



Published in final edited form as:

Curr Oncol Rep. ; 22(5): 45. doi:10.1007/s11912-020-00904-z.

Developing Real-world Evidence-Ready Datasets: Time for Clinician Engagement

James M. Snyder¹, Jacob A. Pawloski¹, Laila M. Poisson²

¹Department of Neurosurgery, Hermelin Brain Tumor Center, Henry Ford Hospital, 2799 West Grand Boulevard, Detroit, MI 48202, USA

²Department of Public Health Sciences, Hermelin Brain Tumor Center, Henry Ford Hospital, 2799 West Grand Boulevard, Detroit, MI 48202, USA

Abstract

Purpose of Review—Real-world data (RWD) applications in healthcare that support learning health systems and pragmatic clinical trials are gaining momentum, largely due to legislation supporting real-world evidence (RWE) for drug approvals. Clinical notes are thought to be the cornerstone of RWD applications, particularly for conditions with limited effective treatments, extrapolation of treatments from other conditions, or heterogenous disease biology and clinical phenotypes.

Recent Findings—Here, we discuss current issues in applying RWD captured at the point-of-care and provide a framework for clinicians to engage in RWD collection. To achieve clinically meaningful results, RWD must be reliably captured using consistent terminology in the description of our patients.

Summary—RWD complements traditional clinical trials and research by informing the generalizability of results, generating new hypotheses, and creating a large data network for scientific discovery. Effective clinician engagement in the development of RWD applications is necessary for continued progress in the field.

Keywords

Real-world evidence; Bioinformatics real-world data; Pragmatic clinical trials; Common data elements; Learning health systems; RWE; RWD; CDE; Big data; Neuroinformatics; Neuro-oncology; Electronic health record; Precision medicine; Point-of-care

Introduction

Widespread adoption of electronic health records coupled with advances in computational analysis and recent legislative changes have sparked tremendous enthusiasm for “real-world data” (RWD) applications in healthcare. RWD has gained popularity in recent years due to

✉ James M. Snyder, jsnyder5@hfhs.org.

Conflict of Interest The authors declare that they have no conflict of interest.

Human and Animal Rights and Informed Consent This article does not contain any studies with human or animal subjects performed by any of the authors.

interest in “real-world evidence” (RWE) as a pathway for drug approval. Neuro-oncologic conditions are an anatomical group of heterogenous neoplasms that typically harbor rare subtypes and have limited treatment options; thus, RWD may provide insight into clinical phenotypes and therapeutic responses through aggregation of data. Clinical notes are thought to be the cornerstone of RWD applications; however, there is variability in clinical notation, including which elements are documented, language and structure used, core data recorded, and definition of key terms. Justifiably, there is apprehension towards establishing norms for point-of-care (POC) data standardization due to increased workload on an overextended population (healthcare providers) who are wary of the achievements of such a system [1]. There is also an unheralded opportunity to accelerate healthcare discovery through adoption of scalable RWD applications that can aggregate clinical information across institutions and ultimately have a meaningful impact on our understanding of rare diseases and clinical phenotypes. This potential paradigm shift is dependent on “ground-truth-data” obtained at the POC recorded directly by care providers. Advances in computational power that can now handle massive volumes of health data and the capability for connected electronic health records through use of a common language have created a platform to accelerate care discovery resulting in improved health outcomes. These advances have provided impetus for the next evolution in clinical documentation: an electronic health record (EHR) designed not only for the patient in front of you but also for a learning healthcare network that informs care of future patients.

What Is Real-world Evidence?

The objective of this manuscript is to introduce key events driving RWE applications in healthcare and to propose a clinical framework for POC, ground-truth-data generation in neuro-oncology. A singular definition of RWD and/or RWE has yet to be agreed on; therefore, we chose to use the Food and Drug Administration (FDA) definition because a recent catalyst for RWE implementation has been alternative forms of approval for medical therapies by the FDA. RWE as per the FDA is the “clinical evidence about the usage and potential benefits or risks of a medical product derived from analysis of RWD.” RWD is defined as “data relating to patient health status and/or the delivery of healthcare routinely collected from a variety of sources” [2]. RWD can come from a variety of sources including electronic health records (EHRs), medical claims, billing data, disease registries, digital phenotyping, patient-generated data, internet activity, and other sources (see Fig. 1). Digital phenotyping as defined by Torous et al. refers to the “moment-by-moment quantification of the individual-level human phenotype in situ using data from smartphones and other personal digital devices” [3]. Use of RWE is not new; however, contemporary applications such as market expansion for drug approvals and rare disease phenotyping have reappraised the impact of RWE.

RWE Complements Traditional Clinical Trials

Randomized prospective clinical trials are the backbone of medical progress and will continue to be the gold standard for efficacy and safety driving evidence-based medicine. Yet, some argue that clinical trial participants represent only a subset of the general population and may not always reflect “real-world” populations. The claim is made that

through the process of maximizing internal validity, clinical trials are more likely to encompass patients with better performance status and less comorbidities. Additionally, there are concerns with narrowed demographic diversity in clinical trial participants, exclusion of advanced age patients, and elevated socioeconomic status when compared with population-level disease incidence [4–7]. Clinical trials are typically offered at large academic medical centers frequently clustered in major urban centers, whereas the majority of patients in the USA are treated in community hospitals, raising awareness of geographic barriers to clinical trial access and patient selection [5, 8••]. Prospective clinical trials harbor tremendous financial investment with an average cost for phase 1, 2, and 3 studies in the USA of \$3.8, \$13.8, and \$19.9 million dollars, respectively, and an average of 7.2 years for glioblastoma (GBM) trials to progress from completion of phase 1 through phase 3 [9, 10]. Multiple estimates suggest that less than 5% of cancer patients in the USA are treated on a clinical trial [7]. RWE may serve as a complement to existing clinical trials by adding data dimensionality over time that extends beyond the therapeutic window of a specific trial agent, refining clinical trial designs, reducing costs through use of RWD control groups, and facilitating implementation of a healthcare learning environment informing intervention reproducibility in expanded populations.

A Developing Framework for Point-of-Care Health Data to Power RWE

The 2016 21st Century Cures Act was intended to accelerate drug development and included a requirement for the FDA to create a RWE program to support new indications for approved drugs [11]. This legislation is also interpreted to support the use of patient experience data, biomarkers, and other surrogate markers in the approval process, paving the way for alternatives to prospective randomized clinical trials for drug approval [11]. Stakeholders across healthcare, including academia, private sector biotechnology, patient-centered clinical research networks, health systems, and consortia, are investing in POC data and other systems to capture health data in a reliable fashion [12–15].

In 2009, the Health Information Technology for Economic and Clinical Health (HITECH) Act catalyzed rapid adoption of EHRs and development of interoperable health technology systems through financial incentives tied to meaningful use [16]. Widespread use of highly functional EHRs would create an environment where the long-held aspirations of a learning health system(s) of curated patient data could be created at scale. In 2009, the National Cancer Policy Forum of the Institute of Medicine (IOM) presented a “rapid learning system for cancer care” workshop which outlined a plan to transform cancer care through an ever-evolving learning healthcare data network anchored in “collecting data in a planned and strategic manner” and a patient-centric, multi-step iterative process for data analysis [17]. In total, there were 12 IOM meetings on learning healthcare systems that resulted in the generation of a comprehensive strategic vision emphasizing POC health data curation to harness new computationally driven analytics tools that inform care decisions at the POC through an ever-evolving dissemination of knowledge [18, 19].

Since the HITECH Act, adoption of EHRs in US hospitals and outpatient offices has become the norm including urban, rural, community, and academic centers. The Healthcare Information and Management Systems Society (HIMSS) developed an 8-stage sequential

model of EHR adoption that reflects functionality called the Electronic Medical Record Adoption Model (EMRAM) [20]. EMRAM stage 5 and higher requires some aspect of the physician note to use structured templates and discrete fields which would greatly advance RWD quality and scope; however, there was no further EHR meaningful-use incentives in the HITECH act once stage 4 was achieved [21]. In 2015, the Office of the National Coordinator for Health Information Technology reported that 83.8% of US health systems had adopted EHR use that includes access to clinical notes. While tremendous progress has been made through the current EHR system, the next epoch of progress may be achieved through federated clinical notes, which requires either tedious abstraction processes or POC-structured data with discrete data elements embedded in the clinical narrative. Updated EHR adoption data was not available at the time of publication; however, predictive modeling suggests that stage 5 should have reached its peak in 2019, while stage 7, which indicates complete EHR use with a central data warehouse that facilitates data analytics, is not anticipated to reach peak adoption until 2035 [21]. True POC clinical phenotyping by the provider will require a transformation in how clinical notes are recorded, which will be rewarded through scientific progress.

In Pursuit of a New Taxonomy of Disease Based on Molecular Biology

In 2011, the National Academy of Sciences presented a heroic multi-decade framework to establish “a new taxonomy of disease” intended to marry medicine and research through an “information commons” of medical data and a “knowledge network” integrating molecular, environmental, and phenotypic data. This effort attempts to create new phenotypic characterizations of health anchored in the intrinsic biology of disease which is meant to serve as a distinction from the established system-based and histologic classifications [22]. The Cancer Genome Atlas (TCGA) and other efforts redefined our understanding of the intrinsic biology of disease, and the 2016 WHO Classification of CNS Tumors adopted molecular features as core elements in neuro-oncologic diagnosis [23]. In many sectors of healthcare, there are groups racing to develop molecularly annotated comprehensive clinical phenotyping repositories; however, none has drawn the attention like the acquisition of Flatiron by Roche Pharmaceuticals for \$1.9 billion US dollars [24]. Flatiron Health was a privately held, US cancer analytics company invested in oncology RWD, reported to have a federated RWD repository of more than 1 million patients, many of whom had linked molecular testing results through a partnership with a large commercially available molecular laboratory. While the race to achieve the vision for learning health systems continues, there remains a data disconnect across competing and fragmented data repositories.

Barriers to RWD Amalgamation

To achieve high-quality RWD, specific conditions need to be met so that the data is consistent, reproducible, and accessible. Semantic heterogeneity can be overcome through use of common data elements (CDEs) that follow a defined data dictionary. For example, glioblastoma, a relatively rare malignant primary brain tumor, is referred to in the literature as “glioblastoma,” “glioblastoma multiforme,” “GBM,” and “grade IV astrocytic tumor” with more labels used in clinic notes. The use of CDEs is necessary to reduce data

fragmentation and improve aggregation, so that specific phenotypes and molecular subgroups can be identified, a core objective of the “information commons” model. There is a collaborative effort to establish CDEs in oncology as well as neurology; however, a defined set of neuro-oncology-specific CDEs has yet to be determined and will likely need to be aggregated across individual efforts to achieve scale [25, 26]. Issues related to technical heterogeneity also need to be overcome so that data can be compiled across software platforms and research tools. In many institutions including EMRAM stage-4 centers, pathology reports, clinic notes, radiation oncology records, and other key documents may be unstructured in the EHR. Some of these documents may have originated in a structured source application, but through the process of export into the EHR, discreet elements may have been lost. Additional processing or abstraction is then required to again reveal the clinically meaningful data within the EHR. This extra step should be eliminated once the EHR is configured to receive structured data. The health IT community is adopting frameworks to overcome the technical limitations of data aggregation, interoperability, and application design. Adoption of technical standards facilitates RWD and includes efforts such as SMART (Substitutable Medical Applications, Reusable Technologies) that enables development of healthcare IT tools that can function without specific knowledge of individual EHRs and HL7’s Fast Healthcare Interoperability Resources (FHIR) which attempts to facilitate health information exchange through implementation of core technical standards.

Clinician-Generated Point-of-Care Health Data

Resources are aligning to create the next generation of learning healthcare systems through RWD, but to achieve this vision, POC-structured clinical records that can be amalgamated at scale are needed. Even in centers with robust EHRs, clinical notes are typically free-text and exhibit significant semantic and structural heterogeneity. The lack of consistent terminology in the phenotypic description of our patients prevents reliable data capture. This absence of ground-truth-data at the POC undermines the process of a learning healthcare system. While there is a growing perspective that RWD can be compiled without clinician guidance through computational language processing, which now permeates our everyday lives (i.e., talk-to-text), clinician engagement in RWD is necessary to ensure accuracy of clinical data input, to govern data use, and to maximize clinical impact. Applications of computational language processing in healthcare pose unique challenges that limit adoption including patient privacy, data security, consequence of error, regulatory compliance, and clinician liability. Incentives to promote implementation of structured data using contemporary technical standards in clinical encounters are needed, as this is likely the rate-limiting step toward the development of a learning health system, particularly in oncology.

Neuro-oncology and other highly specific clinics that care for a narrow scope of clinical phenotypes are ideal sites to initiate RWD and POC-structured notation as most patients will have a limited and shared set of clinical presentations and may have similar treatment paradigms. Efforts to characterize and standardize the metrics used in prospective clinical trials across disease types are needed to guide RWD [10, 27, 28]. The areas of highest yield include patient demographics, comorbidities, pathologic diagnoses and molecular testing results, disease interventions, functional status, patient-reported measures, and disease

outcomes. Some variables, such as patient comorbidities, medications, and central laboratory tests, are typically recorded in structured fields within the EHR in the post-HITECH meaningful-use era.

A major limitation in the development of effective and scalable POC solutions is the uncertainty as to which data are required to define the disease course sufficiently to support a learning healthcare system. In clinical trials, where structured data are expected, patient data capture forms are largely designed around individual therapies and protocols without uniform structure across studies [10]. For instance, while most neuro-oncology trials capture overall survival (OS) and progression-free survival (PFS) as endpoints, there is no agreement as to which additional factors should be consistently captured and to what detail [10, 29, 30]. Consensus is needed about the types of data and level of detail that, if collected consistently, would provide meaningful information. A point of diminishing returns is expected, at which additional detail is captured unreliably thereby negating its utility. As was done for adoption of the EHR, a staged implementation of standardized data collection could be considered. An additional complication in the developing consensus around structured data elements is the changing landscape of clinical measures. For instance, outcome-defining factors are evolving with recent trials prioritizing health-related quality of life (HRQOL) and other measures of the patient experience as endpoints [29, 31]. Thus, systems of standardization must be flexible to support clinical evolution and designed efficiently with backwards compatibility in mind.

Efforts are underway to establish clinical data acquisition standards harmonization (CDASH) across commonly used elements in therapeutic trials [27]. CDASH organizes clinical data into the following categories: interventions, events, findings, and special purpose. For POC data in neuro-oncology, the authors propose to organize information from a clinical perspective into the following categories: diagnosis, interventions, surveillance, and outcome modifiers (Table 1). Use of a CDE dictionary is necessary to ensure interoperability. Data will need to evolve over time with ability to update with changing care paradigms. All fields should preserve longitudinal analysis through collection of metadata that informs CDE definitions, point in time, and unit of measure if appropriate. For example, a pathology specimen used for molecular analysis should state the date specimen was harvested and the date of analysis for molecular features, as well as protocol or testing used to provide the results.

Discussion

At the time of publication, there is neither an established standard for POC data in neuro-oncology nor an agreed-upon CDE dictionary; however, several tools are available to guide this process. The Radiologic Assessment in Neuro-Oncology (RANO) group has published several manuscripts to standardize interpretation of clinical data in neuro-oncology trials, with the bulk of this work focused on imaging that adheres to rules based on histopathology or treatment type (i.e., immunotherapy) [30]. Efforts are underway to automate neuro-oncologic radiographic interpretation based on RANO criteria through segmentation and computational analysis [32]. RANO imaging criteria use specific parameters that may differ from interpretations by a radiologist, and therefore, in many neuro-oncology trials, RANO

interpretation is interpreted by the study investigator and/or a central authority. At our center, we record RANO imaging interpretation in the clinical note and label with the specific RANO criteria applied. The RANO group also has established a standardized and graded neurologic clinical exam used in many clinical trials titled “neurologic assessment in neuro-oncology” (NANO) that could be adopted as a RWD standard [33]. Patient performance status is commonly used to guide treatment recommendations in glioma and is often a screening requirement for clinical trial participation and is monitored longitudinally throughout the disease [28, 29, 34]. Both the Karnofsky Performance Status (KPS) and the Eastern Cooperative Oncology Group (ECOG) score are in common use [33]. Patient-reported outcomes (PROs) are extremely valuable in characterizing the patient experience and HRQOL as recognized in the 21st Century Cures Act and their increasing weight in contemporary clinical trials. A consensus agreement as to which specific PROs should be adopted for the neuro-oncology community has yet to be established; however, the RANO group has outlined a plan to accomplish this task [35]. Involvement of a patient advisory group to provide patient important insight and guide outcome measures is advised. Common Terminology Criteria for Adverse Events (CTCAE) reporting is the standard in clinical trials and should also be followed in RWE generation [36].

There is a growing momentum towards establishing disease-specific CDEs that harmonize within the broader health landscape which in time will provide the universal language of RWD in neuro-oncology and lead to discussions of the core POC attributes for longitudinal neuro-oncology RWD. From there, an agreement will be necessary to create networks that contribute to RWD. Immediate adoption is limited by technical heterogeneity, significant resource investment to create structured disease-specific clinical data, and fragmentation of data networks. Clinicians may be able to overcome EHR barriers with computational language processing by implementing clinical notes designed to facilitate semantic analysis with adherence to CDEs using fixed sentence formats, labeling data in a manner that permits parsing the information into a recognizable CDE subject with associated modifiers [37•]. This would only need to be done for the prioritized variables, while the remainder of the clinical note retains the ability to capture any information with any format the clinician wishes. Additional challenges exist in getting clinicians to participate, ensuring the data is reliable and consistent, and generating meaningful insights.

The National Cancer Institute (NCI) and other bodies have committed to provide resources that facilitate RWD collection, RWE generation and analysis, and a broad adoption of data sharing with the aim to foster collaboration [38, 39••, 40]. NCI-funded research utilizes FAIR (findable, accessible, interoperable, reusable) principles with requirements to share source data and expand access to bioinformatics tools [38, 39••].

Pragmatic clinical trials investigate the outcome of specific interventions in unselected populations that closely reflect real-world situations, whereas, an explanatory trial aims to show the efficacy of an intervention under ideal circumstances [41, 42]. Patients with neuro-oncologic conditions often have additional factors that limit clinical trial participation (brain metastases, multiple disease recurrences, compromised functional status) resulting in uncertainty as to the benefit and risk of applying interventions studied in highly controlled environments [8••, 41]. Neuro-oncology encompasses many rare diseases where there are

little or no approved therapies, unsatisfactory outcomes, and few clinical trials for patients to enroll in or to guide treatment as seen in malignant meningioma, ependymoma, and recurrent low-grade glioma. Such conditions are poised to benefit from RWD studies, pragmatic clinical trials, and learning health systems to inform care.

Conclusions

Leveraging POC clinical data for RWE applications is gaining momentum as a means to complement traditional clinical trials. Neuro-oncologic conditions and other rare diseases that are heterogenous, intimately linked to rare molecular subtypes, and have few effective therapies may benefit from RWE applications and learning healthcare systems. RWE and POC data in healthcare are still in the developmental stages; however, the anticipated improvement in care quality and health discovery can only be achieved with ground-truth-data generation at the POC, which requires new perspectives on clinical notation.

Funding Information

This article is supported in part by National Cancer Institute grant no. R01CA222146.

References

Papers of particular interest, published recently, have been highlighted as:

- Of importance
 - Of major importance
1. Shanafelt TD, Dyrbye LN, Sinsky C, Hasan O, Satele D, Sloan J, et al. Relationship between clerical burden and characteristics of the electronic environment with physician burnout and professional satisfaction. *Mayo Clin Proc.* 2016;91:836–48. [PubMed: 27313121]
 2. Commissioner O of the Real-World Evidence [Internet]. FDA. 2019 [cited 2019 Nov 26]. Available from: <http://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence>.
 3. Torous J, Kiang MV, Lorme J, Onnela J-P. New tools for new research in psychiatry: a scalable and customizable platform to empower data driven smartphone research. *JMIR Ment Health* [Internet]. 2016 [cited 2018 Jun 23]:3. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4873624/>.
 4. Murthy VH, Krumholz HM, Gross CP. Participation in cancer clinical trials: race-, sex-, and age-based disparities. *JAMA.* 2004;291: 2720–6. [PubMed: 15187053]
 5. Sateren WB, Trimble EL, Abrams J, Brawley O, Breen N, Ford L, et al. How sociodemographics, presence of oncology specialists, and hospital cancer programs affect accrual to cancer treatment trials. *J Clin Oncol.* 2002;20:2109–17. [PubMed: 11956272]
 6. Unger JM, Gralow JR, Albain KS, Ramsey SD, Hershman DL. Patient income level and cancer clinical trial participation in a prospective survey study. *JAMA Oncol.* 2016;2:137–9. [PubMed: 26468994]
 7. Unger JM, Cook E, Tai E, Bleyer A. Role of clinical trial participation in cancer research: barriers, evidence, and strategies. *Am Soc Clin Oncol Educ Book Am Soc Clin Oncol Meet.* 2016;35:185–98.
 - 8•• Lee EQ, Chukwueke UN, Hervey-Jumper SL, de Groot JF, Leone JP, Armstrong TS, et al. Barriers to accrual and enrollment in brain tumor trials. *Neuro-Oncol.* 2019;21:1100–17. [PubMed: 31175826] Lee et al. provide a thorough discussion of barriers to clinical trial participation in neuro-oncology and offer possible solutions. With such poor outcomes for malignant glioma,

every effort should be made to improve access to clinical trials—and this article provides a path to address this gap.

9. Sertkaya A, Wong H-H, Jessup A, Beleche T. Key cost drivers of pharmaceutical clinical trials in the United States. *Clin Trials Lond Engl*. 2016;13:117–26.
10. Vanderbeek AM, Rahman R, Fell G, Ventz S, Chen T, Redd R, et al. The clinical trials landscape for glioblastoma: is it adequate to develop new treatments? *Neuro-Oncol*. 2018;20:1034–43. [PubMed: 29518210]
11. Kesselheim AS, Avorn J. New “21st Century Cures” legislation: speed and ease vs science. *JAMA*. 2017;317:581–2. [PubMed: 28056124]
12. Asher AL, McCormick PC, Selden NR, Ghogawala Z, McGirt MJ. The National Neurosurgery Quality and Outcomes Database and NeuroPoint Alliance: rationale, development, and implementation. *Neurosurg Focus*. 2013;34:E2.
13. Collins FS, Hudson KL, Briggs JP, Lauer MS. PCORnet: turning a dream into reality. *J Am Med Inform Assoc*. 2014;21:576–7. [PubMed: 24821744]
14. Axon Registry [Internet]. [cited 2019 Nov 27]. Available from: <https://www.aan.com/practice/axon-registry>.
15. Home | ASCO CancerLinQ [Internet]. [cited 2019 Nov 27]. Available from: <https://www.cancerlinq.org/>.
16. Blumenthal D. Launching HITECH. *N Engl J Med*. 2010;362:382–5. [PubMed: 20042745]
17. Abernethy AP, Etheredge LM, Ganz PA, Wallace P, German RR, Neti C, et al. Rapid-learning system for cancer care. *J Clin Oncol*. 2010;28:4268–74. [PubMed: 20585094]
18. Sledge GW, Hudis CA, Swain SM, Yu PM, Mann JT, Hauser RS, et al. ASCO’s approach to a learning health care system in oncology. *J Oncol Pract*. 2013;9:145–8. [PubMed: 23942494]
19. Smith MD, Institute of Medicine (U.S.), editors. Best care at lower cost: the path to continuously learning health care in America. Washington, D.C: National Academies Press; 2013.
20. Electronic Medical Record Adoption Model [Internet]. HIMSS Anal. - N. Am 2017 [cited 2019 Nov 9]. Available from: <https://www.himssanalytics.org/emram>.
21. Kharrazi H, Gonzalez CP, Lowe KB, Huerta TR, Ford EW. Forecasting the maturation of electronic health record functions among US hospitals: retrospective analysis and predictive model. *J Med Internet Res* [Internet]. 2018;20 [cited 2019 Nov 2]. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6104443/>.
22. National Research Council (US) Committee on a framework for developing a new taxonomy of disease. Toward precision medicine: building a knowledge network for biomedical research and a new taxonomy of disease [Internet]. Washington (DC): National Academies Press (US); 2011. [cited 2018 Jul 27]. Available from: <http://www.ncbi.nlm.nih.gov/books/NBK91503/>.
23. Louis DN, Perry A, Reifenberger G, von Deimling A, Figarella-Branger D, Cavenee WK, et al. The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. *Acta Neuropathol (Berl)*. 2016;131:803–20. [PubMed: 27157931]
24. Petrone J. Roche pays \$1.9 billion for Flatiron’s army of electronic health record curators. *Nat Biotechnol*. 2018;36:289–90. [PubMed: 29621225]
25. Grinnon ST, Miller K, Marler JR, Lu Y, Stout A, Odenkirchen J, et al. National Institute of Neurological Disorders and Stroke Common Data Element Project – approach and methods. *Clin Trials J Soc Clin Trials*. 2012;9:322–9.
26. mCODE: Creating a set of standard data elements for oncology EHRs [Internet]. ASCO. 2019 [cited 2019 Sep 30]. Available from: <https://www.asco.org/practice-guidelines/cancer-care-initiatives/mcode-creating-set-standard-data-elements-oncology-ehrs>.
27. Gaddale JR. Clinical data acquisition standards harmonization importance and benefits in clinical data management. *Perspect Clin Res*. 2015;6:179–83. [PubMed: 26623387]
28. Snyder J, Wells M, Poisson L, Kalkanis S, Noushmehr H, Robin A. INNV-15. Clinical data that matters: a distillation of neuro-oncology clinical trial inclusion criteria using machine learning. *Neuro-Oncol*. 2019;21:vi133.
29. Reardon DA, Galanis E, DeGroot JF, Cloughesy TF, Wefel JS, Lamborn KR, et al. Clinical trial end points for high-grade glioma: the evolving landscape. *Neuro-Oncol*. 2011;13:353–61. [PubMed: 21310734]

30. Wen PY, Chang SM, Van den Bent MJ, Vogelbaum MA, Macdonald DR, Lee EQ. Response assessment in neuro-oncology clinical trials. *J Clin Oncol Off J Am Soc Clin Oncol*. 2017;35:2439–49.
31. Frank L, Basch E, Selby JV. The PCORI perspective on patient-centered outcomes research. *JAMA*. 2014;312:1513–4. [PubMed: 25167382]
32. Chang K, Beers AL, Bai HX, Brown JM, Ly KI, Li X, et al. Automatic assessment of glioma burden: a deep learning algorithm for fully automated volumetric and bidimensional measurement. *Neuro-Oncol*. 2019;21:1412–22. [PubMed: 31190077]
33. Nayak L, DeAngelis LM, Brandes AA, Peereboom DM, Galanis E, Lin NU, et al. The Neurologic Assessment in Neuro-oncology (NANO) scale: a tool to assess neurologic function for integration into the Response Assessment in Neuro-Oncology (RANO) criteria. *Neuro-Oncol*. 2017;19:625–35. [PubMed: 28453751]
34. Nabors LB, Horbinski C, Robins I. NCCN Guidelines Index Table of Contents Discussion 2019;163.
35. Dirven L, Armstrong TS, Blakeley JO, Brown PD, Grant R, Jalali R, et al. Working plan for the use of patient-reported outcome measures in adults with brain tumours: a Response Assessment in NEURO-Oncology (RANO) initiative. *Lancet Oncol*. 2018;19: e173–80. [PubMed: 29508764]
36. Common Terminology Criteria for Adverse Events (CTCAE) | Protocol Development | CTEP [Internet]. [cited 2019 Nov 29]. Available from: https://ctep.cancer.gov/protocolDevelopment/electronic_applications/ctc.htm.
37. Ching T, Himmelstein DS, Beaulieu-Jones BK, Kalinin AA, Do BT, Way GP, et al. Opportunities and obstacles for deep learning in biology and medicine. *J R Soc Interface*. 2018;15:20170387. [PubMed: 29618526] Deep learning applications in healthcare are rapidly developing and promise to drastically alter medical practice. Ching et al. provide a stunning overview of deep learning applications in healthcare with relevance to clinically integrated RWD capture and analysis.
38. Alexander B, Barzilay R, Carpten JD, Haddock A, Hripacsak G, CO-CHAIR Sawyers Charles L.. Data science opportunities for the National Cancer Institute. Report of the National Cancer Advisory Board Working Group on Data Science. <https://deainfo.nci.nih.gov/advisory/ncab/workgroup/DataScienceWG/WGJune2019recommendations.pdf>. Accessed 27 Feb 2020.
39. Grossman RL, Heath AP, Ferretti V, Varmus HE, Lowy DR, Kibbe WA, et al. Toward a shared vision for cancer genomic data. *N Engl J Med* 2016;375:1109–1112. [PubMed: 27653561] This manuscript eloquently illustrates the mission of the Genomic Data Commons and outlines a plan for a truly collaborative and inclusive research environment to address multidimensional omics data and leverage available resources.
40. NCI Data Catalog | CBIIT [Internet]. [cited 2019 Nov 30]. Available from: <https://datascience.cancer.gov/resources/nci-data-catalog>.
41. Ford I, Norrie J. Pragmatic trials. *N Engl J Med*. 2016;375:454–63. [PubMed: 27518663]
42. Thorpe KE, Zwarenstein M, Oxman AD, Treweek S, Furberg CD, Altman DG, et al. A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers. *J Clin Epidemiol*. 2009;62:464–75. [PubMed: 19348971]

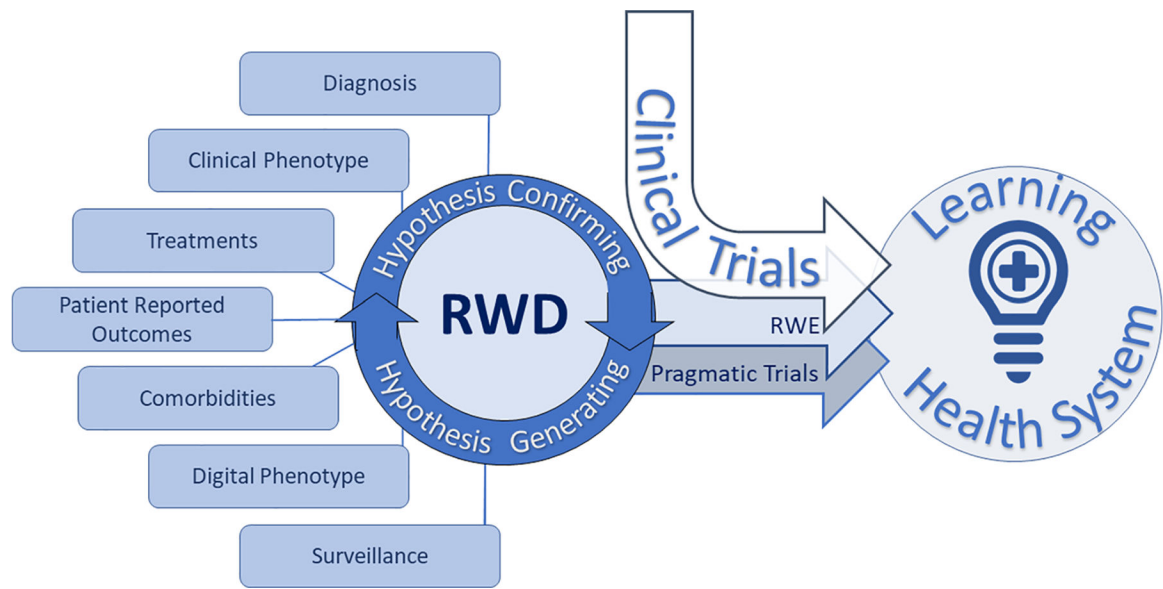


Fig. 1.

Real-world evidence (RWE) is derived from real-world data (RWD) systematically obtained from a multitude of sources loosely defined as patient information obtained outside of traditional research settings. RWE can expand the impact of prospective interventional clinical trials through evidence generation of the patient experience in conditions most similar to everyday healthcare situations (i.e., patients with additional comorbidities treated in routine clinical environments). Pragmatic clinical trials are a study type aimed at showing impact of an intervention in broad patient populations with minimal deviation in standard practice and are thus well-suited to RWE-based analysis. Classical clinical trials, RWE, and pragmatic trials inform learning healthcare systems

Table 1

Clinical data for prioritization in neuro-oncology point-of-care capture

Category	CDASH domain	Class	Factor	
Diagnosis	Findings	Pathology	Histology	
			Molecular features	
	Intervention	Exposure	Imaging	Molecular testing platform
				WHO grade
			Biomarker	Location
				Modality
			Staging	<i>Finding</i>
				Test
			Surgery	<i>Finding</i>
				<i>Cancer stage</i>
Surveillance	Findings	Radiation	<i>Extent</i>	
			Location	
		Disease-modifying therapy	Modality	
			Dose	
		Patient-reported outcomes	Patient-reported outcomes	Agent
				Start date
		Exam	Exam	Completion date
				Cycles completed
		Score	Score	<i>Adverse event</i>
				<i>Modification*</i>
Score	Score	<i>Outcome</i>		
		Modality		
Score	Score	<i>Outcome</i>		
		<i>Functional status score</i>		
Score	Score	<i>Clinical exam findings</i>		
		<i>Neurocognitive testing</i>		
Score	Score	<i>Test</i>		
		<i>Score</i>		

Category	CDASH domain	Class	Factor
Outcome modifiers	Events	Associated comorbidities/risks	Seizure
			DVT/PE
		Other cancers/tumors	
		Comorbidities	From EHR
		Medications	Steroids
			Others
	Special purpose	Social factors	<i>Support</i>

Variables in italics are least likely to have an existing structured data field in the EHR and are therefore most critical for clinician reporting. All reported factors require a minimum of an associated date and unit of measure

CDASH Clinical Data Acquisition Standards Harmonization, *EHR* Electronic health record, *DVT* Deep vein thrombosis, *PE* Pulmonary embolism, *WHO* World Health Organization

*"Modification" is a CDASH "special purpose" domain