The Current Landscape and Emerging Applications for Real-World Data in Diagnostics and Clinical Decision Support and its Impact on Regulatory Decision Making

Elodie Baumfeld Andre[1,*], Nate Carrington[2], Flora S. Siami[3], Jo Carol Hiatt[3], Carly McWilliams[4], Carolyn Hiller[3], Andy Surinach[5], Alejandro Zamorano[6], Chris L. Pashos[5] and Wade L. Schulz[7]

Real-world data (RWD) and real-world evidence (RWE) are becoming essential tools for informing regulatory decision making in health care and offer an opportunity for all stakeholders in the healthcare ecosystem to evaluate medical products throughout their lifecycle. Although considerable interest has been given to regulatory decisions supported by RWE for treatment authorization, especially in rare diseases, less attention has been given to RWD/RWE related to *in vitro* diagnostic (IVD) products and clinical decision support systems (CDSS). This review examines current regulatory practices in relation to IVD product development and discusses the use of CDSS in assisting clinicians to retrieve, filter, and analyze patient data in support of complex decisions regarding diagnosis and treatment. The review then explores how utilizing RWD could augment regulatory body understanding of test performance, clinical outcomes, and benefit-risk profiles, and how RWD could be leveraged to augment CDSS and improve safety, quality, and efficiency of healthcare practices. Whereas we present examples of RWD assisting in the regulation of IVDs and CDSS, we also highlight key challenges within the current healthcare system which are impeding the potential of RWE to be fully realized. These challenges include issues such as data availability, reliability, accessibility, harmonization, and interoperability, often for reasons specific to diagnostics. Finally, we review ways that these challenges are actively being addressed and discuss how private-public collaborations and the implementation of standardized language and protocols are working toward producing more robust RWD and RWE to support regulatory decision making.

Increasingly, real-world data (RWD) and real-world evidence (RWE) are becoming essential tools for regulatory decision making in health care and offer an opportunity for all stakeholders in the healthcare ecosystem (e.g., medical providers, Health Technology Assessment authorities, patient organizations, manufacturers, and regulators) to evaluate medical products throughout their lifecycle, both prior to and following regulatory approval. RWD are the data relating to patient health status and the delivery and outcomes of health care collected from various nonexperimental sources. RWD are collected as part of or in conjunction with the routine care of patients, and can take the form of electronic health records (EHRs), claims and billing documentation, patient, product and disease registries, laboratory results and biomarker data, information on social determinants of health, and information gathered from other sources that can inform on health status, such as wearable or mobile devices.[1] RWE is the clinical evidence regarding the usage and potential benefits or risks of a medical product or other aspects of health care derived from analysis of RWD.[1]

Although much attention has been given to regulatory decisions supported by RWE for treatment authorization, especially in rare diseases, less attention has been given to RWD/RWE related to *in vitro* diagnostic (IVD) products as well as the use of clinical decision support systems (CDSS).

In vitro diagnostics encompass any reagents, instruments, and systems used to diagnose disease or other conditions through the collection, preparation, and examination of specimens taken from the human body.[2] IVDs play a vital role in the healthcare ecosystem, and their importance has been particularly amplified during the coronavirus disease 2019 (COVID-19) pandemic. The Cobas severe acute respiratory syndrome-coronavirus 2 (SARS-CoV-2) and influenza A/B assay, which can detect and differentiate SARS-CoV-2, influenza A, and influenza B viruses in one test,[3] is an example of an IVD.

CDSSs are health information technologies that provide knowledge and person-specific data to clinicians, staff, and patients that has been filtered to provide the relevant information at the appropriate time to enhance healthcare decisions.[4] CDSS can promote significant improvements in the safety, quality, efficiency, and effectiveness of health care due to the variety of tools contained within it that act to enhance decision making in the

[1]Roche Diagnostics, Santa Clara, California, USA; [2]Roche Diagnostics, Indianapolis, Indiana, USA; [3]Medical Device Innovation Consortium, Arlington, Virginia, USA; [4]Roche Diagnostics, Washington, DC, USA; [5]Genesis Research, Hoboken, New Jersey, USA; [6]Datavant, San Francisco, California, USA; [7]Yale School of Medicine, New Haven, Connecticut, USA. *Correspondence: Elodie Baumfeld Andre (elodie.baumfeld_andre@roche.com)

Received February 26, 2021; accepted February 3, 2022. doi:10.1002/cpt.2565

clinical workflow.[4] In the 1980s, clinicians began using CDSS to enhance health care delivery by enhancing medical decisions with targeted clinical knowledge, patient information, and other health information. For example, DXplain was a computer-based CDSS that would accept a list of clinical manifestations and then propose diagnostic hypotheses to support physicians.[5] These systems have seen a rapid evolution in their use to support overall clinical decision making, and, depending on their intended use and country/region of commercialization, may be considered as medical or *in vitro* diagnostic devices.[4]

The use of RWE in regulatory decision making for IVDs has unique considerations, as IVDs have specific regulatory requirements that differ from traditional medical devices (e.g., randomized controlled trials are not usually used for clinical validation of IVD devices; leftover or surrogate samples[6] may be used in many cases). There remains a unique opportunity to develop best practices, tools, and methodologies specifically for the acquisition and appraisal of relevant IVD data generated through routine clinical practice, such that they can be evaluated with scientific rigor and meet the necessary parameters for data quality assurance to support regulatory decisions. However, the quality, reliability, and availability of RWD does not always align with regulatory needs, and thus steps are required to facilitate better access to reliable and appropriate RWD.

This review focuses on current regulatory practices in relation to IVD product development and explores how utilizing RWD could augment regulatory body understanding of a test performance, clinical outcomes, and benefit-risk profiles. Similarly, we will discuss the use of CDSS in assisting clinicians to retrieve, filter, and analyze patient data in support of complex decisions regarding diagnosis and treatment, as well as how RWD may be leveraged to augment CDSS and improve safety, quality, and efficiency of health care practice. Although we present examples of RWD assisting in the regulation of IVDs and CDSS, we also highlight key challenges within current healthcare systems that are preventing more effective use of RWE in this area. Challenges include data availability, reliability, accessibility, harmonization, and interoperability. At the same time, we review ways that these challenges are actively being addressed and discuss how private-public collaborations and the implementation of standardized language and protocols are working toward producing more robust RWD and RWE to support regulatory decision making.

THE STATE OF REGULATORY REQUIREMENTS FOR IVDS AND CDSS

Regulation of IVDs and device CDSS is important to ensure that all products available to patients are safe and effective, fulfill their intended use, and have an appropriate benefit-risk profile. However, obtaining the evidence needed to support robust decision making is often time-consuming and costly.[7] Regulatory decision making is subject to local and regional differences, which may affect the type of RWD needed. In the United States, medical devices, including IVD products, are regulated by the US Food and Drug Administration (FDA), and are classified based on the level of risk associated with them.[8] In the European Union, regulation is slightly different due to the varied responsibilities of the European Medicines Agency (EMA) and the national competent authorities within each of the member countries.[9,10] In China, there is a similar system of classification based on potential risk, and medical device registration is regulated by the National Medical Products Administration (NMPA).[11]

The United States of America

In the United States, the FDA's Center for Devices and Radiological Health (CDRH) is responsible for the regulation of medical devices. The FDA classifies medical devices, including IVD products, into class I, II, or III according to the level of risk and regulatory control necessary to reasonably assure safety and effectiveness, taking into account factors such as the target population for whose use the device is intended and the predicted benefits weighed against the predicted risks (**Table 1**).[8] Non-exempt class I and II devices require a 510(k) submission in order to demonstrate a level of safety and efficacy comparable with those of a similar legally marketed device (a predicate device). A 510(k) should summarize the device characteristics, provide predicate devices to which substantial equivalence is claimed, and describe any nonclinical bench performance tests and analytical studies.[12] Class III devices are the highest risk devices regulated by the FDA. They require Premarket Approval instead of 510(k) clearance and are subject to the most stringent premarket regulatory requirements in the United States.[13] Additionally, for situations where a low or moderate risk medical device has no legally marketed predicate or a new device is not considered substantially equivalent, it must undergo the *de novo* classification process through which it is assessed and granted a relevant class designation.[14] IVDs are defined as devices in §201(h) of the Federal Food, Drug, and Cosmetic Act and, for certain intended uses, may also be considered biological products subject to §351 of the Public Health Service Act. IVDs are generally also subject to categorization under the Clinical Laboratory Improvement Amendments of 1988 (CLIA '88). As with all other medical devices, IVDs are subject to premarket and postmarket controls.[8]

For IVDs, the FDA considers a range of analytical studies during premarket review, including evaluations of accuracy, precision, specificity, and sensitivity.[15] In instances where analytical data are deemed insufficient, determination of substantial equivalence may

Table 1 FDA Medical Device Classification and Approval

Device class	Risk	Application Required for FDA Approval
Class I:		
Exempt	Low	None, but must comply with general controls
Non-exempt	Low	510(k) submission
Class II		
Exempt	Moderate	None, but must comply with general and special controls
Non-exempt	Moderate	510(k) submission
Class III	High	Premarket Approval Application (PMA)

FDA, US Food and Drug Administration.

be achieved by providing clinical performance data. In order for a new device to be considered substantially equivalent to predicate devices, the devices must have the same intended use as the predicate and the same technological characteristics, or the same intended use and different technological characteristics.[15] Clinical performance data could include clinical data establishing that the new IVD falls within the intended use of the predicate device, despite having different indications, such as over-the-counter vs. prescription, or targeted to patients with similar symptoms but different diseases.[15] Similarly, the FDA may request clinical data for new IVDs that use different technologies in order to demonstrate that the performance of the device is equivalent to that of the predicate.[15]

The FDA is currently implementing a national unique device identification (UDI) system which, once fully applied in September 2022, will mean that the labels of most medical devices will possess a UDI, which will improve interoperability across the healthcare system and allow for improved identification of device safety concerns and long-term quality and performance.[16] The plain-text portion of the UDI can be listed as a single line or multiple lines and should be affixed below or near the automatic identification and data capture (AIDC) barcode or other AIDC type. This format is critical because "the easily readable plain-text form allows healthcare professionals (HCPs), patients, the FDA, and other users of the UDI system to read and enter the UDI into data systems, such as patient records or reports to the FDA, without technological assistance."[17]

In the United States, some CDSS are considered as non-devices if they are intended for use by HCPs, but the HCP can independently review the basis for the recommendations, as with software that provide HCPs with current treatment guidelines. CDSS that qualify as software as a medical device (SaMD)[18] are regulated in the same manner as traditional medical devices and IVDs and are classified according to the FDA's three-tiered classification system. However, over the last several years, the Agency engaged in a number of domestic and international efforts to advance regulatory frameworks for SaMD products. In 2013, the International Medical Device Regulators Forum (IMDRF), a group of international medical device regulators, including the FDA, formed the Software as a Medical Device Working Group to provide harmonized guidance to support the innovation and access to effective and safe SaMD.[19] Chaired by the FDA, the SaMD Working Group issued several foundational guidance documents for SaMD regulation, including key definitions,[18] framework for risk categorization,[20] application of a quality management system,[21] and clinical evaluation.[22] Domestically, the FDA has conducted a pilot in relation to a Software Precertification Pilot Program,[23] introduced the concept of predetermined change control plans,[24] and published an AI Action Plan[25] in an effort to advance SaMD regulatory frameworks.

The European Union

The adoption in April 2017 of Regulation (EU) 2017/745 on Medical Devices Regulation (MDR; EU) 2017/746 on *in vitro* diagnostic medical devices regulation (IVDR) changed the EU legal framework for medical devices, introducing new responsibilities for the EMA, for national competent authorities, and for device manufacturers. Both MDR and IVDR entered into force in May 2017 and have a staggered transition period.[9,10] For IVDs, devices are classified as classes A, B, C, and D depending on their intended purposes and inherent risks (**Table 2**). Prior to placing a device on the market, manufacturers must undertake an assessment of the conformity of that device, specifying and justifying the level of clinical evidence necessary to demonstrate conformity with the relevant general safety and performance requirements.[9] Manufacturers must also plan, conduct, and document a performance evaluation that demonstrates scientific validity (the association of an analysis with a clinical condition or physiological state), analytical performance (e.g., analytical specificity, bias, precision, accuracy, etc.) and clinical performance (e.g., diagnostic sensitivity, diagnostic specificity, positive/negative predictive value, etc.). This performance evaluation must be updated throughout the lifecycle of the device, and manufacturers are required to maintain and update a postmarket surveillance system in order to assess the quality, performance, and safety of a device throughout its entire lifetime.[9]

Table 2 EU IVD Device Classification and Approval

Device class	Risk	Application required EU approval
Class A	Low	Self-assessment only
Class B	Moderate	Notified body approval
Class C	High	Notified body approval
Class D	Very high	Notified body approval

EU, European Union; IVD, *in vitro* diagnostic.

Under the IVDR, devices must be labeled with a UDI consisting of both a device identifier (UDI-DI), specific to the manufacturer and device, and a production identifier (UDI-PI), identifying the unit of device production, which should enable unambiguous device identification and increase traceability.[9,10] Additionally, the European Commission implemented a new European Databank on Medical Devices (EUDAMED) to strengthen market surveillance and transparency for competent authorities and to disseminate information to the general public.[26]

In the European Union, CDSS is considered to be a medical device if it has an intended purpose that fulfills the "medical device" definition according to the MDR[9] and IVDR.[10,27] However, there is no specific regulatory term for CDSS or SaMD. Instead, the MDR- and IVDR-related guidance documents use the term medical device software (MDSW).[28] As with the FDA, the European Commission Directorate is a member of the IMDRF.[9,29]

China

The NMPA is responsible for the registration of medical devices in the Chinese market. Medical devices and IVDs are classified as class I, II, or III based on their potential risks (**Table 3**). When registering a device in China, a submission must include product risk analysis data, technical product requirements, a product inspection report, clinical evaluation data, a product manual, label sample draft, quality management system documents related to product development and production, and any other materials

Table 3 Filing and Registration of Chinese Medical Device Products

Device class	Risk	Application required NMPA approval
Class I	Low	Filing materials submitted to the department in charge of drug supervision and management of the municipal people's government with districts (Municipal Bureau)
Class II	Moderate	Registration application materials submitted to the drug regulatory department of the people's government of the province, autonomous region, or municipality directly under the Central Government (Provincial Bureau)
Class III	High	Registration application materials submitted to the drug regulatory authority of the State Council (NMPA)

NMPA, National Medical Products Administration.

required to prove product safety and effectiveness.[11] In terms of SaMD regulation, the NMPA regulates SaMD products in the same manner as traditional medical devices and is also a member of the IMDRF.[30]

In 2019 through 2020, the NMPA ran a UDI pilot program in collaboration with the National Health Care Security Administration (NHCSA) as well as the National Institutes of Medical Device Standards Management. This included three initial pilot groups: medical device manufacturers, distributors, and users.[31] Following this pilot, the Center for Medical Device Evaluation (CMDE) announced that compliance with UDI requirements will now be enforced; applicants with devices listed under the medical device categories provided by the CMDE must now complete and upload UDI-DI information when submitting a registration, renewal, or modification application for their device.[32]

To summarize, in the United States, IVDs are regulated by the FDA, which classifies them into class I, II, or III based on whether they have low, moderate, or high risk, with each class requiring different applications in order to gain the FDA's approval. In the European Union, the main regulatory body is the EMA, which classifies devices into class A, B, C, or D based on whether they have a low, moderate, high, or very high risk. In China, devices are regulated by the NMPA, using a similar classification system to the United States, which places devices into class I, II, or III depending on their risk category. The FDA, the EMA, and the NMPA all require manufacturers to provide evidence that their device conforms to the requirements for safety and performance before it can be accepted to the market. In addition, all three regions are seeing an increasing implementation of UDIs to improve interoperability across healthcare systems and enable postmarketing surveillance.

THE ROLE OF RWD IN MEDICAL DEVICE REGULATORY DECISIONS

Current Guidance for Utilizing RWD for IVDs and CDSS

The FDA's recent publication "Examples of Real-World Evidence (RWE) Used in Medical Device Regulatory Decisions" includes 90 publicly available examples from 2012 to 2019 of RWE use in medical device regulatory decisions,[33] however, only eight examples provided in this document relate to IVDs. Furthermore, although the same criteria of RWE acceptability were applied, most examples included are from prior to the issuance of the FDA CDRH Guidance, "Use of Real-World Evidence to Support Regulatory Decision-Making for Medical Devices: Guidance for Industry and Food and Drug Administration Staff," (hereinafter "FDA RWE Medical Device Guidance").[29,33] Nonetheless, the document provides an important reference point for leveraging RWE in regulatory submissions.

The Medical Device Innovation Consortium (MDIC) *In Vitro* Diagnostic Real-World Evidence Framework (hereinafter "IVD RWE Framework") was developed following the release of the FDA RWE Medical Device Guidance to provide additional contextual information as issues pertain to clinical validation of RWD in premarket and postmarket regulatory decision making of IVDs more specifically.[29,34] The FDA Medical Device RWE Guidance focuses on the use and potential value of RWE to support regulatory decision making for medical devices and was informed by an analysis of current and historical experiences.[29] The IVD RWE Framework describes when and how appropriate study designs and analytical methods may be applied to relevant and reliable RWD to generate valid scientific RWE to inform or augment regulatory decisions in support of clearance or approval of an IVD.[34] It discusses general study design and methodology considerations when making regulatory decisions based on data generated from a routine clinical environment, including addressing data quality issues (e.g., missing data, transparency in data generation processes, and systematic biases) and approaches to establishing criteria for data evaluation in a benefit/risk context.[34] The IVD RWE Framework also describes both hypothetical and previously approved regulatory case examples of applications of RWE in the IVD setting, such as when data for clinical validation may be difficult or burdensome to collect.[34]

The introduction of RWD into Chinese IVD clinical evaluation is relatively new in contrast to the FDA, and as such there is limited information and few examples of use cases available. However, in November 2020, the NMPA published the document "Announcement on the Publication of the Technical Guidelines for the Use of Real-World Data in Clinical Evaluation of Medical Devices (Trial)" with the aim of normalizing and guiding the application of RWD to support medical device regulation. The guidance provides information on the advantages and limitations of RWD sources and discusses evaluation of RWD quality.[35] The trial implementation of these guidelines is a promising step toward implementation of regular RWD use to support future medical device regulation in China.

The guidance for RWD for CDSS is much less robust than for IVDs. The draft guidance from the FDA is primarily focused on defining what constitutes a CDSS.[17] Generally, countries defer to IMDRF guidance,[22] which briefly describes how RWD may be leveraged to monitor software safety, effectiveness, and performance.

IVD submissions supported by RWD

In this section, we introduce existing IVD submissions that were supported by RWE. One such submission is DEN170058, which relates to the MSK-IMPACT assay indicated as a next-generation sequencing-based tumor profiling test, was supported by clinical data from an electronic medical record database of patients with advanced cancer as part of routine workflow at Memorial Sloan Kettering Cancer Center. Retrospective analysis of these records provided evidence to support a pan-cancer claim, to validate a test cutoff, and to provide data on somatic mutation prevalence.[33] Submission number P160052 relates to the Placental Alpha Microglobulin-1 Immunoassay and encompasses a total-product lifecycle example supported by clinical evidence in the form of patients' medical records. The sponsor submitted an observational clinical utility study of patients tested using the assay for premarket clinical evidence and as a condition-of-approval.[33] Submission number P140020 involves a germline gene mutation test (trade name BRACAnalysis CDx). As a condition-of-approval, the company (Myriad Genetic Laboratories) is required to collect, in a sponsor database, data regarding all IVD results during commercial use.[33] A personal genome service from 23andMe (submission number DEN160026) supported a *de novo* classification request using peer-reviewed, real-world literature as a primary source of clinical evidence for each of the 10 conditions included in the Genetic Health Risk tests. Information from the CFTR2 database, a publicly maintained next generation sequencing database, was used as the sole source of clinical evidence supporting a 510(k) for both the Illumina MiSeqDx Cystic Fibrosis Clinical Sequencing Assay and the Illumina MiSeqDx Cystic Fibrosis 139-Variant Assay (Illumina, Inc. submission numbers K132750 and K124006).[33] An example of premarket Pediatric RWE use is the SEEKER System (Baebies, Inc. submission number DEN150035), which was supported by a pivotal trial embedded in a state-run routine screening program, the Missouri State Public Health Laboratory, and Missouri Department of Health and Senior Services (MDHSS) Surveillance Program.[33]

CDSS submissions supported by RWD

RWD offers the potential for the FDA and other international regulatory bodies to augment their understanding of CDSS benefit-risk profiles and could even provide new insights into potential clinical outcomes and device performance. Three instances in which RWD has been used to support premarket regulatory decision making for digital health technologies, including CDSS, are provided in the FDA's "Examples of Real-World Evidence (RWE) Used in Medical Device Regulatory Decisions."

Specifically, PeraHealth Inc. leveraged data from retrospective medical records of adult and pediatric patients to demonstrate the safe and effective performance of the PeraTrend System (submission number K172959).[33] In another example, a radiological computer-aided triage/notification system (ContaCT, Viz AI, Inc., submission number DEN170073) used standard-of-care notification time extracted from radiologist reports against a comparable metric from standalone testing of the SaMD to support a secondary analysis.[33] Additionally, Natural Cycles Nordic AB (submission number DEN170052) performed a retrospective analysis of 15,000 users of their mobile application as a primary source of clinical evidence to support its *de novo* classification request.[33]

CURRENT CHALLENGES OF RWD FOR IVDS AND CDSS

Despite efforts underway to better facilitate the integration of RWD use in regulatory decision making, utilizing RWE for IVDs and CDSS is not without its challenges. RWD sources should be fit for purpose, and the evidence generated should be robust to drive decision making. The relevance and reliability of RWD, RWD sources, and resulting analyses must be evaluated to determine whether the data either partially or fully addresses the regulatory question, if the data accrual can be adequately confirmed to have minimal errors, and if the data assurance is sufficient.[29]

Data relevance and reliability

The FDA guidance specifies that RWE should be supported by an appropriate study design and execution of statistical RWD analysis plan that addresses (minimizes) differential confounders and baseline risks. However, the data needed for regulatory analysis might not always be available; for instance, RWD could be missing information on important confounders and patient follow-up which, in turn, could impact the ability to predict the risk of potential adverse clinical outcomes.[36] Data to support IVD regulatory decisions can be sourced from EHRs, laboratory information systems, disease and medical device registries, and administrative claims and hospital discharge databases.[29] Each of these data sources have their own advantages and limitations when it comes to acquiring the full information needed for product assessment. For example, laboratory information may not include all the relevant results or any details on modifications made to the test by a clinical laboratory.[36] Moreover, EHR, medical device registries, and administrative claims data may not contain information on patient follow-up. For example, when patients change insurance providers, follow-up is often logged in different claims databases.[36]

Problems with data availability are relevant to the current COVID-19 pandemic: the rapid and severe onset of the disease restricts availability of information related to modes of transmission, clinical presentation, and clinical outcomes postinfection. Compounding these challenges are the additional caveats that not all patients who contract COVID-19 are tested, and those that are may be subject to false-positive or false-negative results due to heterogeneity in the accuracy of different available tests on the market.

The COVID-19 pandemic has also highlighted challenges in data reliability. During the pandemic, there have been several controversies surrounding poor standards in laboratory data reflected by the existence of broad variations in case number estimates and the results of seroprevalence studies. This is often a result of laboratory data being inappropriately used to generate RWE rather than the data themselves. One of the most controversial cases was the Santa Clara County study,[37] which claimed that the true number of COVID-19 cases was 50 to 85 times higher than officially recorded in the Santa Clara area. However, concerns about the testing kits used, recruitment into the study, and statistical treatment of the data compromised the reliability of the conclusions. This highlights the vital importance of ensuring IVD data reliability.[38] A harmonized system for IVDs would provide more information

on the mechanism of data generation for laboratory results and only improve the strengths of upcoming studies. To this end, guidance on the use of RWD is still evolving in order to provide standardization across databases.[39]

Data sharing and accessibility

Data access is another significant issue faced when compiling RWD. Much of the promise of RWE stems from its potential to achieve statistical significance by amalgamating population-based data sets. In countries like the United States, where the culture of privacy protection is strong and public scrutiny over digital privacy protection continues to grow with increasing awareness, the sharing of clinical data to unlock important insights remains limited and difficult. The risk of data theft, manipulation, and other malignant use are becoming more apparent with every news cycle.[40] There are a number of de-identification strategies being explored, but fundamentally may not be possible for every type of data. Strategies to avoid the transfer of actual data, such as synthetic data and federated learning, are being developed but are largely in their infancy.[40] The synthetic data approach enables generation of a realistic synthetic record that has the same statistical and time-dependent properties as the original data, but does not link the data to individuals, leading to full anonymization of the data set.[41] Federated learning is a learning paradigm that trains algorithms collaboratively without actually exchanging the data itself; in a healthcare setting, federated learning is able to form a consensus model without moving patient data beyond institution firewalls.[42] However, tiered access to data seems unavoidable, with many stages of permissioned access likely to come between tightly regulated to widely open access.

Similar challenges exist in the European Union, where the General Data Protection Regulation (GDPR) is in place.[43] The GDPR plays a critical role in protecting the interests of individuals over their data and establishes rules for the processing of personal data (including health data). However, the implementation and the interpretation of this Regulation, including its inconsistent application across EU Member States and the existence of parallel national Member State laws with regard to the processing of health data, create considerable uncertainties which have resulted in barriers to health data use. For example, under the GDPR, medical device and IVD companies face difficulties in accessing and sharing data for secondary use purposes, which can lead to significant challenges in leveraging RWD for regulatory decision making (among other challenges). Such barriers may potentially be addressed through the European Health Data Space (EHDS), a European Commission priority focused on promoting better exchange and access to different types of health data (EHRs, genomics data, data from patient registries, etc.) to support healthcare delivery (primary use of data) and health research and health policy making (secondary use of data).[44] This initiative should aim to establish a firm legal basis for processing health data and provide legal certainty regarding the rules under which health data can be collected, processed, and shared.[45]

In 2017, the Chinese National Scientific Data Sharing Platform for Population and Health (NSDSPPH) released 49.1 TB of clinical data collected from hospitals, which is now available to the public and researchers.[46] However, data accessibility is still a limitation to RWD utilization in China, with EHRs primarily used for clinical practice and thus largely containing unstructured data that is not required to be shared with other healthcare systems.[47] Another limitation of data sharing in China is that patient privacy is jeopardized by the fact that there is no clear regulation for de-identifying health system data. A 2020 study using patient data from Chinese hospitals evaluated privacy risks based on the US Health Insurance Portability and Accountability Act (HIPAA) safe harbor and limited data set policies. They reported that patient data might be vulnerable to re-identification risks under different policies and lay the foundation for further privacy risk studies using Chinese patient data.[48]

Data fragmentation

Patient interactions with the healthcare system generate large volumes of data, which are stored in disconnected systems. Fragmented RWD is a key barrier in generating RWE for IVDs and CDSS. Fragmentation of RWD provides an incomplete picture of a patient's healthcare journey, leading to gaps in RWE. For example, due to the decentralized nature of the US healthcare system, data on a patient's healthcare journey are distributed across many different data stewards. A patient's healthcare data can be scattered across any clinical trial(s) in which they participated, the EHR of the hospital(s) the patient visited, the prescriptions that they picked up at one or more pharmacies, the laboratory test(s) they received in the outpatient clinic, and their wearable devices.

Further complicating matters is that international privacy protections require that patient-level data can only be shared on a de-identified basis. As a result, before exchanging healthcare data, RWD sources are required to remove all Personally Identifiable Information and represent them with an ID of some form. This is problematic in that these IDs are different for the same individual across RWD sources, making the linking of data across RWD sources difficult.[49]

For example, in traditional clinical studies, participants meeting the eligibility criteria during screening are enrolled at individual sites. As they enroll, participants give consent for their data to be used for specific purposes related to the study. They are then randomly assigned to an arm of the study. In many studies, this process is designed to be "blinded," meaning that neither the participants nor the investigators know if the participant is receiving the treatment, a comparator, or a placebo. In order to protect the integrity of the study, each participant's identifying information is removed and replaced with a subject ID. Investigators use the subject ID to track all of the patient's site visits and to create a longitudinal record. Once this de-identification occurs, the sponsor might be left with a very expensive siloed data set that cannot be linked to other RWD sources.[50]

Interoperability

The most significant and problematic challenges hampering health data interoperability are the nonstandardized approaches to data collection and lack of an industry-wide interoperability measurement standard. The general problems with RWD collection of IVDs are

that test results are not always included in routine claims databases or are provided in an unstructured format, such as free text added to the record. Moreover, reimbursement codes may be "cross-walked" from one version to another, which means information on specific tests may be missing.[36] A key challenge of laboratory interoperability is the limited use of terminology standards that would reduce the time spent exchanging, tracking, and reporting tests. Translating observational data findings from a provider, laboratory data, and other diagnostic information to standard terminologies would allow them to be understood by all information systems.[51]

In the United States, the Interoperability Standards Advisory (ISA) process is designed to bridge interoperability gaps in the healthcare industry by coordinating the identification, assessment, and public awareness of interoperability standards and implementation specifications.[52] However, the adoption of standard and implementation specification orders related to IVDs have been minimal to date.[53] In the European Union, the ISA[2] running from 2016 through 2020 offered a similar facility of interoperability solutions for public administrations, businesses, and citizens, aiming to implement interoperable cross-border and cross-sector public services.[54] However, an evaluation of the program recognized awareness-raising beyond national administrations, transitioning from user-centric to user-driven solutions, and preservation of the program's achievements as areas in need of improvement.[55] The eHealth Action Plan 2012–2020 was designed to implement widespread use of eHealth across the European Union, but barriers to deployment included lack of awareness of, as well as limited confidence in eHealth among HCPs and patients, lack of interoperability, and regional differences in the access to information and communication technology services, particularly in deprived areas.[56] The eHealth system utilizes information and communication technology in healthcare systems and covers the interaction between patients and health-service providers, institution-to-institution transmission of data, or peer-to-peer communication between patients and/or HCPs.[56]

China has a significant problem with data interoperability, as adoption of individual EHRs is hindered by incompatibility between different hospital systems; over 300 commercial providers of hospital information systems exist in China, and each use different technical structures and data standards. Healthcare systems are not required to exchange data with each other, and there is a distinct lack of consistent medical terminology.[47] In 2002, use of the International Classification of Diseases Revision 9 (ICD-9, and more recently ICD Revision 10 (ICD-10)) was mandated by the National Health and Family Planning Commission for all hospital patients, but the coding of other clinical terms beyond diagnosis varies greatly among hospital information systems and thus makes data exchange difficult.[47]

OPPORTUNITIES FOR RWD USE TO SUPPORT IVD AND CDSS REGULATION AND INNOVATION

RWD are a source of data containing a wealth of information that can be used to support medical device regulatory decisions. The FDA has already demonstrated its willingness to leverage RWD in premarket submission determinations, as indicated in the previously cited examples, and use of RWD for regulatory decision making supports the FDA's mission to help patients gain timely access to high-quality, safe, and effective medical devices.[57] Supporting these efforts, pending legislation in the US Senate seeks to enable more efficient use of RWD and RWE.[58] But to further achieve this goal, the FDA and other regulatory authorities must continually improve and increase the efficiency of RWD regulatory processes.

Promoting RWD value to evolve methodology and improve fit-for-purpose data availability

One approach is to apply guiding principles that are "least burdensome," which is defined by the FDA as "the least amount of information necessary to adequately address a relevant regulatory question or issue through the most efficient manner at the right time."[57] Similarly, in the European Union, more significant premarket and postmarket regulatory requirements under the MDR and IVDR have created opportunities for more widespread use of RWD for regulatory decision making. With these definitions and foundational principles in mind, depending on the intent, RWD could offer a less burdensome source of data as compared with traditional clinical studies as the data are generated during typical healthcare activities.

For instance, when designing medical device clinical studies, prospective enrollment of subjects is the traditional way to acquire a required sample size. If the disease or condition of interest has a low prevalence, then many patients would need to be screened, thus incurring high associated costs and demands on the time of investigators, patients, and the overall study period. However, by leveraging data about subjects from RWD sources, it is possible to augment data collection for traditional studies to reduce the length and costs of clinical studies. In some instances, RWD may serve as the sole clinical study data source.

In a recent study, Chen *et al*. introduced a statistical method to potentially expedite clinical studies through the use of RWD in situations where there is a low disease prevalence, which serves as a strong example of the possibilities for leveraging RWD.[59] By using the propensity score methodology defined in the study, Chen *et al*. demonstrated that it was possible to define a group of real-world subjects who matched their prospectively enrolled counterparts and confirmed that the model was transferable to the evaluation of any diagnostic test.[59]

Tokenization to combat data fragmentation

To solve the data fragmentation challenges, one strategy is to use a common token across various RWD sources. Tokenization utilizes the underlying identifying information in a dataset and uses encryption to create a de-identified, unique, and irreversible alphanumeric string that is unique to a patient. This token, unlike an ID, is the same for the same patient across all RWD sources. As a result, the token can be used as the key to link disparate RWD sources together, resulting in a fuller picture of a patient's healthcare journey.[60]

In the context of clinical trials, tokens can be used by sponsors to link their clinical trial data to other RWD sources resulting in exciting use cases such as:

- Retrospective Subcohort Analysis: When reviewing study data, investigators often identify a particular subcohort of patients that responded better to the treatment than others, or that might have experienced more adverse side effects. By linking

in EHRs, historical diagnoses, and demographic data, investigators can gain further insight into what factors might have driven differential response to treatment.[61]

- Long-Term Surveillance Monitoring: In other cases, sponsors may want or be required to track patients for a number of years after a study is completed, seeking to better understand long-term outcomes and ensure there are no unanticipated safety events when the treatment is deployed in the real world. To do so, they could link their clinical trial data to EHRs and claims data.[62]

Standardization to improve interoperability

Another important consideration related to the use of RWE in regulatory decision making for IVDs is that such RWD use is only possible if the data is interoperable, of sufficient quality, and able to be extracted and presented in a structured format, which, for example, could be achieved using natural language processing.[63] As such, it is vital that healthcare systems, electronic medical software companies, and the IVD industry adopt a common data model for IVD test results. To achieve this, we recommend that stakeholders implement a digital format consisting of the data elements described in the LIVD (Logical Observations Identifiers Names and Codes (LOINC) for IVD)[64] digital format proposal (with the addition of data elements for identifying IVD instrument serial number and IVD assay lot number) and UDI. Such an approach will enable unambiguous identification of IVD test result information to facilitate routine collection of IVD RWD. This will benefit IVD manufacturers and regulators in relation to regulatory decision making, and it will also benefit patients, HCPs, health care institutions, payers, and multiple other stakeholders by enabling greater evidence generation and insight into the real-world performance of these products.

Private-public partnerships to support RWD utilization

The National Evaluation System for health Technology Coordinating Center (NESTcc) was established in 2016 by the MDIC to support the sustainable generation and use of timely, reliable, and cost-effective RWE to support the entire medical device lifecycle. In 2018 and 2019, the NESTcc announced 21 test-case projects that reflect the diversity of the types of medical devices available and their different uses, following them throughout the medical device Total Product Life Cycle.[65] The objective of examining these test-case projects is to explore the feasibility for medical device ecosystem stakeholders working with RWD sources, and to help identify areas where NESTcc could assist with reducing transaction costs, such as data sharing agreements and public policies.

Another NESTcc initiative is an evaluation of the uptake of UDIs by health systems, which seeks to understand current implementation and use of UDIs in clinical care in order to identify potential barriers to more consistent deployment.[66] Use of UDIs across healthcare systems would greatly improve interoperability and therefore increase the availability of high-quality RWE available for medical devices and permit long-term safety surveillance. Once UDIs are well integrated into RWD, they will enable RWE to be generated from a wider number of sources that can support the total evidence required for initial approvals and label expansions. Additionally, broad UDI adoption will allow device manufacturers to link clinically meaningful outcomes related to use of their devices in order to generate RWE to demonstrate the value of their product to payers.

In addition to UDIs, efforts directed at solving the challenges of RWD use in RWE and to promote accessibility, interoperability, and linkage in defragmentation of stakeholder data sources include several collaborative working groups. One example is NEST, which the FDA played a significant role in the establishment of, and is designed to generate better evidence more efficiently across the total product lifecycle of a medical device through the use of RWE and advanced analytics. The impetus for NEST was primarily the FDA's interest in improving the quality of available RWE so that it can be used to better inform HCPs and patients about the devices available to them.[67]

Another example of a public-private partnership created to introduce laboratory information interoperability is the Systemic Harmonization and Interoperability Enhancement for Lab Data (SHIELD) initiative, which developed (LOINC) that will be consistently linked to the same type of IVD by manufacturers, laboratories, and HCPs.[68] Michael Waters, PhD, succinctly described SHIELD's principles as "Simply, describing the same test the same way. By improving the semantic interoperability of laboratory data within and between institutions, diagnostic information can be used to better support clinical decisions."[69]

Whereas attempts to improve interoperability are ongoing, emergency situations reinforce the value and urgency of being able to use RWE to drive the adoption of healthcare data interoperability. The COVID-19 pandemic has motivated the fast-tracking of data availability/interoperability internationally. The Interoperability Proving Ground (IPG) is an open community platform designed to allow countries to share information on interoperability projects. To date, there are 96 project entries related to the COVID-19 Novel Coronavirus Pandemic covering topics such as interoperability of patient care, clinical workflow tools, and designing patient care plans.[70] Created in the United States, the majority of IPG projects stem from numerous states across the country, but there are additional projects from Canada, Australia, the Netherlands, Ukraine, and Finland.[70] Although this is far from being a strong international interoperability platform, it demonstrates an opportunity and an interest in facilitating healthcare harmonization at an international level.

Within the United States itself, the recent cooperative arrangement agreed between competing HCPs HealthPartners and Allina Health[71] represents an exciting example of interoperability between two companies that could broaden the scope and quality of available RWD. Progress toward a connected healthcare system is slow, but the COVID-19 pandemic highlighted the detrimental impact that a lack of interoperability can have on healthcare provision. In a pandemic situation, where an urgent care response is required, the healthcare system is severely hindered by a lack of harmonization resulting in inconsistent statistics, and the requirement for time-consuming manual data processing, which is not appropriately reactive to healthcare demands.[51]

The MDIC is a public-private partnership organization in the United States that facilitates collaboration among industry, government, and non-profits to advance medical device regulatory science

for patients. MDIC's IVD RWE Framework focuses on issues related to the clinical validation of RWD in pre- and postmarket regulatory decision making and assists stakeholders in identifying appropriate RWD/RWE that is "fit-for-purpose."[34] The MDIC is leading the "Open Hand" process to provide a transparent process to evaluate new technology and methods, in particular, converting emergency use authorizations for a COVID-19 diagnostics to full 510(k) or *de novo* submissions.[72]

One of the main challenges encountered when collecting RWD, which sometimes deters researchers from using RWD, is the potential for variation in the language used to report health measurements and observations. Using traditional research methods, such as randomized clinical trials (RCTs) synergistically with RWE requires researchers to consider hybrid research approaches, such as pragmatic trials or cluster-randomized designs.[40] These approaches too, have challenges because they require the engagement of the health system and health providers, incorporation of disparate RWD sources, and analysis techniques that may be more complicated than those traditionally used in RCTs. However, these approaches can achieve the same level of rigor of an RCT, reduce patient burden, and achieve cost and time savings. For this to be effective, clinical researchers and regulators alike need support, training, and tools to choose appropriate data sources and methodologies to facilitate their use of RWE and hybrid methods.

A final collaboration example is the IMDRF, introduced earlier in this review as a group of international medical device regulators, who have a mission to strategically accelerate international medical device regulatory convergence to promote an efficient and effective regulatory model for medical devices that is responsive to emerging challenges in the sector while protecting and maximizing public health and safety.[73] As such, it is important for medical device manufacturers and regulators to come together to address the challenges and opportunities identified in this paper. Principles such as regulatory agility and harmonized regulatory approaches will advance regulatory science in relation to the use of RWD in regulatory decision making. This will transform the healthcare sector and lead to new insights and improvements in patient care, during both "normal" times and times of crisis.

CONCLUSIONS

RWD holds great potential to facilitate rapid deployment of IVDs and CDSS, support new approvals, lead to expanded device indications, support label enhancements, and fulfill postmarketing monitoring requirements to assess safety with improved accuracy in larger populations. However, in order to reach these goals, there needs to be concerted effort among researchers, developers, and regulators to enhance data harmonization and interoperability. Researchers and developers must thoughtfully design real-world studies with appropriate analytical and statistical methods. Unfortunately, the challenges with data exchange and interoperability are significantly limiting for IVD and CDSS; however, there are clear steps that can be taken to address these challenges and unleash the use of RWE for IVDs and CDSS. Through the use of common data standards, methods, and nomenclature, as well as widespread implementation of UDI and LIVD, IVD instrument serial numbers and assay lot numbers, there can be unambiguous identification of IVD test results and significant improvement in the quality and usability of IVD RWD. Additionally, tokenization could help combat data fragmentation, which is a core barrier to generating RWE for IVDs and CDSS. These steps, in combination with the continuation of private-public collaborations, such as NEST, SHIELD, and MDIC, will drive forward the traceability, transparency, patient protection, and privacy of RWD for IVDs and CDSS.

ACKNOWLEDGMENTS

This study was funded by Roche Diagnostics. Eve Hunter-Featherstone, PhD and Amy Glenwright, PhD from Genesis Research, LLC provided editorial and writing assistance funded by Roche Diagnostics in accordance with GPP3.

FUNDING

This review was funded by Roche Diagnostics.

CONFLICTS OF INTEREST

All authors report no conflicts of interest.

1. FDA. Real-World Evidence <https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence>. Accessed May 24, 2021.
2. FDA. In Vitro Diagnostic Products for Human Use. 21 CFR § 80932020.
3. Diagnostics, R. Cobas® SARS-CoV-2 & Influenza A/B Assay Rapid answers to manage the unknown. Updated June 16, 2021 <https://diagnostics.roche.com/global/en/products/params/cobas-sars-cov-2-influenza-a-b-nucleic-acid-test.html>. Accessed June 16, 2021.
4. HealthIT.gov. Clinical Decision Support <https://www.healthit.gov/topic/safety/clinical-decision-support>. Accessed May 24, 2021.
5. Barnett, G.O., Cimino, J.J., Hupp, J.A. & Hoffer, E.P. DXplain: An evolving diagnostic decision-support system. *JAMA* **258**(1), 67–74 (1987).
6. MDIC. *Surrogate Sample Framework – A report of the Surrogate Sample working group of the Medical Device Innovation Consortium (MDIC)*. MDIC <https://mdic.org/resource/surrogate-sample-framework/> (2017).
7. Hampson, G., Towse, A., Dreitlein, W.B., Henshall, C. & Pearson, S.D. Real-world evidence for coverage decisions: opportunities and challenges. *J. Comp. Eff. Res.* **7**(12), 1133–1143 (2018).
8. FDA. Medical Device Classification Procedures. 21 CFR § 8603 2020.
9. EU. Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC. *Official Journal of the European Union L117* **60** (2017) <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32017R0745>.
10. EU. Regulation (EU) 2017/746 of the European Parliament and of the Council of 5 on in vitro diagnostic medical devices and repealing Directive 98/79/EC and Commission Decision 2010/227/EU. *Official Journal of the European Union L117* **60** (2017) <https://eur.lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:02017R0746-20170505>.

11. NMPA. Regulations on the Supervision and Administration of Medical Devices <https://www.nmpa.gov.cn/xxgk/fgwj/flxzhfg/20210319202057136.html>. Accessed May 11, 2021.
12. FDA. Establishment Registration and Device Listing for Manufacturers and Initial Importers of Devices. 21 CFR § 807922020.
13. FDA. Premarket Approval (PMA) <https://www.fda.gov/medical-devices/premarket-submissions/premarket-approval-pma>. Accessed June 17, 2021.
14. FDA. De Novo Classification Request <https://www.fda.gov/medical-devices/premarket-submissions/de-novo-classification-request>. Accessed June 17, 2021.
15. The 510(k) Program: Evaluating Substantial Equivalence in Premarket Notifications [510(k)] (Center for Devices and Radiological Health) (2011).
16. Daniel, G., McClellan, M., Gardina, S., Deak, D., Byran, J. & Streit, C. *Unique Device Identifiers (UDIs): A Roadmap for Effective Implementation*. (The Brookings Institution, Washington, DC, 2014).
17. FDA. Unique Device Identification System: Form and Content of the Unique Device Identifier (UDI) (2021).
18. Group ISW. Software as a Medical Device (SaMD): Key Definitions. In: Forum IMDR, editor. (2013).
19. FDA. Software as a Medical Device (SaMD) <https://www.fda.gov/medical-devices/digital-health-center-excellence/software-medical-device-samd>. Accessed June 30, 2021.
20. Group ISW. "Software as a Medical Device": Possible Framework for Risk Categorization and Corresponding Considerations. In: Forum IMDR, editor. (2014).
21. Group ISW. Software as a Medical Device (SaMD): Application of Quality Management System e International Medical Device Regulators Forum (2015).
22. Group ISW. Software as a Medical Device (SaMD): Clinical Evaluation. In: Forum IMDR, editor. (2017).
23. FDA. Developing a Software Precertification Program: A Working Model.
24. FDA. Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML) - Based Software ad a Medical Device (SaMD) <https://www.fda.gov/files/medical%20devices/published/US-FDA-Artificial-Intelligence-and-Machine-Learning-Discussion-Paper.pdf2019>.
25. FDA. Artificial Intelligence/Machine Learning (AI/ML) - Based Software as a Medical Device (SaMD) Action Plan.
26. European Databank on Medical Devices Commission– (EUDAMED) <https://ec.europa.eu/health/md_eudamed/overview_en>. Accessed June 18, 2021.
27. FDA. Clinical Decision Support Software: Draft Guidance for Industry and Food and Drug Administration Staff. (2019).
28. MDCG. MDCG 2019-11 Guidance on Qualification and Classification of Software in Regulation (EU) 2017/745 – MDR and Regulation (EU) 2017/746 – IVDR (2019).
29. FDA. Use of Real-World Evidence to Support Regulatory Decision-Making for Medical Devices. (2017).
30. FDA. International Medical Device Regulators Forum (IMDRF) <https://www.fda.gov/medical-devices/cdrh-international-programs/international-medical-device-regulators-forum-imdrf>. Accessed June 30, 2021.
31. Eisenhart, S. Chinese regulators preparing medical device UDI pilot <https://www.emergobyul.com/blog/2019/07/chinese-regulators-preparing-medical-device-udi-pilot>.
32. Herr, T. Chinese regulators announce classification adjustments and UDI submission guidelines <https://www.emergobyul.com/blog/2021/01/chinese-regulators-announce-classification-adjustments-and-udi-submission-guidelines>.
33. FDA. Examples of Real-World Evidence (RWE) Used in Medical Device Regulatory Decisions (2019).
34. MDIC. Real-World Clinical Evidence Generation: Advancing Regulatory Science and Patient Access for In Vitro Diagnostics (IVDs) (2020).
35. NMPA. *Announcement of the State Food and Drug Administration on the Technical Guidelines for the Release of Real-World Data for the Clinical Evaluation of Medical Devices (for Trial Implementation).* State Food and Drug Administration. (2021) <https://www.nmpa.gov.cn/xxgk/ggtg/qtggtg/20201126090030150.html>.
36. O'Neill, T. *et al*. ISPOR, the FDA, and the evolving regulatory science of medical device products. *ISPOR* **22**(7), 754–761 (2019).
37. Bendavid, E. *et al*. COVID-19 antibody seroprevalence in Santa Clara County, California. *Int. J. Epidemiol.* **50**(2), 410–419 (2021).
38. The Scientist. How (Not) to Do an Antibody Survey for SARS-CoV-2 <https://www.the-scientist.com/news-opinion/how-not-to-do-an-antibody-survey-for-sars-cov-2-67488>. Accessed May 24, 2021.
39. Mullins, D.C. Evolving use of real-world evidence for devices: good for patients, Good for Policy Makers. *Value Health* **22**(7), 751–753 (2019).
40. Rudrapatna, V.A. & Butte, A.J. Opportunities and challenges in using real-world data for health care. *J. Clin. Investig.* **130**(2), 565–574 (2020).
41. Reiner Benaim, A. *et al*. Analyzing medical research results based on synthetic data and their relation to real data results: systematic comparison from five observational studies. *JMIR Med. Inform.* **8**(2), e16492 (2020).
42. Rieke, N. *et al*. The future of digital health with federated learning. *NPJ Digit. Med.* **3**, 119 (2020).
43. EU. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). In: Union OJotE, editor. 2016.
44. European Commission. European Health Data Space <https://ec.europa.eu/health/ehealth/dataspace_en>.
45. MedTech Europe. Unlocking the full benefits of health data Recommendations from MedTech Europe <https://www.medtecheurope.org/wp-content/uploads/2021/06/mte_160621_positionpaper_health-data.pdf>.
46. Zhang, Z. Release of the national healthcare big data in China: a historic leap in clinical research. *AME Med. J.* **2**(2), 19 (2017).
47. Zhang, L., Wang, H., Li, Q., Zhao, M.H. & Zhan, Q.M. Big data and medical research in China. *BMJ*. **360**, j5910 (2018).
48. Gong, M. *et al*. Evaluation of privacy risks of patients' data in China: case study. *JMIR Med Inform*. **8**(2), e13046 (2020).
49. Morris, G., Farnum, G., Afzal, S., Robinson, C., Greene, J. & Coughlin, C. Patient Identification and Matching Final Report. HealthIT.gov <https://www.healthit.gov/sites/default/files/patient_identification_matching_final_report.pdf>.
50. Tucker, K. *et al*. Protecting patient privacy when sharing patient-level data from clinical trials. *BMC Med. Res. Methodol.* **16**(S1) (2016). https://doi.org/10.1186/s12874-016-0169-4.
51. Futrell, K. COVID-19 highlights need for laboratory data sharing and interoperability. Medical Laboratory Observer <https://www.mlo-online.com/information-technology/lis/article/21210723/covid19-highlights-need-for-laboratory-data-sharing-and-interoperability>. Accessed May 24, 2021.
52. ONC. About the ISA <https://www.healthit.gov/isa/about-isa>.
53. ONC. Exchanging InVitro Diagnostics (IVD) Test Orders & Results <https://www.healthit.gov/isa/exchanging-invitro-diagnostics-ivd-test-orders-results>.
54. European Commission. ISA² - Interoperability solutions for public administrations, businesses and citizens <https://ec.europa.eu/isa2/isa2_en>. Accessed July 2, 2021.
55. European Commission. Interim evaluation concluded good performance of ISA² Programme <https://ec.europa.eu/isa2/news/interim-evaluation-concluded-good-performance-isa%C2%B2-programme_en>. Accessed July 2, 2021.
56. European Commission. eHealth Action Plan 2012-2020 - Innovative healthcare for the 21st century 2012.
57. FDA. The Least Burdensome Provisions: Concept and Principles (2019).
58. Marshall, R. Text - S.1508 - 117th Congress (2021-2022): A bill to provide for the use of emergency use authorization data and real world evidence gathered during an emergency to support premarket applications for drugs, biological products, and devices, and for other purposes. legislation. 2021/04/29/ (2021).

59. Chen, W.C. *et al*. Evaluation of diagnostic tests for low prevalence diseases: a statistical approach for leveraging real-world data to accelerate the study. *J Biopharm Stat* **31**, 375–390 (2021).
60. Burcu, M., Manzano-Salgado, C., Butler, A. & Christian, J. A Framework for extension studies using real-world data to examine long-term safety and effectiveness. *Ther. Innovation Regul. Sci.* **56**, 15–22 (2022).
61. Bolislis, W.R., Fay, M. & Kuhler, T.C. Use of real-world data for new drug applications and line extensions. *Clin. Ther.* **42**(5), 926–938 (2020).
62. Katkade, V.B., Sanders, K.N. & Zou, K.H. Real world data: an opportunity to supplement existing evidence for the use of long-established medicines in health care decision making. *J. Multidiscip. Healthc.* **11**, 295–304 (2018).
63. Chen, B., Dolatabadi, E., Kwong, J. & Azimaee, M. Deep Learning and NLP for knowledge extraction from laboratory reports. *Int. J. Population Data Sci.* (2020). https://doi.org/10.23889/ijpds.v5i5.1637.
64. Centers for Disease Control and Prevention (CDC). LOINC In Vitro Diagnostic (LIVD) Test Code Mapping for SARS-CoV-2 Tests.
65. NESTcc. NESTcc Test-Cases <https://nestcc.org/test-cases/>.
66. NESTcc. Evaluation of Uptake of Unique Device Identifiers (UDIs) by Health Systems.
67. FDA. National Evaluation System for health Technology (NEST) <https://www.fda.gov/about-fda/cdrh-reports/national-evaluation-system-health-technology-nest>. Accessed May 24, 2021.
68. ASPE. Improving the Quality, Interoperability and Portability of Laboratory Data within and Between Institutions <https://aspe.hhs.gov/shield-standardization-lab-data-enhance-patient-centered-outcomes-research-and-value-based-care>. Accessed May 24, 2021.
69. Waters, M. SHIELD Harmonizing Standards Application to Accelerate Innovation <https://mdic.org/wp-content/uploads/2020/02/SHIELD-Harmonizing-Standards-Application-to-Accelerate-Innovation.pdf>. (2002). Accessed January 13, 2022.
70. HealthIT.gov. Interoperability Proving Ground <https://www.healthit.gov/techlab/ipg/>. Accessed July 1, 2021.
71. Carlson, J. HealthPartners, Allina partnership aims to lower costs by making care more effective <https://www.startribune.com/healthpartners-allina-partnership-aims-to-lower-costs-by-making-care-more-effective/600045529/>.
72. Public-private initiative focusing on Using RWE to Convert IVD EUAs. MDIC <https://mdic.org/news/openhand/>. Accessed February 18, 2021.
73. International Medical Device Regulators Forum Committee. IMDRF Strategic Plan 2021-2025 <http://www.imdrf.org/docs/imdrf/final/procedural/imdrf-proc-n39-strategic-plan-200925.pdf>.