Research and Applications

Implementation of a data control framework to ensure confidentiality, integrity, and availability of high-quality real-world data (RWD) in the NeuroTransData (NTD) registry

Knut Wehrle[1], Viola Tozzi[2], Stefan Braune[1], Fabian Roßnagel[1], Heidi Dikow[1], Silvia Paddock[2], Arnfin Bergmann[1], and Philip van Hövell[2]

[1]NTD Study Group, NeuroTransData GmbH, Neuburg an der Donau, Germany, and [2]PricewaterhouseCoopers AG, Zurich, Switzerland

Corresponding Author: Knut Wehrle, NeuroTransData GmbH, Bahnhofstraße 103B, 86633 Neuburg an der Donau, Germany; kwehrle@neurotransdata.com

ABSTRACT

Objective: To implement a dynamic data management and control framework that meets the multiple demands of high data quality, rigorous information technology security, and flexibility to continuously incorporate new methodology for a large disease registry.

Materials and Methods: Guided by relevant sections of the COBIT framework and ISO 27001 standard, we created a data control framework supporting high-quality real-world data (RWD) studies in multiple disease areas. We first mapped and described the entire data journey and identified potential risks for data loss or inconsistencies. Based on this map, we implemented a control framework adhering to best practices and tested its effectiveness through an analysis of random data samples. An internal strategy board was set up to regularly identify and implement potential improvements.

Results: We herein describe the implementation of a data management and control framework for multiple sclerosis, one disease area in the NeuroTransData (NTD) registry that exemplifies the dynamic needs for high-quality RWD analysis. Regular manual and automated analysis of random data samples at multiple checkpoints guided the development and implementation of the framework and continue to ensure timely identification of potential threats to data accuracy.

Discussion and conclusions: High-quality RWD, especially those derived from long-term disease registries, are of increasing importance from regulatory and reimbursement perspectives, requiring owners to provide data of comparable quality to clinical trials. The framework presented herein responds to the call for transparency in real-world analyses and allows doctors and patients to experience an immediate benefit of the collected data for individualized optimal care.

Key words: data accuracy, registries, reference standards, information technology, trust

Lay Summary

Workflows of doctors treating patients in medical offices have changed dramatically in recent years. Computers have replaced paper files, and physicians rely on sophisticated digital systems to keep track of data from many sources, including laboratory results and patient-reported outcomes collected using tablets and smartphones. Data collected through these efforts are termed "real-world" data, because they reflect actual healthcare encounters in daily practice. Too often, such data constitute an untapped resource for learning, because the results from individual patients are kept decentralized. When kept separately, it is not possible to understand if certain groups of patients may respond better or worse to certain treatments. Physicians in the NeuroTransData network in Germany decided to remove the barriers for learning by connecting their practices. They now all contribute data to a central data collection called a "registry." Personally identifying information is replaced with codes to protect privacy, and the data are checked automatically for errors to ensure that no mistakes are introduced. We describe the processes to collect, handle, and analyze real-world data in this network to support scientific studies with data of high quality and to allow insights derived from the data to directly benefit the patients in the network.

INTRODUCTION

Background and significance

There is an urgent need for real-world data (RWD) to fill the information gap between randomized controlled trials (RCTs) and evidence based, post-registration clinical use of innovative treatments in increasingly complex disease landscapes.[1,2] Analysis of fit-for-purpose RWD can address long-term information gaps on comparative effectiveness and post-authorization safety.[3,4] It can also answer questions on patient segmentation, treatment pathways, resource allocation, and pharmacoeconomic issues for health-policy decisions and regulatory affairs. As in RCTs, high integrity and transparent quality of RWD are the foundation for valid and meaningful results.

Because data capturing and analysis must be integrated in daily practice of multiple centers with different environments, qualifications, work processes, and time constraints, RWD acquisition encompasses specific challenges. While RCTs have long been governed by established Good-Clinical-Practice (GCP) standards that are monitored during the study, RWD capture and management is still a developing field. To work toward common standards, pharmaceutical regulators recently embarked on the definition of acceptable RWD operational processes. In December 2021, the US Food and Drug Administration (FDA) released a new draft guidance for industry regarding the use of RWD and real-world evidence (RWE),[5,6] and in October 2021, the European Medical Agency (EMA) issued novel guidance on registry-based studies.[7,8]

We identified 4 critical "pillars" of a solid large-scale registry that can respond to the demands of regulators and mitigate known challenges to RWD collection: (1) data management procedures need to be sufficiently nimble to allow for fast learning and support individual treatment decisions yet robust enough to guarantee high data quality; (2) the registry needs to be large and representative of the population of interest; (3) biases (eg, selection bias, unknown confounders) need to be accounted for to compensate for the lack of randomization in RWE studies, and (4) all analyses conducted on the data need to be logged and reported, including feasibility studies, to support full transparency. We herein address the first of these pillars by reporting on the design and implementation of a data process and control framework. The other 3 pillars will be the topic of separate publications (manuscripts in preparation).

NTD real-world disease registry database and patient-management system

NeuroTransData GmbH (NTD) is a nationwide network of doctors in Germany in the field of neurology and psychiatry.[9] Partners in the network are recruited from modern, digitalized practices providing comprehensive state-of-the-art care with large patient numbers, which is monitored via yearly certification based on ISO norm 9001. The company currently consists of 60 practices and 120 partners. Each year, more than 600 000 patients are treated by the network's practices. Since 2008, the NTD network has built disease registries currently active in the indications of bipolar disorder, dementia, epilepsy, migraine, multiple sclerosis (MS), Parkinson's disease and movement disorders, depression, and schizophrenia. The most extensive registry covers MS with 25 242 MS patients currently documented and an average observation period of 6.11 years.

To satisfy stakeholder demands for timely delivery while adhering to the above operational standards, registries need to implement state-of-the-art data process technology to ensure confidentiality, integrity, and availability of RWD. An effective data control framework is an essential requirement for the generation of timely high-quality data. To this end, the NTD doctors' network initiated a collaboration in 2016 with health data management experts at a data processing site (DPS), managed by PricewaterhouseCoopers AG in Zurich, Switzerland, to implement a data process and control framework for RWD captured daily in real-time during clinical visits in neurological and psychiatric practices belonging to the NTD practice network in Germany.

Collaborating with the centers of excellence that constitute the network, NTD links medical, content and patient-oriented aspects of care, offering physicians and patients an integrated clinical decision support environment (Figure 1). The core of this environment is the DatabasE-aSsisted Therapy decIsioN support sYstem ("DESTINY®"),[10] jointly developed by NTD and the DPS. The NTD disease registry databases provide the technical and organizational foundation for this multi-modular patient management system, which enables bidirectional data exchange. DESTINY® supports communication between physicians and patients and encourages active involvement of the patient in data documentation processes.

The DPS regularly performs data analyses based on pre-defined study protocols to allow stakeholders (eg, clinicians, scientists, pharmaceutical companies, pharmacies, healthcare providers, regulatory authorities) to gain meaningful insights from the registry's large-scale datasets, the timely generation and delivery of which is critical to its usefulness.[11,12] In the field of MS, the clinical utility includes the development and continuous support of several clinical support tools, including PHREND®,[13] a personalized predictive algorithm for therapy optimization in relapsing remitting MS and an algorithm to evaluate the risk of progression to secondary progressive MS. The algorithms are based on the registry data and updated every quarter to support the shared clinical decision process of doctors and their patients through the DESTINY® patient management platform (Figure 1).
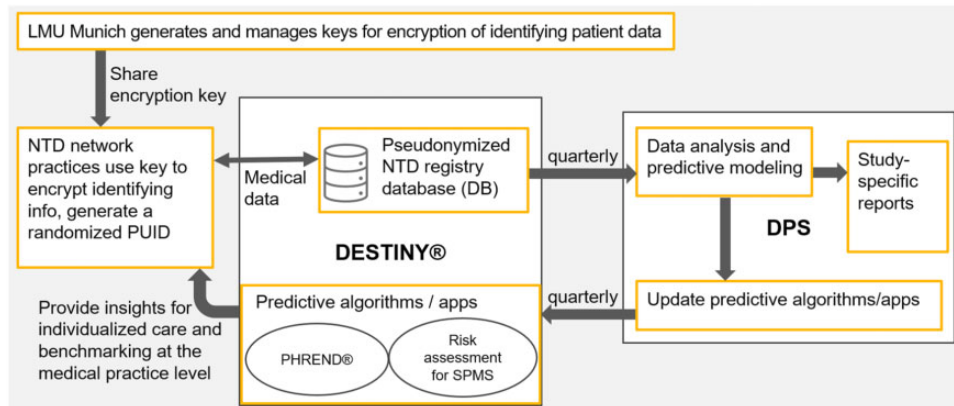
Figure 1. Overview of the components of the registry and analytic site and the data flow in the field of multiple sclerosis (MS). Results from predictive modeling feedback into the DatabasE-aSsisted Therapy declsioN support sYstem (DESTINY®) every 3 months to support treatment decision making in the real world. DPS: data processing site; LMU: Ludwig Maximilians Universität; NTD: NeuroTransData; PUID: patient unique identifier; SPMS: secondary progressive multiple sclerosis.

MATERIALS AND METHODS

Mapping the data journey

Understanding data processing and identifying potential threats in data handling are fundamental first steps of the development of a control framework. We therefore first created detailed maps of the entire data flow. As shown in the overview presented in Figure 1, the data process starts with the data input coming from NTD network practices and ends after data model output derived at the DPS feeds back into DESTINY® to assist in treatment decision-making. The detailed maps created in this process were then used as the basis for the control framework.

Implementation of the data control network

After mapping the data journey, we systematically identified potential risks for data loss or inconsistencies in each step of the journey. We then created the data control framework based on relevant sections of the COBIT[14] and ISO 27001[15] standards and implemented the following 5 elements:

1. Access and change management
2. Business continuity management
3. Data operations regarding data input
4. Transfer of patient data
5. Controls for completeness and accuracy

By systematically applying this framework to the mapped process, we evaluated whether the controls adhered to best practices described in COBIT and ISO 27001 to ensure the confidentiality, integrity, and availability of data from the registry. To test the effectiveness of the controls, random data samples were taken and analyzed during the creation process of the framework, and these checks are now being conducted on a regular basis.

Creation of a strategy board

To monitor the implementation and continuous evolution of the framework, we created an internal strategy board to review the results and regularly assess and implement potential improvements. Board members include medical doctors, data scientists, mathematicians and statisticians, and project management leaders of the registry and the DPS; meetings are being held every 3 months.

RESULTS

Complete maps of the data journey

Overviews

Figure 2 provides a generic summary of the data journey from input to analysis, while Figure 3 shows specific elements from the field of MS that includes the calculation of the PHREND® treatment optimization algorithm made available not only to NTD members but also external users through a stand-alone web-based application (https://phrend.neurotransdata.com; user registration through the following website: https://www.neurotransdata.com/en/destiny#phrend).

Patient ID management and pseudonymization process

NTD does not have any access to personal patient information collected by the member practices, which pseudonymize the data before saving them to the NTD registry database (Figure 1). Each medical data record stored in the NTD database (registry) thus has a randomized unique identifier (PUID), which is used to assign data to the patient. Based on this PUID, pseudonymized data from different sources are assignable to the correct patient. The encryption keys linking the patients with their identifiable data are managed by the Institute for medical information processing, biometry, and epidemiology (Institut für medizinische Informationsverarbeitung, Biometrie und Epidemiologie [IBE]) at the Ludwig Maximilians University in Munich, Germany, acting as an external honest broker.

All data acquisition and management protocols were approved by the ethical committee of the Bavarian Medical Board (Bayerische Landesärztekammer; June 14, 2012, ID 11144) and re-approved by the ethical committee of the Medical Board North-Rhine (Ärztekammer Nordrhein, April 15, 2017, ID 2017071). Compliance with European and German legislation (General Data Protection Regulation [GDPR], Bundesdatenschutzgesetz [BDSG]), including patient rights and informed consent requirements, is documented in detail in an internal data security handbook and confirmed in yearly audits.

Data entry and pseudonymization process and entry into the NTD registry database

The data processing system provides different methods to collect data from different sources during routine patient care in the daily work of the NTD medical practices. Data are either automatically uploaded (eg, laboratory data) or entered manually into the database via the web-based platform DESTINY®.
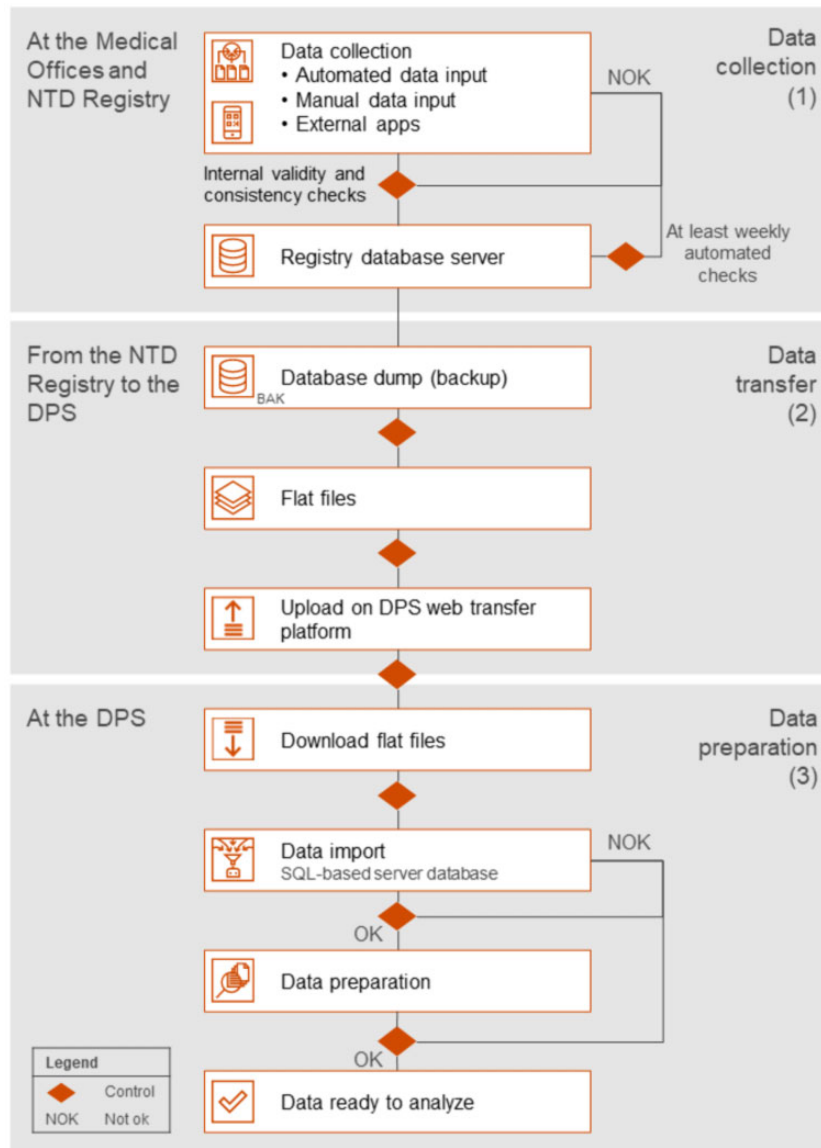
Figure 2. Journey of real-world data from capturing in the NTD registry to analysis at the DPS. Diamonds indicate checkpoints of the data control framework. (1) Already during the data collection process, dropdown menus ensure consistency during the input procedures and multiple automated checks ensure that accurate data are uploaded to the server. (2) The data transfer process from the registry to the data processing site (DPS) contains multiple checkpoints to ensure the integrity of the transfer. (3) After arriving at the DPS, the data are uploaded into a secure and strictly access-controlled structured query language (SQL) database and undergo additional control steps in a process that ends with a dataset ready for analysis. DESTINY®: DatabasE-aSsisted Therapy decIsioN support sYstem; DPS: data processing site; NTD: NeuroTransData.

- Automated data input
  - Data from external sources (eg, laboratory results, patient-reported outcomes) is uploaded automatically into the NTD database. NTD provides an application programming interface (API) which is used for the different sources. To assign the data to the correct patient, the PUID is mandatory for the API which uses different methods to ensure data security:
    - ■ Transport Layer Security/Secure Sockets Layer (TLS/SSL)-encryption: transport-layer encryption with state-of-the-art cipher suites is implemented.
    - ■ Authentication information: for every automated input, an authentication field is necessary.
  - Due to this security process, access to the data is restricted and an authentication key is needed to send data via the

API. The authentication keys required to write data into the database are managed by NTD, so that only data from intended sources can be loaded into the database.
- Manual data input
  - Data collected at the doctors' practices are entered manually into the database via a modern, browser-based single-page application.
    - ■ The user interface and the backend with the database are encrypted via TLS/SSL.
    - ■ Doctors and nurses must authenticate with personal credentials to view or insert data of visits and results of diagnostic exams in the registry database.
    - ■ Dropdown menus ensure consistency regarding nomenclature and avoid typos during entry.
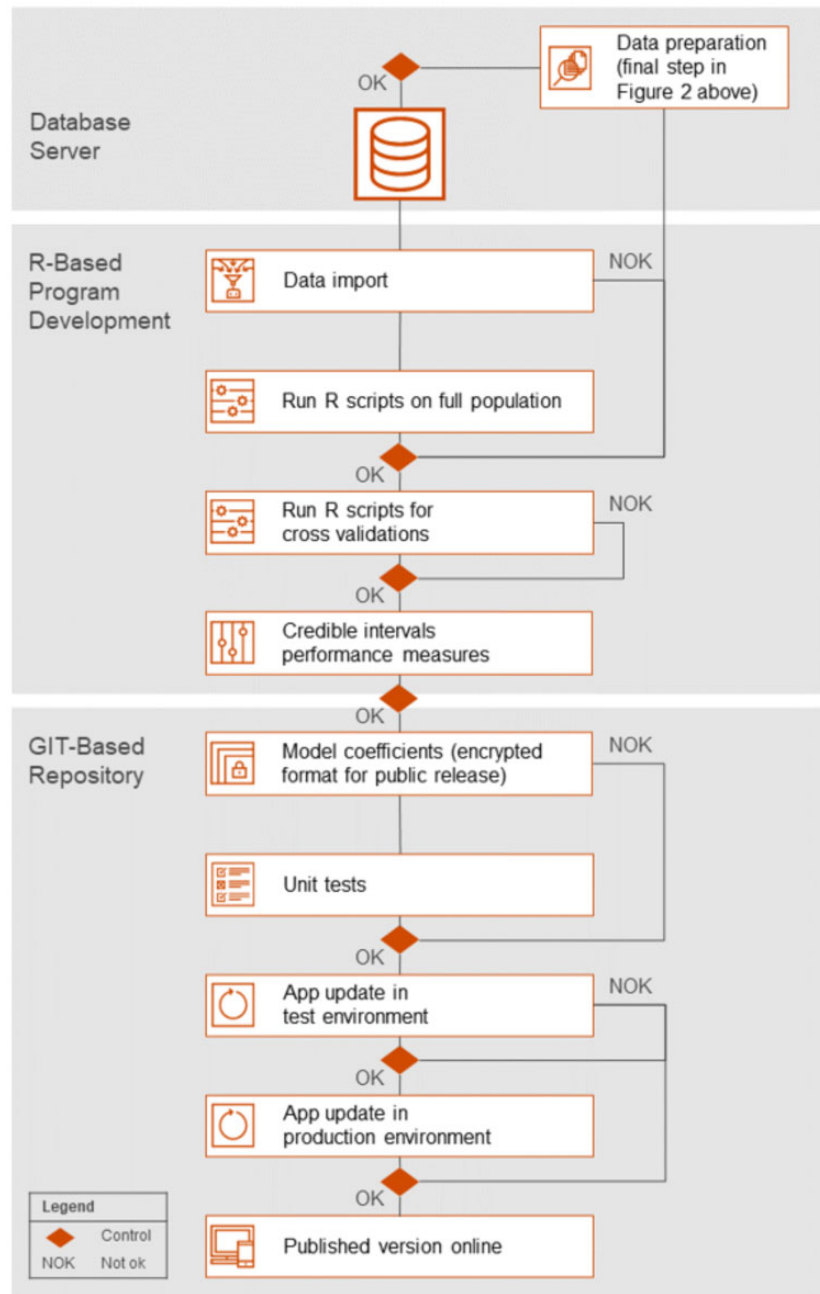
Figure 3. Analysis process in an example from multiple sclerosis (MS). Diamonds indicate checkpoints of the data control framework. After arriving at the DPS, the files undergo multiple checks based on version-controlled scripts written in the programming language R. Errors discovered in these checks lead to exclusion of records from further processes. The clean data are then passed on to the next step and used to calculate the new model parameters. The processes that update the model also occur in a version-controlled ("GIT") environment to ensure that a complete audit trail of changes is available. The updated App with new model predictions then undergoes series of tests for feasibility before the final version is uploaded to the public website.

■ Free-text input fields (eg, measured values) are checked upon entering based on regular expression patterns.
- Logs are written for each database change, constituting a complete audit trail of all data entered in the database.

Automated and manual data inputs occur through an application programming interface (API) and a web interface, the front-end of the DESTINY® platform. Core data sets have been defined for each indication, which are requested to be captured when the patient is first included into the registry and at every visit afterwards. Data in-

put and registry databases are managed by dedicated departments of the NTD administration and NTD IT services, respectively. Detailed user handbooks for the different systems are available and updated regularly as part of the ISO 9001 certification and GDPR compliance processes.

Data backup
The NTD registry has an implemented backup procedure and data restorage plan. The databases, which are hosted by an external pro-

vider, are backed up daily. The encrypted and compressed database backup archives are stored locally and are synced to an (offsite) cloud storage.

Transfer of the data from the registry to the DPS

Once every quarter of the year, the registry data are extracted by NTD as a set of flat files from the restored backup of the first day of each quarter (Figure 2). The extracted files are uploaded to a secured web-transfer platform provided and run by the DPS.

Overview of the system landscape at the DPS

The DPS processes the data in a virtual desktop environment with restricted access (Figure 2). All data remain pseudonymized and stored on an internal server with closely monitored access and read/write restrictions. All activity is logged by the system administrators. Additional structured query language (SQL) audit logs can be turned on for individual projects to keep track of changes of data tables and create complete records of all analyses conducted on a specific dataset. A commercial database management system is used for the preparation, cleaning, and transformation of the raw data.

Detailed data journey characteristics using the example of MS

Using the example of MS, Figure 3 shows the detailed data flow to process quarterly data exports and generate parameters for the treatment optimization algorithm (PHREND®) available through the DESTINY® platform. The process starts with further data preparation and ends with steps to update the predictive model and the web-based treatment-optimization application.

The NTD data dictionary for the MS registry is updated regularly and contains all items featured in the registry database. Patient-related outcomes, clinical variables, co-morbidities, medications, adverse events, socio-economic parameters, and family planning or pregnancies are captured in real time during clinical visits, including the MS core data set (eg, ICD-10 codes, anatomical therapeutic chemical [ATC] medication codes) as recommended by the European medical agency (EMA/548474/2017).[16] MedDRA adverse events coding is currently being implemented and will be available in 2022.

Each module contains information on a specific topic divided in items. The imported data within the database is transformed as a preparation step for the model fitting in R, results of which are used in the PHREND® web-based application. In this process, the dictionary data is joined to the master data by the item ID. The data are then denormalized, preparing 1 table per module (ie, patient data, visit data, incidence data, and therapy data).

Data quality controls

Overview of the control framework at NTD

The purpose of the framework is to ensure that the data is stored securely ("confidentiality"), cannot be accidentally modified in unpredicted ways ("integrity") and is ready for analyses in a timely manner ("availability"). At the NTD registry, the data quality framework is characterized by 3 components:

1. Identification of quantitative data quality key performance indicators (KPIs);
2. A data governance framework including processes and controls; and
3. Implementation of an actionable data quality dashboard in DESTINY®.

As part of the implementation process of the data control framework, KPIs were identified to monitor the integrity of the registry data, and processes were put in place to monitor their results and track changes in the indicators over time. The data governance framework ensured that all mapped processes had assigned owners and quality metrics assigned, and the dashboard finally ensured that any errors identified in the data checking processes could be addressed by NTD members in a systematic and timely manner.

Implemented controls for data input

Data input into the NTD database is controlled by automatically implemented measures and regularly conducted manual controls during the ISO 9001 audits. For each data transferred automatically to the database, performance testing of the API is completed. Furthermore, data transfer is only possible when the User ID and the PUID are correct to ensure that automatically processed data are consistent and assigned to the correct patient.

Once in a year, an on-site primary source data monitoring takes place in a randomly chosen sample of NTD offices, performed by an external auditor commissioned by NTD. A random sample of 10 patients in each office is selected and consistency of source data documentation in the electronic health record and practice management software system and in the NTD registry is investigated. This annual source data monitoring is performed within the annual recertification for the ISO 9001 certification[8] of NTD offices.

Furthermore, automatic quality assurance queries are in place for all data entries. The objective of these controls, examples of which are shown in Table 1, is to identify inconsistencies, gaps, or erroneous data. A version control system of the automated controls running on the database is in place for documentation purposes and internal governance.

Missing data trigger notifications in a weekly report shown to the doctor in the registry web-application dashboard in context-specific form (eg, the doctors can see all alerts that correspond to a given patient) to support data completeness.

Monitoring of data backup quality

Access to the backups is limited to the NTD IT team. The backups are tested on a quarterly basis by NTD's IT service employees during the process of data transfer to the DPS.

Access management

Roles and responsibilities for the data journey are defined and an inventory of involved staff and systems exists. Based on these roles, access rights to the individual systems (ie, code repositories, servers, data) are provided. Up-to-date lists of active users per systems are reviewed regularly.

Data quality controls at the management level

NTD has implemented the following measures to mitigate risks while capturing and processing the data:

- Implementation of standardized user interfaces and masks for data entry and data definitions to avoid inconsistencies during entry.
- IT support provided for all users and error management is in place. In an NTD-internal online forum, questions on functionalities are answered and detailed information on specific workflows of DESTINY® are documented.

Table 1. Example of quality assurance queries

Category	Example messages
Error	[Medication name] was prescribed and canceled on the same day ([date]).
	[Medication name] was canceled on [Date] without giving a discontinuation reason.
	Relapse without date.
	Anamnesis: The first diagnosis ([Date]) cannot possibly be before a first manifestation ([Date]).
Information	No first diagnosis date found.
Warning	Patient did not have any visit in the last quarter.
	The patient does not have a single visit. Please complete the relevant core data sections or delete the patient.

- The version history for the registry database with information about added or changed features and solved bugs is also available in the forum.
- Processes are described and available for all users and regular training sessions are provided.
- Availability of resources for data management and security is checked regularly.
- A joint registry steering committee meets to check and discuss data quality issues.

Data controls for the transfer from the registry to the dedicated DPS

Each step of the transfer of data from NTD to the DPS follows a regulated and fully documented process. Once every quarter of the year, the registry data are extracted by NTD as a set of flat files from a database backup. A control is implemented to ensure completeness and accuracy of the data export. The change in data quantity is recorded per number of entries, patients, visits, and therapies. A consistent increase of data over time indicates timely data capture and constant data density, and deviations from the expected increase can be identified and investigated.

Data preparation process

All code for the data preparation process is stored on version-controlled repositories stored in access-controlled servers on DPS premises with clearly assigned roles and responsibilities for each step of the process. Only the directly responsible staff members at the DPS have access to the web-transfer platform to download the data. Access to the data is limited only to relevant persons throughout the entire transfer process.

Controls for data processing at the DPS

After downloading the flat files generated by NTD, checks for accuracy and completeness are performed by the DPS based on the checksum string of the downloaded data package. All file sizes are also compared to the sizes of the last quarter using an automated R-script, since the file size should increase steadily. The reconciliation results are then stored.

If all data import controls have been passed, the flat files are imported by a bulk insert statement as part of a set of SQL scripts to a dedicated database management system, whereby each year's data is stored in a separated database. The bulk insert statement is stored on a version-controlled repository and manually executed. A further completeness check is performed after the data import, whereby the export summary report provided by NTD is compared to the imported data.

The data preparation process now starts via a set of SQL scripts performing data cleaning and transformation. Duplicate records are identified and removed by comparing all attributes. For all data tables, additional data cleaning steps are performed based on plausibility checks: in these checks, records are removed based on either missing information or implausible dates (eg, diagnosis dates in the future).

Plausibility of results

Once the process is finished, an analysis of change (AoC) is being performed on the original and the cleaned data to compare the cleaning process for each quarter and the development of the main modules, ie, number of patients, therapies, events, relapses, and visits. In case these changes are not plausible, a cross check with NTD is performed using the issue-tracking system (eg, an office may have conducted an enrollment initiative).

Change management using agile processes

Changes made are defined and collected as requirements and shared among the joint NTD and DPS strategy board. New requirements are either added by the DPS or NTD and agreed upon in quarterly status update meetings. Changes are discussed, authorized, and prioritized by the strategy board. The change requests are then translated into technical requirements and saved within an issue tracking system. A project timeline is created according to preliminary time estimations and the work is planned assigned to sprints according to the priorities. The version-controlled code files have restricted access. Each change to the source code is tracked and connected to a change request. Changes are deployed on the test environment and must be validated by the strategy board before they can be deployed in production.

Following this process, changes are implemented and reviewed. Results are discussed with the respective product and process owners and then presented to the NTD stakeholders for feedback. If approved, the code changes are implemented and deployed in a version-controlled repository and used for the new quarterly update. A sprint is considered complete when all changes have been deployed to the test server and at least 3 NTD stakeholders have confirmed the correct functionality. Once the final quality-check gate is passed, the release is deployed to the production system.

DISCUSSION

Importance of high-quality data

The requirements by payers to demonstrate a treatment's cost-effectiveness have increased for pharma companies, and the healthcare sector is increasingly required to measure and demonstrate patient-related outcomes. Although RCTs still represent the gold standard for evidence and regulatory decision-making regarding a drug's tolerability, efficacy and outcomes, regulators are increasingly open to RWD/RWE as important source of information in post-authorization processes. In addition, doctors require high-quality data to make decisions about optimal individual care in an increasingly complex treatment landscape. The example of MS, an indication with a plethora of treatment alternatives in which doctors may struggle to identify the best treatment choice for individual patients, illustrates the need for timely data generation to support the feedback loop providing quarterly updates of the clinical decision support algorithms to DESTINY® and to support shared treat-

ment decision making during visits of individual patients in a timely manner.

Need for a robust data control framework

To achieve high-quality data, standardized processes and regular controls must be established to ensure accurate data management in a constantly evolving real-world environment and address emerging data quality risks. Through a defined, transparent, and standardized data control framework, risks to the data process can be identified and controls implemented to address them. Feedback loops ensure the detection of data inconsistencies or implausible results at any stage of the process. Several standard security measures are limiting direct access to the data to protect its confidentiality, and multiple controls are in place to prevent the loss of data and thus ensure data integrity. Updates are performed by standard operating procedures and reviewed for plausibility, and all changes to the code are reviewed and tested.

Further required developments

As expectations of regulatory authorities in Europe and US regarding post-authorization drug monitoring are becoming increasingly more detailed and the political will to utilize RWD in health resources allocation decisions gains momentum,[17,18] the necessity to develop living structures for qualified RWD data capturing needs to be met by reliable concepts and longstanding financing to make data acquisition and use a natural part of patient care in daily practice. While it is neither feasible nor desirable that RWE replace all RCTs in the foreseeable future,[19] registry-based real-world studies hold great promise to allow for high-quality evidence generation in situations in which RCTs may not be feasible of prohibitively expensive.[20] The crucial element for meaningful analysis is to capture the spectrum of medical information, patient outcomes and resource utilization simultaneously.

CONCLUSION

Data quality enabling reliable and robust information to supplement information from RCTs can be achieved by adaptation of established data acquisition and processing procedures for the real world, based on established standards. Implementation of such a processing and control framework in medical disease real-world registries is feasible via a dedicated interdisciplinary cooperation between experts in data acquisition, management, and analytics. Sustained resources supporting these activities are needed for a longer period than for RCTs to allow for the generation of meaningful and robust long-term data and real-world insights.

In our understanding, this sustainability can only be achieved if RWD capturing is part of private (eg, licensing of data, analyses, and models) and public (eg, digital health apps, DIGA in Germany) reimbursement concepts and if there is the experience of an immediate benefit of the collected data for doctors and patients in their daily efforts to realize individualized optimal care. This underlines the necessity that RWD data capturing needs to be embedded into routine care procedures by advanced IT techniques and needs to provide immediate feedback for imminent clinical decisions. Rigorous data control frameworks will become critical to ensure that all this can be done with a level of data accuracy that mirrors the quality of GCP-based data processing in RCTs.

CONFLICT OF INTERESTS STATEMENT

All authors are supported by either NeuroTransData GmbH or PwC AG, as indicated in the author affiliations, and performed the work without the use of any external funding sources. AB has received consulting fees from advisory board, speaker, and other activities for NeuroTransData; project management and clinical studies for and travel expenses from Novartis and Servier. SB has received honoraria from Kassenärztliche Vereinigung Bayern and HMOs for patient care; honoraria for consulting, project management, clinical studies, and lectures and from Biogen, Lilly, MedDay, Merck, NeuroTransData, Novartis, Roche, and Thieme Verlag; honoraria and expense compensation as board member of NeuroTransData. KW, FR, and HD are employed by the company NeuroTransData GmbH. VT, SP, and PvH. are employees of PricewaterhouseCoopers AG. The funders had the following involvement with the study: decision to publish, preparation of the manuscript.

DATA AVAILABILITY

Please contact the corresponding author regarding access to data from the registry.

REFERENCES

1. Cline Amin S. Can Real-World Evidence Transform Healthcare? Recent FDA Activities Indicate Yes. https://www.clinicalleader.com/doc/can-real-world-evidence-transform-healthcare-recent-fda-activities-indicate-yes-0001. Accessed December 17, 2021.
2. Flynn R, Plueschke K, Quinten C, *et al*. Marketing authorization applications made to the European Medicines Agency in 2018–2019: what was the contribution of real-world evidence? *Clin Pharmacol Ther* 2022; 111 (1): 90–7.
3. Cave A, Kurz X, Arlett P. Real-world data for regulatory decision making: challenges and possible solutions for Europe. *Clin Pharmacol Ther* 2019; 106 (1): 36–9.
4. Graff J. What the Rise of Real-World Evidence Means for the Pharmaceutical Industry: A Closer Look. ISPOR | International Society For Pharmacoeconomics and Outcomes Research. https://www.ispor.org/publications/journals/value-outcomes-spotlight/vos-archives/issue/view/unlocking-the-promise-of-real-world-evidence/what-the-rise-of-real-world-evidence-means-for-the-pharmaceutical-industry-a-closer-look. Accessed December 17, 2021.
5. Considerations for the Use of Real-World Data and Real-World Evidence to Support Regulatory Decision-Making for Drug and Biological Products; Draft Guidance for Industry. Federal Register. 2021. https://www.federalregister.gov/documents/2021/12/09/2021-26640/considerations-for-the-use-of-real-world-data-and-real-world-evidence-to-support-regulatory. Accessed December 17, 2021.

6. Commissioner of the FDA. Real-World Evidence. FDA. 2021. https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence. Accessed December 17, 2021.

7. European Medicines Agency. Guideline on Registry-Based Studies. European Medicines Agency. 2021. https://www.ema.europa.eu/en/guideline-registry-based-studies-0#current-effective-version-section. Accessed December 17, 2021.

8. Arlett P, Kjær J, Broich K, Cooke E. Real-world evidence in EU medicines regulation: enabling use and establishing value. *Clin Pharmacol Ther* 2022; 111 (1): 21–3.

9. NeuroTransData. About Us. https://www.neurotransdata.com/ueber-uns. Accessed December 19, 2021.

10. Bergmann A, Stangel M, Weih M, *et al*. Development of registry data to create interactive doctor-patient platforms for personalized patient care, taking the example of the DESTINY system. *Front Digit Health* 2021; 3: 32.

11. Braune S, Grimm S, van Hövell P, *et al.*; NTD Study Group. Comparative effectiveness of delayed-release dimethyl fumarate versus interferon, glatiramer acetate, teriflunomide, or fingolimod: results from the German NeuroTransData registry. *J Neurol* 2018; 265 (12): 2980–92.

12. Yadlowsky S, Pellegrini F, Lionetto F, Braune S, Tian L. Estimation and validation of ratio-based conditional average treatment effects using observational data. *J Am Stat Assoc* 2021; 116 (533): 335–52.

13. Stühler E, Braune S, Lionetto F, *et al.*; NeuroTransData Study Group. Framework for personalized prediction of treatment response in relapsing remitting multiple sclerosis. *BMC Med Res Methodol* 2020; 20 (1): 24.

14. De Haes S, Van Grembergen W, Joshi A, Huygh T. COBIT as a framework for enterprise governance of IT. In: De Haes S, Van Grembergen W, Joshi A, Huygh T, eds. *Enterprise Governance of Information Technology: Achieving Alignment and Value in Digital Organizations*. Cham: Springer International Publishing; 2020: 125–62.

15. ISO/IEC 27000:2018(en), Information Technology—Security Techniques—Information Security Management Systems—Overview and Vocabulary. https://www.iso.org/obp/ui/#iso:std:iso-iec:27000:ed-5:v1:en. Accessed December 19, 2021.

16. European Medicines Agency. Multiple Sclerosis Workshop—Registries Initiative. 2018. https://www.ema.europa.eu/en/events/multiple-sclerosis-workshop-registries-initiative. Accessed December 19, 2021.

17. [A19-43] Development of Scientific Concepts for the Generation of Routine Practice Data and Their Analysis for the Benefit Assessment of Drugs According to §35a Social Code Book V—Rapid Report. IQWIG. https://www.iqwig.de/en/projects/a19-43.html. Accessed December 17, 2021.

18. Redaktion Deutsches Ärzteblatt. Randomisierte versorgungsnahe Studien: Gesetzliche Hürden abbauen. Deutsches Ärzteblatt. 2021. https://www.aerzteblatt.de/archiv/221291/Randomisierte-versorgungsnahe-Studien-Gesetzliche-Huerden-abbauen. Accessed December 17, 2021.

19. Bartlett VL, Dhruva SS, Shah ND, Ryan P, Ross JS. Feasibility of using real-world data to replicate clinical trial evidence. *JAMA Netw Open* 2019; 2 (10): e1912869.

20. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *Am J Epidemiol* 2016; 183 (8): 758–64.