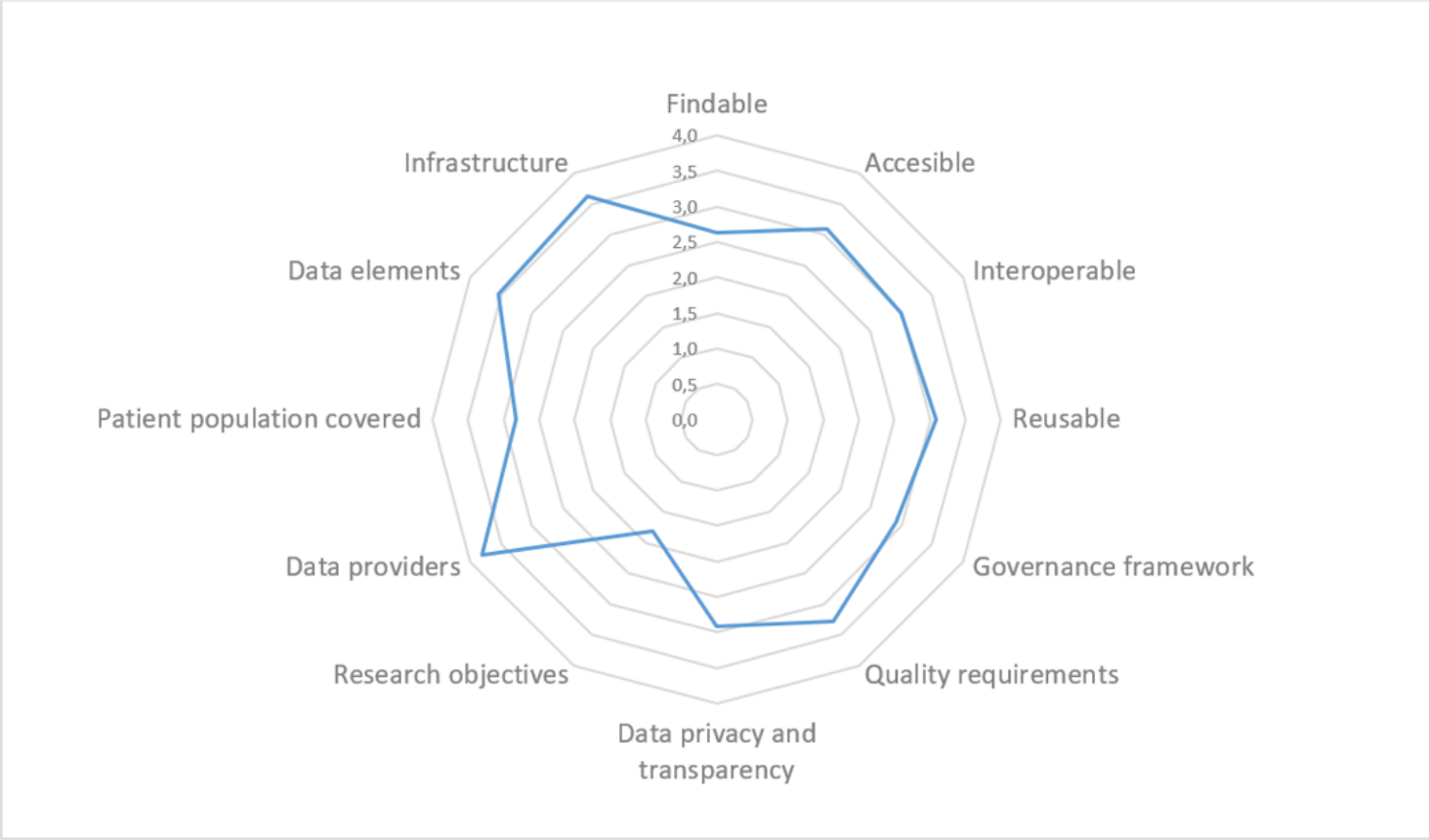

Figure 7

Achieved quality criteria of examined 11 RWD sources.

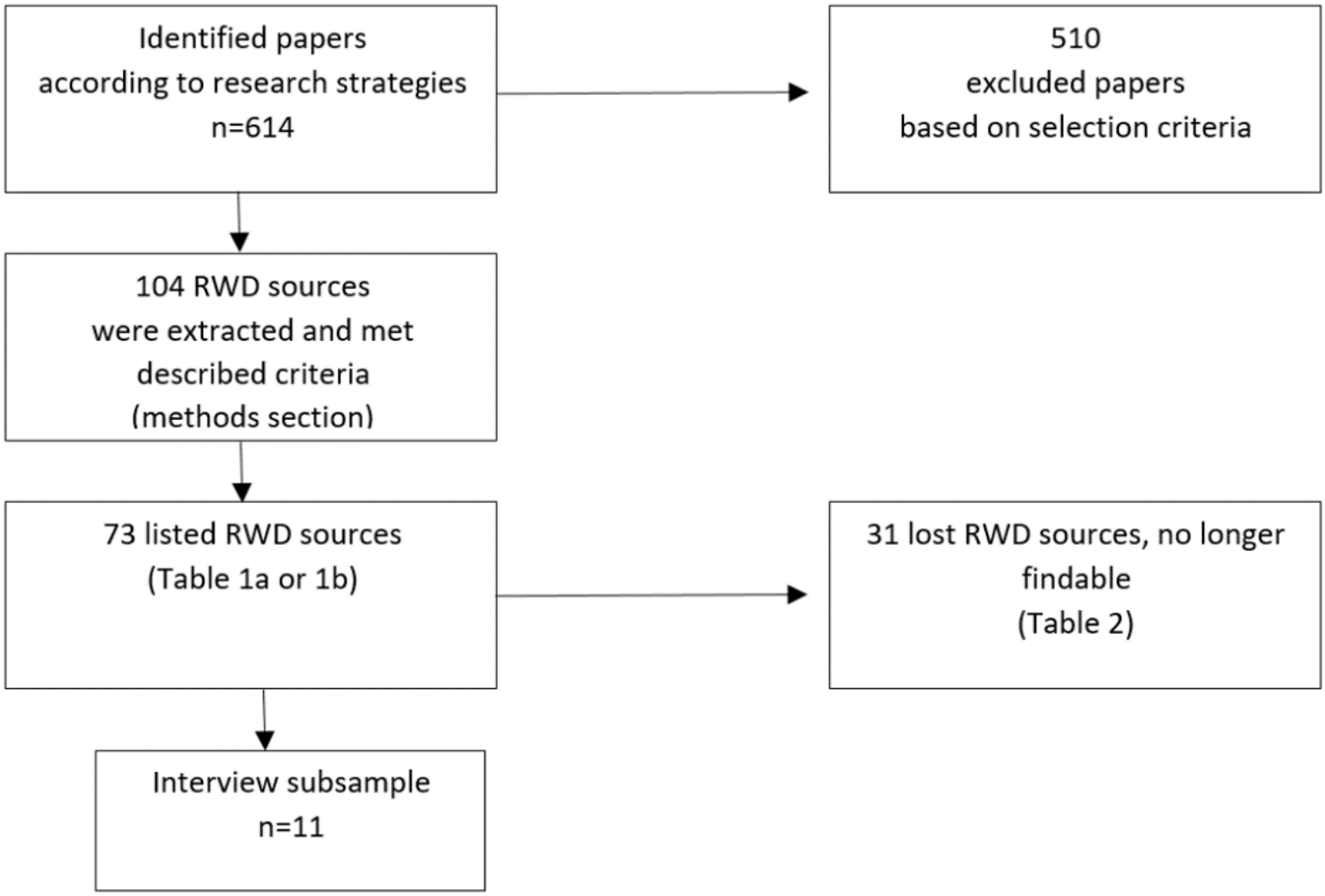

Figure 8

RWD source inclusion and selection process.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Table1.docx
- 4SupplementaryInformationMaintablesnatureSR.docx