

Clinical Researcher

January 2019

HOME STUDY TEST

Data and Technology Trends You're Already Part Of

Earn 2.0 Continuing Education Credits

Two articles from the January 2019 issue of *Clinical Researcher* have been selected as the basis for a Home Study test that contains 20 questions. For your convenience, the selected articles and test questions are combined and posted in the form of this printable PDF at <https://www.acrpnet.org/professional-development/training/home-study/>, where the test may be purchased. *The test will be active until January 31, 2020.* This activity is anticipated to take two hours. **Answers must be submitted using the electronic answer form online (members \$30; non-members \$50).** Those who answer 80% or more of the questions correctly will receive an electronic statement of credit by e-mail within 24 hours. Those who do not pass can retake the test for no additional fee.

The *Clinical Researcher* archive is at <https://www.acrpnet.org/resources/clinical-researcher/>.

CONTINUING EDUCATION INFORMATION

The Association of Clinical Research Professionals (ACRP) is an approved provider of clinical research continuing education credits.

Contact Hours

The Association of Clinical Research Professionals (ACRP) provides 2.0 contact hours for the completion of this educational activity. These contact hours can be used to meet the maintenance requirements for certification programs of the Academy of Clinical Research Professionals. (ACRP-2019-HMS-001)

ACRP DISCLOSURE STATEMENT

The Association of Clinical Research Professionals (ACRP) requires everyone who is in a position to control the planning of content of an education activity to disclose all relevant financial relationships with any commercial interest. Financial relationships in any amount, occurring within the past 12 months of the activity, including financial relationships of a spouse or life partner, that could create a conflict of interest are requested for disclosure. The intent of this policy is not to prevent individuals with relevant financial relationships from participating; it is intended that such relationships be identified openly so that the audience may form their own judgments about the presentation and the presence of commercial bias with full disclosure of the facts. It remains for the audience to determine whether an individual's outside interests may reflect a possible bias in either the exposition or the conclusions presented.

ACRP EDITORIAL ADVISORS

Suheila Abdul-Karrim, CCRA, CCRT, FACRP
Victor Chen, MSc
Fraser Gibson, CCRC, CCRA, CCRP
Stefanie La Manna, PhD, MPH, ARNP, FNP-C
Christina Nance, PhD, CPI
Paula Smailes, DNP, RN, MSN, CCRP, CCRC
Jerry Stein, PhD, ACRP-CP
Shirley Trainor-Thomas, MHSA
Heather Wright, CCRC:
Nothing to Disclose

ACRP STAFF/VOLUNTEERS

James Michael Causey (*Editor-in-Chief*)
Gary W. Cramer (*Managing Editor*)
Jan Kiszko, MD
Deepti Patki, MS, CCRC
Barbara van der Schalie
Nothing to Disclose

Clinical Researcher—January 2019 (Volume 33, Issue 1)

PEER REVIEWED

How Can Real-World Data Support Clinical Trials and Medical Research?

Manfred Stapff, MD, PhD; Jennifer Stacey

Drug development is a long and complex process. A drug's life cycle is not limited to the results of clinical trials in Phase I through Phase IV—it also includes research about the cause and natural history of diseases, clinical outcomes, long-term safety, tolerability, optimal treatment targets, and new indications.

Over the past decade, key performance metrics in the clinical trials area have been constantly disappointing, and even worsening.^{1,2} As the complexity of clinical trial protocols increases, the feasibility to conduct them and the percentage of sites fulfilling the enrollment goal decreases.^{3}

Clinical research, especially as applied to drug development, is a very data-driven industry. Biostatistical methods and the interpretation of study results based on *p*-values dominate the scientific decision-making processes. However, in planning and operating clinical research, objective data and metrics are not used to the extent to which they are available today.

For example, healthcare data have been available in electronic format in recent years. The digitalization of the U.S. healthcare system, driven by the 2014 “Meaningful Use” legislation, can be considered almost complete. By 2015, 96% of all hospitals in the United States had already adopted certified electronic health records.^{4} However, the electronic data thus far have been locked in various systems, are scattered in many locations, and follow different standards of ontology, units, and other characteristics. These electronic data issues, together with the lack of a consumable, user-friendly visualization platform, made the use and interpretation of this vast amount of medical information difficult.

The following is intended to provide an overview of how electronic health data can currently provision the design and conduct of clinical trials, as well as support other medical research areas.

Electronic Medical Records Versus Claims Data

The two major sources for real-world data (RWD) in medical research are electronic medical records (EMRs) and insurance claims data. EMRs can be considered as the more “medical” component of patients’ health information. They contain data about diagnoses, examinations, and treatments as documented by a provider who applies healthcare to a patient.

A patient’s EMR data from a specific healthcare provider may be blind to information collected by other providers. For instance, another hospital visited (or physician consulted) by a patient may not have access to the patient’s prior EMR information, as these records may be in separate repositories unique to each treating healthcare organization.

Meanwhile, claims data represent the more “administrative” part of a patient’s health history. They originate from the interaction between provider and payer, and could include documentation which have been submitted or adjudicated or remitted for payment. Because of the original billing related intention, claims data may be limited to information supporting reimbursement (see Table 1).

Table 1: Claims Data Compared to EMR Data

	Claims Data	EMR Data
Scope of data	Broad: Information from all doctors/providers caring for a patient	Limited: Only the portion of care provided by doctors using the specific EMR of a provider organization
Contained diagnoses	Limited to diagnoses supporting a claim	Complete set of conditions and comorbidities
Included patients	Payers' covered population, U.S.-employed socioeconomic group	All patients of a healthcare provider, including uninsured
Medication	All prescriptions that were filled, including dates of refills	Knows only that a physician prescribed a drug, but not if it was filled
Longitudinality of data	Payer/employer-based: As long as a patient stays with the same insurance	Provider-based: As long as the patient stays with the same healthcare provider
Richness of data	As necessary for reimbursement (diagnoses, procedures, treatments)	More complete medical picture (diagnoses, laboratory results, vital signs, problem list, etc.)
Timeliness	Lag time, delay from submit to close	Often real time, as soon as entered/coded

As a rule of thumb, one can assume that claims data better support studies about the economic effect of a therapy or the cost burden of a disease, while EMR data better capture natural disease history, efficacy, safety of a drug, or the outcome of a disease. Ideally, both data sources, used together on a patient level (also known as “linked data”) and cleaned for duplicate information, would provide an optimal dataset for all applications.

Incorporating additional data sources can enhance patients' health histories and clinical characteristics beyond standard medical coding practices. For example, data from tumor registries often contain tumor stage at diagnosis, histology, and other cancer-specific factors; a genomics database may include details on sample sites tested and variant types. Such sources can open the world of personalized medicine in a data context.

How Can RWD Support Clinical Trials?

The success rate of drug candidates making it all the way from Phase I to launch remains low (approximately at 10%).^{5} Meanwhile, the complexity of clinical trial protocols, notably expressed by number of patient eligibility criteria requirements, is increasing, and this leads to significant enrollment challenges.^{6} Less than 30% of protocols do not need to be amended; 70% need two to three changes over the course of the study, which is an inefficient trend causing damages in terms of costs and time.

Study results are only taken seriously if there is a p -value below 0.05 or an appropriate confidence interval. However, in conducting clinical research, data and analytics are not used to the extent they are available.

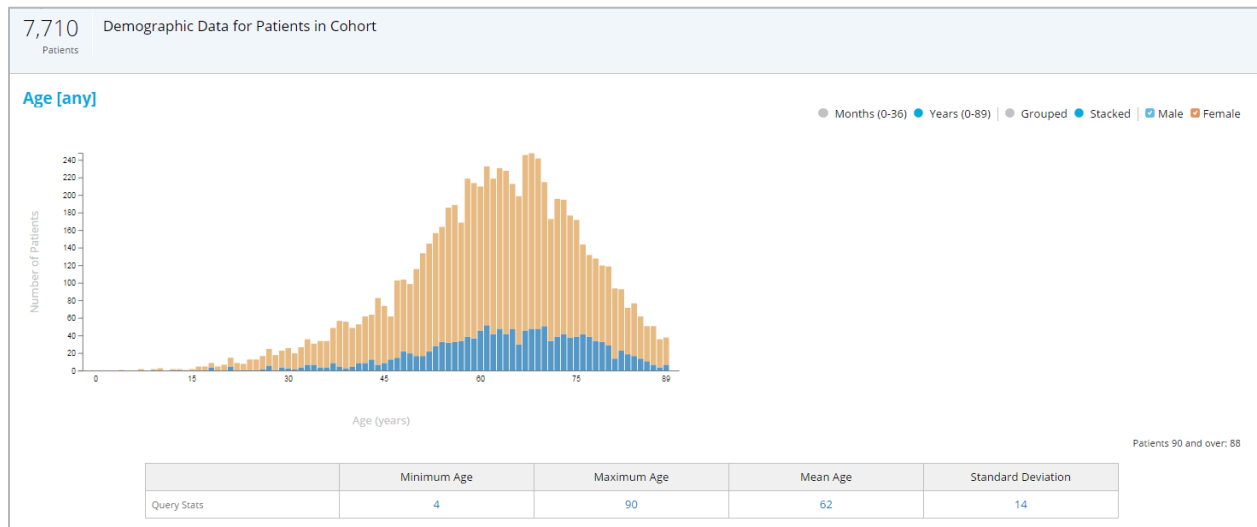
When determining the target population for clinical trial protocol design—from eligibility criteria to whether an amendment would improve a study—protocol authors traditionally rely on literature, experience, or expert opinion. They often do not have access to or use the extensive amount of RWD available in ways which would propel this decision-making process into a more objective and real-world scenario. Many costly and time-consuming amendments could be avoided by proper data-driven strategy planning.^{7}

Where in the Clinical Trial Process Can We Use RWD?

Trial design: Protocol authors need reliable and real-world information about patients, diseases, comorbidities, and concomitant treatments from routine medical practice (i.e., how patients present themselves in a true medical setting). RWD allow these authors to design clinical trial protocols in a realistic manner, including creating a feasible set of eligibility criteria.

A very simple example: It is a well-known fact that elderly patients and minorities are usually underrepresented in clinical trials.^{8} Upper age limits are introduced in trial protocols for safety reasons, but often shift the demographics of the population toward younger patients. A quick look at RWD and the age distribution of patients with the target indication helps to quantify the discrepancy, and to set the age limit to an optimal value with the best balance between safety and representativeness (see Figure 1).

Figure 1: RWD Age and Gender Distribution of Patients with Rheumatoid Arthritis (RA) on Disease-Modifying Therapy



Note: If the upper age limit was set to 75 years versus 90 years, the study would miss almost 20% of patients with RA.

Study Feasibility: Clinical trial eligibility criteria are often compiled arbitrarily and carried forward through development phases by company standards (“tradition”), or come from individual(s) expert input(s). Nevertheless, they are rarely tested against RWD, especially in terms of their effects on the final percentage of eligible patients with all criteria taken into consideration together.

A simulation of a criteria analysis, or “patient funnel,” can help identify the most impactful criteria, predict recruitment hurdles, and test the effect on enrollment if criteria are changed (see Figure 2).

Figure 2: Automated “Patient Funnel” to Simulate Enrollment and Effect of Eligibility Criteria on Patent Availability (Based on RWD)



Note: Patients with cardiovascular events, with controlled hyperlipidemia, receiving statins, and with different cholesterol lab values were compared.

Site Selection: Once the protocol is designed, the study should be placed only in those sites where there is proof of availability of eligible patients. Traditionally, lengthy and time-consuming feasibility questionnaires are used to determine the number of potentially eligible patients at a site, and often this is estimated by an investigator. An RWD system which keeps the link from anonymized patient data back to the site—ideally with a built-in communication feature to the sites—allows the user to select sites with pre-screened patients. It also simultaneously addresses the Good Clinical Practice requirement{9} of the investigator to prove access to suitable study subjects.

Patient Screening: Data privacy regulations require that the collection of patients’ health data happens in an anonymized, or at least pseudonymized, manner that makes re-identification of individuals impossible. A federated network structure, however, allows aggregated statistical counts to be obtained from the data source, keeping the original data at the source together with an identification key. Therefore, this enables the site via an “Honest Broker” to re-identify

eligible patients (those matching inclusion/exclusion criteria) and potentially contact them (after respective institutional review board approval) for study participation.

Many vendors and data providers offer service or systems for different aspects of clinical trial optimization. The more steps of the process, from protocol design to patient enrollment, that can be addressed by the same system, the easier it will be for the corporate sponsor or the contract research organization it is using to implement such support from the procurement, budgeting, efficiency, and training perspectives.

From Randomized Controlled Trials to Real-World Evidence Studies

In decades past, randomized controlled trials (RCTs) represented the one and only gold standard for gaining scientific knowledge in drug development and in medicine in general.^{10} Only with the advent of EMRs has a new method, generally referred to as real-world evidence (RWE) studies, been discussed as reasonable and more representative alternative to RCTs.

RCTs are usually conducted in a very experimental and unrealistic setting. Nowhere in actual medical practice are patients so carefully selected, so closely supervised, and so well cared for as they are in a clinical trial. In some therapeutic areas, such as oncology, the expectations of study subjects achieving efficacy are quite high, and clinical trials are intensively promoted as best treatment options. Thus, a significant placebo effect can occur, and ethical questions are often raised.^{11}

Inclusion of patients who may be more likely to show efficacy and exclusion of patients with certain risk factors dramatically reduce the representativity of the study cohorts and their applicability to the general population.

On the other hand, RWE studies have their flaws, too. The main subject of criticism concerns the quality and completeness of the data, especially in terms of the trustworthiness of data analyses when there has been no randomization of the subjects. In comparative RWE studies, the susceptibility for confounding factors (bias) requires correcting or balancing methods, such as stratification or propensity scoring, to achieve comparable cohorts. See Table 2 for a summary of the general differences between RCTs and RWE studies.

Table 2: Main Differences Between RCTs and RWE Studies

RCT (experimental, “laboratory situation”)	RWE (data from real medical practice)
<ul style="list-style-type: none"> • Extremely selected population • Artificial experimental treatment • High expectation on efficacy by study subjects • Prospective randomization • Ideally, no confounding factors • Impacted by the scientific thinking of the protocol author • Average number of data points collected: 400,000 to 1,000,000 	<ul style="list-style-type: none"> • Population “as in medical practice” • Usual care • No significant placebo effect • Retrospective characterization • Very susceptible to significant imbalance and bias (e.g. physicians’ treatment decision) • Impacted by processes and documentation • Average number of facts in a typical RWD cohort: 500,000 to 900,000,000

The data from an RWE study cohort can contain much more information than the data from an RCT. This is mainly due to the much larger sample size, and often longer observation period, that can easily be achieved in comparison to RCTs with limited durations in which patients are willing to participate.

In RCTs, only a fraction of the available evidence is used. Estimates report that only about 2% of patients with cancer can enroll into a clinical trial, but we use 100% of the information from this highly selected small population for decision making about that specific oncology condition. {12}

Data in RCTs are perfectly validated against the source and checked for errors, as they are heavily regulated and monitored. Contrarily, RWE studies take the data as they are, reflecting actual medical practices. As such, they are influenced by patient characteristics (demographics, comorbidities) and by provider characteristics (prescribing behavior and documentation completeness).

Meanwhile, it is sometimes the case that not enough discipline is applied in differentiating between RWD and RWE. In general, the “data” in RWD are related to the delivery or reimbursement of healthcare to a patient; they only become the “evidence” in RWE if adequate methods of collection, analysis, and interpretation of the data are applied. Only the combination

of high-quality data collection and proper scientific methodology creates RWE out of RWD and makes this RWD/RWE combination “fit for purpose.”

The more accessible RWD become and the more valid RWE analyses are considered as researchers’ capabilities to do so develop and improve, the more questions will be raised over the extent to which RWE studies will one day replace RCTs. In our view, while RCTs are certainly complicated and costly, they will most likely never be replaced completely by RWE studies.

For drug development, especially in early phases when the knowledge about safety and efficacy of an experimental therapy is very limited, researchers may always need the experimental and relatively safe environment of a clinical trial. Yet in the advanced stages of clinical development—in Phase IV or perhaps even late in Phase III—RWE studies can be a much more cost-efficient tool for collecting the necessary knowledge based on a conditional approval for new indications or for long-term safety observations.

Proper Analyses Methods Needed for RWE Studies

Due to the ease of use and cost efficiency, it may be tempting to run repeated analyses on RWD until a desired result is found, and then take this result as scientifically proven. Terms like “data dredging,” “fishing expeditions,” “*p*-hacking,” and “selective publishing” are used to describe this undesirable practice. Therefore, it is extremely important to follow proper scientific methods from concept to planning regarding data collection, analysis, interpretation, and publication.

Ideally, an RWE platform would require a predefined analysis plan to be uploaded and documented, and would have an audit trail which date stamps all analytical steps. This would show that the pre-specified data analysis plan was followed, and no result-driven analysis was conducted.

Proper documentation of analytical steps is important for the overall credibility of the study, and for the use of RWE within the context of meeting regulatory expectations for validity of data comparable to what is seen from RCTs. Guidelines are being developed and standards are currently being defined by several organizations, including the International Society for

Pharmacoeconomics and Outcomes Research (ISPOR) and the International Society for Pharmacoepidemiology (ISPE){13}, as displayed in Figure 3.

Figure 3: ISPOR and ISPE Recommendations for Good Procedural Practices for Hypothesis Evaluating Treatment Effectiveness Studies

1. *A priori*, determine and declare that a study is a Hypothesis Evaluation Treatment Effectiveness (HETE) study or an exploratory study based on conditions outlined below.
2. Post a HETE study protocol and analysis plan on a public study registration site prior to conducting the study analysis.
3. Publish HETE study results with attestation to conformance and/or deviation from the study protocol and original analysis plan. Possible publication sites include a medical journal or a publicly available website.
4. Enable opportunities to replicate HETE studies (i.e., for other researchers to be able to reproduce the same findings using the same dataset and analytic approach). The ISPE companion paper lists information that should be reported to make the operational and design decisions behind an RWE study transparent enough for other researchers to reproduce the conduct of the study.
5. Perform HETE studies on a different data source and population than the one used to generate the hypotheses to be tested, unless it is not feasible (e.g., another dataset is not available).
6. Authors of the original study should work to publicly address methodological criticisms of their study once it is published.
7. Include key stakeholders (patients, caregivers, clinicians, clinical administrators, payers, regulators, manufacturers) in designing, conducting, and disseminating HETE studies.

Studies using RWD have not yet achieved the levels of credibility and sophistication credited to RCTs, with their very detailed guidelines and regulations. While the U.S. Food and Drug Administration has already issued guidelines for the use of RWD in the regulatory process of medical devices, similar guidelines for use in drug development are pending. Therefore, RWE studies will, for the time being, mainly focus on hypothesis-generating projects, signal detection in pharmacovigilance, and obtaining supportive data for new indications (see Figure 4).

Figure 4: Potential Use Cases for RWD in Medical Research and Drug Development

Clinical Development and Operations

Use in design and statistical planning (sample size, variability, event rates)

Test feasibility of eligibility criteria

Select sites with eligible patients

Pharmacovigilance and Patient Safety

Observe safety signals (e.g., previously unknown adverse events)

Identify subpopulations with unique risk profiles

Describe safety in real-world conditions, validate or disprove safety signals

General Medical Research

Generate new hypotheses about diseases, causes, and potential new therapies

Understand real patient demographics, comorbidities, comedication

Describe the natural history of diseases

Medical Services and Health Outcomes

Understand treatment patterns and pathways

Evaluate clinical effectiveness and patient outcomes over time

Discover auxiliary benefits beyond original indication

RWD and associated RWE may constitute valid scientific evidence, depending on the characteristics of the data and the analytical methods used. It will, however, take time for RWD guidelines to be established and the quality of RWD to achieve a high enough standard so that RWE studies can be used as a hypothesis confirming method, especially in the drug regulatory process. Nevertheless, RWD can now answer very important questions for which RCTs would be too time-consuming and costly for individual sponsors to conduct.

From RWD to RWE: An Example

After more than two decades of clinical trials conducted in hypertension, it is still unclear which antihypertensive class is better as first-line therapy for stroke prevention.^{ 14,15 } As most beta blockers (BBs) and angiotensin conversion enzyme inhibitors (ACEIs) are off patent and their clinical outcome results are unpredictable, it is unlikely that a sponsor would ever conduct a lengthy and costly clinical trial to compare BBs to ACEIs in stroke prevention. However, as seen in the following, RWD can provide valuable clinical insight to this comparative research.

The population to be evaluated is defined as having the ICD-10 code for hypertensive diseases (I10 to I15) and never having any cardiovascular medication before (taking either a BB or an ACEI as a first-line therapy). The ACEI group must have an ACEI but no BBs, and the BB group must have a BB but no ACEIs. To focus on relatively recent data, only patients with start of therapy after January 2013 are included (see Figure 5).

Figure 5: Cohort Definition Comparing ACE Inhibitors to Beta Blockers for First-Line Treatment of Hypertension

Cohort 1	Cohort 2
ACEI as of 2013 or later	Betablocker as of 2013 or later
47,857 Patients	65,204 Patients
1 Terms Selected	1 Terms Selected
Event 1A The terms in this event occurred on or before Dec 31, 2012	Event 1A The terms in this event occurred on or before Dec 31, 2012
<input type="checkbox"/> I10-I15 Hypertensive diseases	<input type="checkbox"/> I10-I15 Hypertensive diseases
<input type="checkbox"/> N0000029116 Cardiovascular medications	<input type="checkbox"/> N0000029116 Cardiovascular medications
Event 2A The terms in this event occurred on or after Jan 01, 2013	Event 2A The terms in this event occurred on or after Jan 01, 2013
<input checked="" type="checkbox"/> N0000029130 Ace inhibitors	<input checked="" type="checkbox"/> N0000029118 Beta blockers/related
Event 2B The terms of any instance of Event 2B occurred on or after the most recent instance of Event 2A	Event 2B The terms of any instance of Event 2B occurred on or after the most recent instance of Event 2A
<input type="checkbox"/> N0000029118 Beta blockers/related	<input type="checkbox"/> N0000029130 Ace inhibitors

The start of therapy has been defined as the index event with a three-year observation period, starting 30 days after the index event (to exclude carryover effects from any diagnosis documented at the index date). The browser-based analytics platform then provides the results to a defined outcome (i.e., the risk of experiencing a cerebrovascular event [I60- I69] within the observation period) (see Figure 6).

Figure 6: RWD Results Comparing Risk of Experiencing Any Cerebrovascular Event Up to Three Years After Starting Antihypertensive Therapy with an ACEI or BB

Cohort		Cohort Statistics		
		Patients in Cohort	Patients with Outcome	Risk
1	ACEI as of 2013 or later	47,827	3,084	6.448%
2	Betablocker as of 2013 or later	65,152	6,805	10.445%

Risk Difference		Risk Ratio		Odds Ratio			
Risk Difference	95 % CI	z	p	Risk Ratio	95 % CI	Odds Ratio	95 % CI
-3.997%	(-4.318%, -3.675%)	-23.486	< 0.0001	0.617	(0.593, 0.643)	0.591	(0.565, 0.618)

Since retrospective RWD observations are not based on randomized patients, confounding factors with the potential for introducing bias must be considered. In this case, the BB group had a slightly higher percentage of patients with cardiac arrhythmias, heart failure, atrial fibrillation, and ischemic heart disease, but was otherwise comparable to the ACEI group.

There are several methods to balance for such confounding factors, with two of the most common being stratification and propensity scoring. Stratification was applied in this example, with two more analyses being performed. One subgroup contained only patients with at least one of these cardiac comorbidities; the second excluded all potential confounding cardiac comorbidities. Both strata delivered results comparable to those seen to the initial cohort, indicating that the cerebrovascular advantage of the ACEI may not be confounded by a slight imbalance in cardiac comorbidities.

More sophisticated cohort definitions and biostatistical methodologies may bring this simple example closer to a hypothesis-confirming study, rather than a hypothesis-generating one. This example also shows that the use of RWD in combination with user-friendly analytical tools can very quickly provide a picture about the therapeutic effectiveness of treatments in real medical practice, rather than in the artificial environment of an experimental and costly RCT.

Scientific Use of “Big Data” in Medical Research: General Considerations

Discussions about the use of healthcare data for research are often motivated by two fundamental questions:

1. Which is the optimal source of data?

- RCTs are a very exact methodology to evaluate a treatment in isolation, under ideal conditions, and in a highly selective population. The results depend on patient eligibility criteria and on the details of the study protocol, providing insights to scientists on the effects of a molecule in development.
- RWE studies represent the collective experience from thousands of physicians treating millions of patients. They provide a complete view at the actual use of a product, its effectiveness, and related adverse events in a real medical setting. The results are influenced by treatment standards, physicians’ prescribing behavior, pharmaceutical company marketing, and documentation quality.
- A third data source is gaining more and more attention: patient-reported outcomes (PROs), including those communicated in social media. Here, the “voice of the patient” comes into play, and the information can be considered as if coming from an expanded focus group whose members are providing unfiltered, spontaneous feedback toward an overall view of the treatment’s reputation.
- While very sophisticated and scientific mentalities may still consider RCTs as the only gold standard in clinical research, one must admit that only together with RWE and PRO can a holistic picture about a treatment—its efficacy, tolerability, and effectiveness—be achieved. These three methods should not be viewed as competitive, but rather as complimentary.

2. Who can use a patient’s medical data?

This question can be approached from an ethical standpoint or from a legal definition of “ownership.” The legal aspects, such as those related to privacy, data forwarding, or the analyzing process, must obviously be considered when using patient data for any research purposes. Compliance with every applicable law is a must, and the use of data created by one organization and used by another requires a contractual agreement.

The ethical and scientific views add nuance to this issue. A look at the life cycle of a study from data collection to publication shows that every step in the process adds value in advancing clinical research (see Figure 7):

- The patient allows the healthcare provider to document her or his health data
- The provider joins a data network, so healthcare information can be shared and analyzed by researchers
- These researchers analyze the data and publish the results

Only then can progress in medicine be achieved toward helping patients with the same condition.

Figure 7: Data Value Chain



All participants and steps in this circle make valuable contributions to data-driven medical progress. Patients who will suffer due to a medical condition in the future should be able to benefit from the information obtained from those who suffer the same medical condition today. To define an ethical “data ownership” is rather difficult, but the use of patient data should be available to researchers aiming to improve how clinical research is conducted.

References

1. Lamberti MJ, Chakravarthy R, Getz KA. 2016. Assessing practices & inefficiencies with site selection, study start-up, and site activation. *Appl Clin Trials* August 5.
www.appliedclinicaltrials.com/assessing-practices-inefficiencies-site-selection-study-start-and-site-activation
2. Lamberti MJ, Brothers C, Manak D, Getz K. 2012. Benchmarking the study initiation process. *Therap Innov Reg Sci* 47(1):101–9.
<https://journals.sagepub.com/doi/abs/10.1177/2168479012469947>
3. Tufts Center for the Study of Drug Development. 2018. Rising protocol complexity is hindering study performance, cost, and efficiency. *Impact Report* 20(4).
<https://csdd.tufts.edu/impact-reports/>
4. The Office of the National Coordinator for Health Information Technology. Health IT Data Summaries. <https://dashboard.healthit.gov/apps/health-information-technology-data-summaries.php?state=National&cat9=all+data&cat1=ehr+adoption#summary-data>
5. Thomas DW, Burns J, Audette J, et al. 2016. BIO Industry Analysis Reports. *Clinical Development Success Rates 2006–2015*.
<https://www.bio.org/sites/default/files/Clinical%20Development%20Success%20Rates%202006-2015%20-%20BIO,%20Biomedtracker,%20Amplion%202016.pdf>
6. Getz K. 2014. Improving protocol design feasibility to drive drug development economics and performance. *Int J Env Res Pub Health* 11(5):5069–80.
https://www.researchgate.net/publication/262303511_Improving_Protocol_Design_Feasibility_to_Drive_Drug_Development_Economics_and_Performance
7. Getz K, Stergiopoulos S, Short M, et al. 2016. The impact of protocol amendments on clinical trial performance and cost. *Therap Innov Reg Sci* 1-6.
https://www.mdsol.com/sites/default/files/Abstract_The%20Impact%20of%20Protocol%20Amendments%20on%20Clinical%20Trial%20Performance%20and%20Cost_2016.pdf
8. Shenoy P, Harugeri A. 2015. Elderly patients’ participation in clinical trials. *Perspect Clin Res* 6(4):184–9. <https://www.ncbi.nlm.nih.gov/pubmed/26623388>
9. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH). 2016. ICH Harmonised Guideline. *Integrated Addendum to ICH E6(R1): Guideline for Good Clinical Practice E6(R2)*. Investigator – Adequate Resources (4.2.1).

https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E6/E6_R2_Step_4_2016_1109.pdf

10. Bothwell LE, Greene JA, Podolsky SH, et al. 2016. Assessing the gold standard—lessons from the history of RCTs. *NEJM* 374:2175–81.

<https://www.nejm.org/doi/10.1056/NEJMms1604593>

11. Howie LJ, Peppercorn JM. 2014. The ethics of clinical trials for cancer therapy. *NC Med J* 75(4):270–3. www.ncmedicaljournal.com/content/75/4/270.abstract

12. Murthy VH, Krumholz HM, Gross CP. 2004. Participation in cancer clinical trials: race-, sex-, and age-based disparities. *JAMA* 291(22):2720–6.

<https://jamanetwork.com/journals/jama/fullarticle/198896>

13. Berger ML, Sox H, Wilke RJ, et al. 2017. Good practices for real-world data studies of treatment and/or comparative effectiveness: recommendations from the joint ISPOR-ISPE Special Task Force on real-world evidence in health care decision making. *Pharmacoepidemiol Drug Saf* 26(9):1033–9. <https://www.ncbi.nlm.nih.gov/pubmed/28913966>

14. Ravenni R, Jabre JF, Casiglia E, Mazza A. 2011. Primary stroke prevention and hypertension treatment: which is the first-line strategy? *Neurol Int* 3(2):e12.

<https://www.pagepress.org/journals/index.php/ni/article/view/ni.2011.e12/0>

15. Hong K-S. 2017. Blood pressure management for stroke prevention and in acute stroke. *J Stroke* 19(2):152–65. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5466289/>

Manfred Stapff, MD, PhD, is Senior Vice President and Chief Medical Officer at TriNetX, Inc.

Jennifer Stacey (Jennifer.Stacey@trinetx.com) is Vice President of Clinical Sciences at TriNetX, Inc. in Cambridge, Mass.

Clinical Researcher—January 2019 (Volume 33, Issue 1)

PEER REVIEWED

Risk-Based Strategies and Technology: Streamlining Site Activities

Emily Palmisano Holliday, MACPR; Mary Raber Johnson, PhD, RAC

In the past several years, clinical trials have continued to increase in number and complexity. Site infrastructure must acclimate to technology and create adaptive coping methods to meet these demands.^{1} Resource adaptation within changing protocol requirements is a critical skill, particularly as technology advances in the global clinical market.^{2} The streamlining of processes requires embracing new technology and new applications of current technology for best practices in training, reporting, and documentation.

The challenges and advances of the 21st century are a give-and-take system, accommodating both the increasing detail of research and providing technological solutions for excellence in safety and quality.^{3} By using a risk-based approach to site operations that is rooted in technology, site staff can be supported by best practice clinical trial conduct.^{4}

While current literature is abundant on risk-based monitoring application for sponsor/contract research organization (CRO)–level processes, the use of these strategies at a site level has not been fully explored. This discussion outlines essential considerations for small and/or inexperienced sites using a two-pronged approach: using technology to streamline site activities, while also reallocating resources to the areas of highest risk.

The Site Role

Implementation of a risk-based approach to site processes is key in smaller practice trials, where funding for complex software aids and specialized staff may be lacking. Within practices of one to three investigators, research support staff often maintain numerous roles to accommodate the protocol requirements. Crucial skills of a research coordinator include “...organizational skills, attention to detail, interpersonal skills, and computer and database experience.... They can make or break the clinical trial.”{5}

When communicating complex situations, site-provided details create the data story. An accurate story allows downstream data managers/statisticians to correctly categorize and interpret adverse events, relationships, and investigational product comparisons. The information must be reliable in detail and well-representative of the real-life situation.{6}

The site focus should be on crucial success factors for growing and maintaining a clinical research team: proper knowledge, practical skills, and appropriate attitude. These skills allow site staff to effectively work in teams, stay motivated for self-directing tasks, and accurately handle information.{7}

Resource Allocation in Trainings—Strategic Use of Recorded Sessions

Rather than conveying a complete spread of equal resources to each aspect of a study, a risk-based approach has been shown to improve data quality at a lower cost.{8} As outlined by the U.S. Food and Drug Administration (FDA) and the European Medicines Agency (EMA), a risk-based strategy identifies the areas of greatest risk for each specific protocol and allocates the highest priority and resources to those high-risk areas.{1,9}

It may be necessary to first discuss the purpose and rationale for risk-based strategies with site staff in order to prevent negative assumptions.^{10} While larger and experienced sites may afford risk-based software, smaller and/or inexperienced sites will need to tackle trainings from a simplified, risk-based perspective.

With the amount of protocol/procedural education required in clinical studies, sites should consider recording narration of slides for minor trainings and for staff self-review purposes. While major trainings may require scheduling live presenters, recording such sessions allows for simple refreshers as needed.

Creating Visibility with Documentation

It is the sponsor's responsibility to ensure the quality of study data reported from each site. Inaccuracies or inadequacies in case report forms (CRFs) have been one of the most common reasons for FDA Warning Letters to clinical investigators.^{6} Good documentation practices (GDPs) aim to prevent these types of violations by reducing/eliminating queries within the patient record. Once a site's risk areas are identified, transparent procedures can be built into daily activities.

In novice research sites, insufficient documentation could be due to a lack of training and/or experience.^{6} Examples of essential GDP training include attention paid to documentation of Good Clinical Practices, study protocols and related procedures, and standard-of-care practices onsite. Evidence of these trainings should be tracked in dedicated electronic logs or forms, to create transparency and enhance reporting.

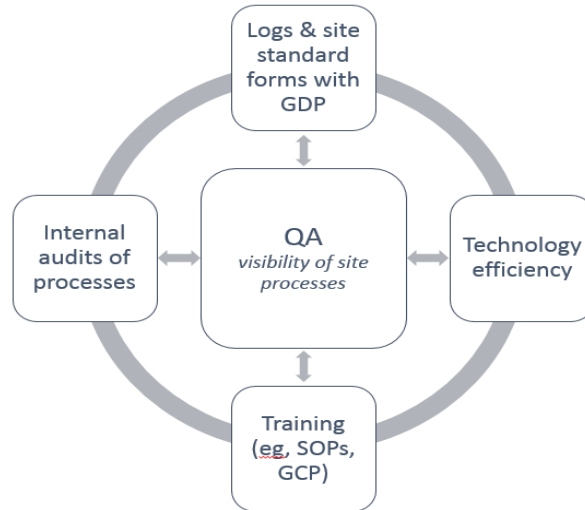
Note the difference between the task quality and procedure of documentation. In this case, “quality” refers to training completeness, while the documentation “procedure” refers to the correctness or detail of the associated log.

Figure 1: Example of a Log with GDP Elements

V1.0 15April2016		Version number and date.		
Dr. Smith's General Practice 1234 Street Road Boston, MA 555-555-1234				
Name of site and contact information.				
Site Training Log	Description of log.		Information on each person who participated as well as a confirmation signature.	
	The following staff members employed at Dr. Smith's General Practice participated in a training session.			
Name	Title	Signature	Date	Comments
Suzy Adams	Trial coordinator	<i>Suzy Adams</i>	4/20/17	none -
 				
SA 4/20/2017				
Date of training (dd mmm yyyy): <u>20 April 2017</u>				
Training completed by (name, title): <u>Dr. Smith, Medical Director</u>				
Training Topic: <u>Training covered good clinical practice guidelines and ICH guidelines current as of April 2017. The session lasted approximately 2 hours and included time for questions.</u>				
Signature of individual completing training:		<i>Dr. Jane Smith</i>		
Description of training, with date and name/signature of person who completed training.				

Figure 1 illustrates a well-constructed log using GDP elements. Two common research adages are worthy of memorization by site staff: “What is not documented is not done” and “Document what is done, as well as what is not done.” These phrases are excellent reminders to improve the visibility, providing a preventative and ongoing feedback loop for quality management. {6}

Figure 2: Interactions Between Site Process to Ensure Quality Assurance



When determining the interaction between site processes, Figure 2 may provide clarity.

Log/form standardization, training, and technology efficiency are dynamically involved with each other to ensure quality assurance (QA) at the site. With documentation of the research activities, specific movement of staff, patients, and data throughout the practice becomes measurable.

Consider an alternate application of technology to facilitate mandatory staff training, thereby allowing visibility within site process; for example, staff are required to review slides on a minor protocol amendment. Upon completion, staff are instructed to complete a brief electronic quiz on SurveyMonkey[®] to demonstrate understanding. After the training due date, the quiz results are downloaded and filed, along with the slide deck.

A QA reviewer of the quiz will report any concerns about individual results to management, and a discussion on future risk mitigation could commence—fulfilling certain safety requirements and guidelines established in 21 CFR Part 50 (Protection of Human Subjects) in the *Code of Federal Regulations* and in the International Council for Harmonization’s E6(R2) guideline on GCP.{11,12}

Specific strategies on documentation should be integrated into operational site processes, enabling routines to support FDA compliance. This coupling of corrective/preventative action is the best practice for efficient and high-quality site management.{1,12}

Integrating GDPs into Your Site

The specific mechanism of action to establish visibility is outlined through GDPs. Per GCPs, GDP standards should be integrated into mandatory site standard operating procedures (SOPs). The FDA ensures data quality with the acronym ALCOA, which stands for attributable, legible, contemporaneous, original, and accurate.{13} The EMA later enhanced ALCOA with four new elements (complete, consistent, enduring, and available) when considering electronic documentation, changing the acronym to ALCOA-CCEA.{9} These aspects are wholly grounded in the traceability of data origination, which are more readily examined from electronic sources.

A second document that is often required for clarity at a clinical research site is a memo- or note-to-file (MTF or NTF). These are documents included to explain errors or issues that have occurred. M/NTFs should not be used as a cover-all on deviations, but as a tool to explain exceptions to total protocol compliance. They allow for clarity on the accidental deviations that

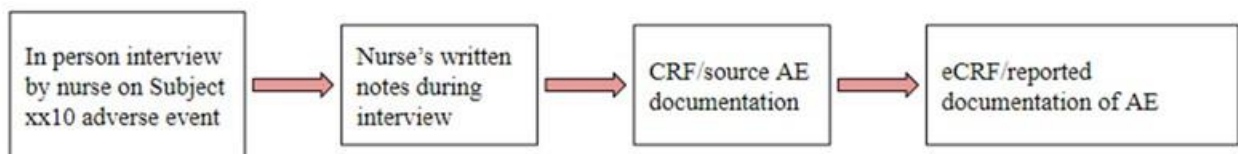
can be expected within real-world clinical trials. Additionally, original source pages or visit notes should be amended where possible to prevent unnecessary documentation. {6,14,15}

Technological Advantages within GDPs

Technology is a broad category of advancements that can be applied within various categories of clinical trials at the site level. Consider the use of paper source documents versus a low-cost, Internet-capable tablet. For the former, Figure 3 shows an example workflow for moving patient adverse event (AE)-related information reported at visits to electronic CRFs (eCRFs) in the electronic data capture (EDC) platform. The three sequential arrows represent the transcription of information from one form into another, with each having the potential for increased error rates.

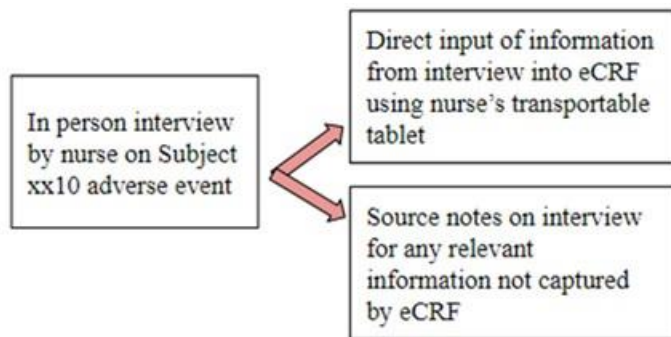
Unnecessary handling of data introduces risk and increases with the number of added variables, such as the number of staff members involved or dependent actions. In this example, a clinic nurse completes the patient interview, a research assistant transcribes notes from the clinic chart to participant binder, and a study coordinator enters information into the eCRF. Within part of the workflow, an error could occur and be repeated/exacerbated in dependent steps. { 16}

Figure 3: Example Workflow of Newly Reported AE, Without Application of Technology



Now consider Figure 4, which uses a tablet for the same scenario. The nurse is able to effectively enter the information directly while interviewing the patient. This has the minimum amount of risk, with a single transformation of data from the verbal patient report into the relevant eCRF as original data, without the need for a paper copy. Instead of a physical paper trail, the site process would be supported by the site SOP explaining this streamlined data collection process and rationale in risk mitigation.

Figure 4: Example Workflow of Newly Reported AE, With Use of Technology



With the tablet in use, workflow is immediately shortened by staff EDC data entry. Duplicate collection of information via secondary paper documents is removed from the equation, providing two benefits: 1) time-intensive creation/validation of paper source documents becomes unnecessary, and 2) information entry is completed only once between source and eCRF for each piece of data.

Tablet usage also allows the staff members completing patient interviews to be the ones entering data, eliminating the concern for data misrepresentation. Finally, there is also easy transportation

of the exact, protocol-required eCRF entry fields, ensuring that all necessary information is collected without extraneous or unessential details.

Considerations for Technology Use in Documentation

While embracing technology, it is important to note the requirements for security when utilizing software and digital site files. Title 21 CFR Part 11 outlines the specific requirements of electronic systems and how to translate the previous requirements for physical documents into the e-version of these pages. This FDA regulation also verifies that the use of electronic records is acceptable in place of physical records. Just as hard copies of patient information must be in restricted access areas, the e-versions of trial files must be located in protected workflows, with unique, confidential passwords for each employee.

The same requirements also apply in terms of audits, such as e-files being easily accessible to the FDA. For electronic items, there should be an audit trail to provide evidence of all changes, username of the change initiator, the date, and the changes made. Unique identifying passwords can act as the signature electronically and create an appropriate imprint to verify the authenticity of the user inciting any changes or approvals. Further, records retention is essential; much of the good practices for clinical trials pertains to the accuracy and assurance of the records.{17}

Sites should also consider how systems already used onsite could be expanded for documentation in clinical trials. For example, Microsoft® Office 365® aims to comply with 21 CFR 11 requirements, and has options for audit trails and cloud-based storage.{18} Adobe™ Acrobat™ can be used for more than document viewing, and includes e-signature capabilities.{19} Identifying existing programs that show FDA compliance with electronic

documentation can be a cost-saving technique to improve efficiency of technology according to GDPs.

Conclusions

Technology and risk-based considerations in site management pair productively with data management to enhance site output quality. Understanding the interaction between site activities in order to create visibility and measurability promotes risk management. Tactical application of the aspects described herein allow for optimization of site and study success in the new era of technology fluency.

References

1. U.S. Department of Health and Human Services. 2013. Oversight of Clinical Investigations — A Risk-Based Approach to Monitoring, Guidance for Industry.
2. Jones CT, Jester PM, Fitz-Gerald M. 2011. Issues in research management: protocol challenges in the era of complexity. *Res Practitioner* 12(4):121–31.
3. Friedman TL. 2016. *Thank you for being late: An optimist's guide to thriving in the age of accelerations*. New York: Farrar, Straus and Giroux; Chapter 3.
4. Arismendy C. 2010. How to start doing office-based clinical trials. *Medscape Business of Medicine*. <https://www.medscape.com/viewarticle/720916>
5. Gosnell JE. 2014. Building your clinical trial research team. In T. M. Pawlik & J. A. Sosa (Eds.), *Success in Academic Surgery: Clinical Trials*. London: Springer London.
6. Bargaje C. 2011. Good documentation practice in clinical research. *Perspect Clin Res* 2(2), 59–63.

7. Rajadhyaksha V. 2010. Training for clinical research professionals: focusing on effectiveness and utility. *Perspect Clin Res* 1(4), 117–9.
8. Agrafiotis DK, Lobanov VS, Farnum MA, Yang E, Ciervo J, Walega M, Baumgart A, Mackey AJ. 2018. Risk-based monitoring of clinical trials: an integrative approach. *Clin Ther* 40(7):1204–12.
9. European Medicines Agency. 2013. Reflection paper on risk based quality management in clinical trials.
www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2013/11/WC500155491.pdf
10. Jeong IS, Jeong JH, Hwang, YS, Youn JH. 2018. Clinical research coordinators' attitude toward risk-based monitoring. *Ther Innov Regul Sci*. Epub ahead of print. doi: 10.1177/2168479018793132
11. Chow S, Liu J, Ohio Library and Information Network, & Wiley Online Library. 2013. Design and analysis of clinical trials. Hoboken, N.J.: John Wiley & Sons.
12. International Council for Harmonization (ICH). 2016. Good Clinical Practice: Integrated Addendum to ICH E6(R1): Guidelines for Good Clinical Practice E6(R2)
https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E6/E6_R2_Step_4_2016_1109.pdf
13. U.S. Department of Health and Human Services. 2016. Data Integrity and Compliance with CGMP, Guidance for Industry.
14. PharmOut. 2016. How to implement good documentation practices [white paper].
<https://www.pharmout.net/downloads/white-paper-how-to-implement-good-documentation-practices.pdf>

15. Hazra A. 2011. Use, abuse and misuse of notes to file. *Perspect Clin Res* 2(1), 38–40.
16. U.S. Department of Health and Human Services. 2013. Electronic Source Data in Clinical Investigations, Guidance for Industry.
17. U.S. Department of Health and Human Services. 2003. Part 11, Electronic Records; Electronic and Application, Guidance for Industry.
<https://www.fda.gov/downloads/RegulatoryInformation/Guidances/ucm125125.pdf>
18. Compliance Certifications for Office 365. 2018. <https://products.office.com/en-us/business/office-365-trust-center-compliance-certifications>
19. Adobe Document Cloud. 2018. <https://acrobat.adobe.com/us/en/acrobat/how-to/electronic-signatures-online-e-signatures.html>

Emily Palmisano Holliday, MACPR, (emrose.palmisano@gmail.com) is a Clinical Trial Associate at Ora, Inc. in Andover, Mass.

Mary Raber Johnson, PhD, RAC, (johnson.6844@osu.edu) is an Assistant Clinical Professor at The Ohio State University.

JANUARY 2019 CLINICAL RESEARCHER

HOME STUDY

Data and Technology Trends You're Already Part Of

How Can Real-World Data Support Clinical Trials and Medical Research?

LEARNING OBJECTIVE

After reading this article, participants will be able to describe how electronic health data can provision the design and conduct of clinical trials, as well as support other medical research areas.

DISCLOSURE

Manfred Stapff, MD, PhD; Jennifer Stacey: *Employees of TriNetX, Inc.*

- 1. Which of the following is true of electronic medical record (EMR) data?**
 - a) They are documented by healthcare providers.
 - b) They are documented by a patient's employer.
 - c) They are limited to prescriptions that were filled for patients.
 - d) They are limited to details about diagnoses supporting healthcare payment claims.

- 2. Which of the following is true of claims data?**
 - a) They include a complete set of conditions and comorbidities affecting a trial.
 - b) They are most often recorded in real time by principal investigators.
 - c) They provide a more complete picture of a patient's financial stability.
 - d) They provide information from all doctors/providers caring for a patient.

- 3. Which of the following descriptions best characterizes randomized controlled trials (RCTs)?**
 - a) They are performed with extremely carefully selected populations.
 - b) Their susceptibility to significant data imbalances and researcher biases is very low.
 - c) They involve nothing more or less than the usual medical care for the condition.
 - d) They carry no possibility of significant placebo effects.

- 4. Which of the following descriptions best characterizes real-world evidence (RWE)?**
 - a) It derives from artificially arranged experimental treatments.
 - b) It carries the highest expectation of efficacy among patients and regulators.
 - c) It involves retrospective characterization of the results achieved from treatment.
 - d) It derives from patients who were prospectively randomized to treatment.

5. The International Society for Pharmacoeconomics and Outcomes Research (ISPOR) and the International Society for Pharmacoepidemiology (ISPE) are working on recommendations to improve:

- a) Randomization processes for blinding researchers and subjects in psychotherapeutic studies.
- b) Site selection criteria for use by sponsors of multinational vaccine studies.
- c) Inclusion/exclusion criteria for potential subjects in trials with waiver of informed consent.
- d) Procedural practices for Hypothesis Evaluation Treatment Effectiveness studies.

6. Identifying subpopulations with unique risk profiles is involved in the potential use of real-world data (RWD) in the area of:

- a) Clinical development and operations
- b) Pharmacovigilance and patient safety
- c) General medical research
- d) Medical services and health outcomes

7. Understanding treatment patterns and pathways is involved in the potential use of RWD in the area of:

- a) Clinical development and operations
- b) Pharmacovigilance and patient safety
- c) General medical research
- d) Medical services and health outcomes

8. Selecting sites with eligible patients is involved in the potential use of RWD in the area of:

- a) Clinical development and operations
- b) Pharmacovigilance and patient safety
- c) General medical research
- d) Medical services and health outcomes

9. Stratification and propensity scoring are used to:

- a) Select clinical trials sites for industry-sponsored studies.
- b) Prepare a clinical trials budget for review by regulatory authorities.
- c) Balance for confounding factors when using RWD.
- d) Randomize clinical trials to ensure inclusion of more women and ethnic minorities.

10. Discussions about the use of healthcare data for research are often motivated by which of the following questions?

- a) Which phase of clinical trial is being conducted?
- b) Who is paying for the clinical trial?
- c) How many participants will be involved?
- d) Who can use a patient's medical data?

Risk-Based Strategies and Technology: Streamlining Site Activities

LEARNING OBJECTIVE

After reading this article, participants will be able to describe technology requirements for risk-based strategies at the site level.

DISCLOSURE

Emily Palmisano Holliday, MACPR: *Employee of Ora, Inc.*

Mary Raber Johnson, PhD, RAC: *Consultant with Raber Communications, LLC and employee of The Ohio State University*

11. What has been shown to improve data quality at a lower cost?

- a) Increasing the number of study processes.
- b) Following a risk-based approach.
- c) Subcontracting training of site staff to external providers.
- d) Eliminating site infrastructure that is too technology-heavy.

12. Who is responsible for the quality of the study data reported from sites?

- a) The CRC assigned to recruit patients to the study.
- b) The EMA or other regulatory authorities.
- c) The sponsor of the study.
- d) The ethics committee or institutional review board that approved the study.

13. Which is one the common reasons that FDA has issued Warning Letters to investigators?

- a) Their study site staff spent too much time in GCP training.
- b) Their study site failed to meet its enrollment targets.
- c) They followed an incorrect version of the protocol provided by the sponsor.
- d) There were inaccuracies or inadequacies in the study's case report forms.

14. How does integrating strategies on documentation into operational site processes benefit the site?

- a) It increases the likelihood of the study's results being published.
- b) It enables routines to support FDA compliance.
- c) It aligns site processes to meet state-specific requirements for trial conduct.
- d) It serves as evidence of training that has taken place at the site level.

- 15. What should be integrated into mandatory site standard operating procedures (SOPs)?**
- a) GCPs
 - b) GLPs
 - c) GDPs
 - d) GMPs
- 16. FDA ensures data quality by the acronym ALCOA, and the EMA enhanced this to ALCOA-CCEA for electronic documentation. What new elements have been added?**
- a) Current, case, evaluations, and approvals.
 - b) Complete, consistent, enduring, and available.
 - c) Consistent, clinical, evaluations, and accuracy.
 - d) Contemporaneous, clinical, and estimation of adverse events.
- 17. What is a note-to-file?**
- a) An appendix to site work instructions.
 - b) An essential document filed in the sponsor TMF.
 - c) A document that is included in the case report form.
 - d) A document that explains errors or issues that have occurred.
- 18. How does unnecessary handling of data impact a trial at a site?**
- a) It introduces risk that increases with the number of added variables.
 - b) Transcription of the same information by more than one individual violates several regulations.
 - c) Time-intensive validation of the source versus the CRF is not considered billable work.
 - d) It interrupts the workflow in ways that could support a class action lawsuit.
- 19. What is the purpose of unique identifying passwords in electronic records?**
- a) They ensure a subject's information cannot be accessed by study sponsors or investors.
 - b) Their use is in compliance with EMA requirements for ALCOA-CCEA and ICH GCP.
 - c) They ensure e-files are easily accessible to the regulatory authorities in non-U.S. studies.
 - d) They act as an electronic signature and create an imprint to verify the authenticity of the user.
- 20. Which system mentioned in the article is 21 CFR 11 compliant?**
- a) Microsoft Office 365
 - b) Apple Software
 - c) Google Chrome
 - d) Microsoft Media