Randomized Controlled Trials Versus Real World Evidence: Neither Magic Nor Myth

Hans-Georg Eichler[1,2,*], Francesco Pignatti[1], Brigitte Schwarzer-Daum[2,3], Ana Hidalgo-Simon[1], Irmgard Eichler[1], Peter Arlett[1,4], Anthony Humphreys[1], Spiros Vamvakas[1], Nikolai Brun[5] and Guido Rasi[1,6]

Compared with drugs from the blockbuster era, recently authorized drugs and those expected in the future present a heterogenous mix of chemicals, biologicals, and cell and gene therapies, a sizable fraction being for rare diseases, and even individualized treatments or individualized combinations. The shift in the nature of products entails secular trends for the definitions of "drugs" and "target population" and for clinical use and evidence generation. We discuss that the lessons learned from evidence generation for 20th century medicines may have limited relevance for 21st century medicines. We explain why the future is not about randomized controlled trials (RCTs) vs. real-world evidence (RWE) but RCTs and RWE—not just for the assessment of safety but also of effectiveness. Finally, we highlight that, in the era of precision medicine, we may not be able to reliably describe some small treatment effects—either by way of RCTs or RWE.

What types of clinical study designs are appropriate to generate knowledge about the good or bad effects of drugs? Should we rely solely on randomized clinical trials (RCTs) because they are less prone to bias or should we push for more real-world evidence (RWE) because it might provide some of the information we really need to guide treatment decisions in routine care? The arguments have been aired extensively on both sides of a sometimes acrimonious debate. The controversy about the value of nonrandomized RWE was recently fanned by an insightful article by Collins *et al*., titled "The Magic of Randomization vs. the Myth of Real-World Evidence"[1]; the authors make a well-justified case for modernizing and improving the feasibility of RCTs but also caution against the use of RWE.

We here share thoughts why we believe that neither magic nor myth need to be invoked and we explain why we believe the future is not about RCTs vs. RWE but RCTs and RWE—not just for the assessment of safety but also of efficacy and relative effectiveness. We highlight that the way evidence was generated for 20th century medicines, like statins, and the lessons learned from these examples may have limited relevance for the medicines being developed now and in the near-term future. Finally, we draw attention to the disconcerting fact that, in the era of precision medicine, we may simply not be able to reliably describe some small treatment effects—either by way of RCTs or RWE. To address these increasingly frequent situations characterized by unavoidably scarce evidence, decision makers, including regulators, payers/providers, and, ultimately, the prescriber and patient, may need to evolve their decision frameworks.

THE CHANGING NATURE OF DRUGS

The great majority of drugs authorized during the "statin era" before the turn of the 21st century were chemicals (with very few biologicals), aiming for "blockbuster status."[2] By contrast, recently authorized drugs present a heterogenous mix of product types: of the 73 products containing a new active substance centrally authorized in the European Union between January 1, 2018, and December 31, 2019, 39 were chemicals, 29 were biologicals, and 5 were Advanced Therapies Medicinal Products (ATMPs; comprising cell, gene, and tissue engineered therapies). Twenty of 73 were designated orphan medicinal products, meaning that the prevalence of the target condition is < 5 of 10.000 in the European Union. An additional 8 were authorized for rare conditions but did not receive orphan status for a range of legal or regulatory reasons; hence, the total number of products for rare diseases is 28, that is close to 40% of all products containing a new active substance (for a complete list of products and characteristics, please see **Table S1**).

Based on the European Medicines Agency (EMA)'s in-house horizon scanning activities and predictions by other organisations,[3] we are confident that the trend toward complex biologicals, ATMPs, and drugs for orphan diseases will accelerate over the coming decade. It has been estimated that by 2025, 10–20 cell and gene therapy products will be approved each year.[4]

The shift in the nature of products is relevant because it entails other secular trends for evidence generation. Most older drugs belonged to classes of compounds that are pharmacologically similar and intended for large groups of eligible patients; examples include statins, angiotensin receptor blockers, or proton pump inhibitors. With these compounds, the pharmacologic target of drug action and the drug-target interactions are essentially the same across most or all patients. Observed differences in clinical effect size across patient subpopulations can

[1]European Medicines Agency (EMA), Amsterdam, The Netherlands; [2]Medical University of Vienna, Vienna, Austria; [3]EMA's Committee for Orphan Medical Products (COMP), Amsterdam, The Netherlands; [4]London School of Hygiene and Tropical Medicine, London, UK; [5]Danish Medicines Agency, Copenhagen, Denmark; [6]University Tor Vergata, Rome, Italy. *Correspondence: Hans-Georg Eichler (Hans-Georg.Eichler@ema.europa.eu)

Dedication: This article is dedicated to the memory of the late Sir Alasdair Muir Breckenridge, CBE, FRCP, FRCPE, FRSE, FMedSci (1937–2019), a dear friend and tireless campaigner for innovation in drug regulation.

Received August 21, 2020; accepted September 30, 2020. doi:10.1002/cpt.2083

usually be explained by differences in down-stream pathologies or other between-patient differences external to the drug-target interaction *per se*. For example, the pharmacological drug targets for statins are the same across a number of subpopulations that can be defined by extra-pharmacological patient characteristics, such as primary prevention (a lower risk group) or secondary prevention (higher risk) groups. Consequently, it has been claimed that "valid estimates of the absolute benefits and harms of a treatment can be obtained by applying reliable randomized evidence for its separate proportional effects on each outcome of interest to the absolute incidence of these outcomes in observational studies conducted within a particular population."[1] Where this assumption holds, results from RCTs may indeed be reasonably extrapolated across subpopulations (e.g., from higher risk to intermediate or even lower risk groups).[5,6]

Contrast this situation with the case of, for example, recently authorized treatments for cystic fibrosis (CF); patient heterogeneity is of a fundamentally different nature. CF is caused in all patients by a well-defined dysfunction in a single physiologic structure, a transmembrane chloride channel (CFTR). However, in each individual patient, the dysfunction itself is the result of one of > 2,000 different known mutations in the *CFTR* gene, which lead to defects in protein production, trafficking, function, misfolding, or premature degradation.[7] Drugs designed to modulate the CFTR product itself (rather than downstream common pathologies like infection or bronchial obstruction) are confronted with different drug targets across different subpopulations. The proteins encoded by some mutations are a better fit for a given molecule than those encoded by other mutations, and with some mutations there is no target protein present at all. It follows that the definition of a treatment population for the design of such drugs and for the generation of evidence about their effects (to be discussed below) must be mutation-specific, or at least specific to clusters of mutations. Extrapolation of efficacy information from one subpopulation to the next is therefore fundamentally different from a statin-type situation and will be quite impossible in some cases.

A very different example of drug-target heterogeneity is presented by cell therapies, which are autologous and thus patient-specific; no two patients receive identical treatments—a situation unheard of with nearly all 20th century treatments. Two autologous chimeric antigen receptor T cell therapies were approved in the past years in several regions for the treatment of defined hematologic malignancies.[8,9] On-market experience shows that their clinical effectiveness is dependent not just on the clinical condition of the patient at the time of treatment but also, to no small extent, on the way the cells are harvested, prepared, and administered.[10,11]

The most extreme form of individualization of the drug-target interaction is now emerging with the development of therapies specifically designed for one single patient. We recently witnessed the development of an antisense oligonucleotide (ASO) therapy (milasen) specifically designed to correct the mis-splicing of mRNA for a single patient with a very rare form of Batten's disease. The researchers emphasized that milasen "is not suited to the treatment of other patients with Batten's disease because its design is customized to our patient's specific mutation."[12]

To date, there is only a very small number of patients, including a person with idiopathic multicentric Castleman's disease, who identified a specific signaling pathway as a target in their own disease. Yet, these cases illustrate how recently developed technologies permit the delineation of pathways for truly individualized drug development.[13]

ASOs are not the only therapeutic platform that lend themselves to the design of individualized treatments. Similarly, protein-based therapeutics (such as Fc-fusion proteins, antibody fragments, and antibody–drug conjugates), small interfering RNAs, or genetic constructs (e.g., for cell-based gene therapy) can be tailored to one individual patient's signaling or other disease pathway as a target in their own clinical condition. The hoped-for result of treatment personalization is not only improvement in effect size, compared with the blockbuster model, but also an expansion of the numbers and types of selectable targets.[7]

Drugs designed for individual patients raise the fundamental question of our definition of the term "drug." Consider a defined "condition," like CLN7 Batten's disease, for which different pathogenic mutations are found in (nearly) each individual patient. Although all of them may receive an ASO with the same chemical backbone and sugar chemistry modifications, each ASO would be chemically different and patient-specific. Are we looking at one archetypal drug, or a drug class, or a multitude of different, individualized drugs? Similar considerations may apply to our definition of what is a "clinical indication." For regulatory approval, it has been suggested that "approvals as variations on a well-characterized archetypal product might be feasible if the interventions are closely related."[13] Although this might be a useful approach to regulatory approval of some families of individualized treatments, we will need to address the separate question of study designs for this growing number of 21st century treatments.

As our understanding of disease mechanisms improves, combination therapy targeting different aspects of disease pathogenesis, already a standard procedure in cancer, will likely become the new standard in a growing number of disease areas. Combination may refer to sequential or combined administration of conventional drugs or even to combined modalities, such as small-molecule conjugations with an antibody. Although combination therapies are not new, and have been investigated in RCTs, personalized combinations, guided by omics or other criteria, may in some conditions result in a situation where (almost) no two patients will receive the same treatment regimens.

A further challenge to evidence generation for 21st century drugs may arise because of the changing nature of the product over time, and therefore of the drug-target interaction. We have discussed above how chimeric antigen receptor T effectiveness may depend on the production process. As experience accumulates, and cell processing is increasingly optimized over time, the biological effects of these cell products will shift and efficacy information from past clinical trials becomes less relevant. Another example is Zynteglo, a product based on autologous CD34+ cells encoding the βA-T87Q-globin gene, recently authorized in the European Union for the treatment of β-thalassemia. Patients enrolled in early pre-authorization trials had received a slightly different version of the product than patients treated immediately pre-launch

and post-launch, requiring bridging studies to reassure regulators of the safety and efficacy of the final approved product.[14] It may be expected that further post-authorization modifications to the manufacturing process will emerge—with this product and most other ATMPs—raising the question whether we are still looking at the same "drug" over time—and about the continuing relevance of early clinical trial data and the need for revised regulatory and other decisions.

A broadly similar issue will arise with the advent of drug-device combinations, with some devices incorporating software. Devices have been known for incremental improvements over time, making such treatments a moving target. How long will the results of a past RCT remain relevant if the treatment itself keeps changing over time?

Contrast this with, for example, simvastatin, whose chemical structure remains identical over decades and where there is no question about the continuing relevance of RCTs conducted decades ago, as long as the definition of target population remains unchanged.

The above examples and secular trends illustrate how 20th century blockbuster drugs are being replaced by niche products,[2] treatment platforms that lend themselves to personalization, individualized (or very-small number) drugs, or individualized drug-drug or drug-device combinations. With these emerging treatments, patient heterogeneity is not only a function of downstream or concomitant pathologies but of the very basis of drug action, the drug-target interaction. As a result, the definitions of "drug" and of "target population" are becoming more fluid and changing over time, with important consequences for drug utilization and evidence generation.

WHAT WILL THESE TRENDS IN THE NATURE OF MEDICAL TREATMENTS MEAN FOR HEALTHCARE SYSTEMS?

The obvious consequence of the shift from blockbuster to niche-buster drugs is a decrease in the size of the treatment-eligible population.[2] This is accompanied by a second, less apparent trend: statin-type drugs were prescribed by large numbers of (primary care) physicians. Novel drugs remain mostly in the hands of specialists, and, increasingly, are confined to use in tertiary care centers. Of the 73 drugs authorized in the European Union in 2018 and 2019, 59 (almost 81%) had some form of prescribing restrictions in their summary of product characteristics (see **Table S1**; e.g., "Treatment should be under the supervision of a physician experienced in the treatment of haemophilia"). Partly, this is due to the nature of the target population, such as patients with rare diseases, who are predominantly seen in specialized centers. However, many novel drugs, like cell or gene therapies, cancer treatments, and other non-small molecule drugs, also require intense immediate and follow up management. An example is the requirement for myeloablative conditioning before infusion of gene-modified human stem cells, which is associated with significant untoward effects. As a consequence, such treatments will mostly be confined to a small number of highly specialized centers, at least initially (some follow-up care might be shifted toward the primary care sector who therefore need to be adequately informed about treatment at the tertiary care facility). The shift in healthcare delivery has immediate consequences for our considerations of RCTs and RWE: in the future, specialized tertiary care facilities should be expected and held accountable to implement a high level of patient documentation that enables generation of high-quality real-world data, and ultimately the development of a "learning health care system" with the ability to provide increasingly robust assessments of drug effects over time.[15] Examples of using specialized centers for knowledge generation about novel drugs include the patient registry managed by the European Society for Blood and Marrow Transplantation,[16] and Cystic Fibrosis patient registries in Europe[17] and in the United States.[18]

We touch upon another consequence of the changing nature of drugs for health care and public health. Collins *et al.*[1] observe that "...the direction of drug development has changed in ways that may adversely affect public health. For example, in the past decade, the revenue from the 10 top-selling drugs in the United States increased by a factor of 2.5, but the patient population that those medications target decreased by a factor of 7.5 [...]. This trend [is] leading to a shift toward seeking treatments with larger effects in less common conditions that could be detected in smaller trials." We fully agree with the observation of the trend—which we have described in more detail above—and with the notion that the observed shift may have unwelcome economic consequences for healthcare systems. However, setting aside the economic aspects, we politely disagree with the conclusion that the long-term trend "may adversely affect public health." It goes without saying that larger effects are preferable to incremental effect sizes, such as those of statins. Some but not all of the personalized treatments discussed above have shown impressive effect sizes that proved immediately life changing for patients. As to the diminishing target populations, we would argue that we are in the early stages of 21st century treatments. Broad application of sequencing technologies and other advances in basic sciences has uncovered the causes of a range of (rare) pathologies and has identified new mutations responsible for previously defined disorders with considerable knock-on effects for similar pathologies. As we learn to manipulate technology platforms[7] to develop personalized, or at least stratified, treatments at an increasing rate and speed, the circle of patients benefiting from such treatments will inevitably widen. The broadening of focus from statin-type drugs to include these novel therapies should be welcomed by patients, healthcare professionals, and the drug development community. The difficulty of conducting (large) RCTs of these therapies should not deter us from seeking to exploit their potential.

WHAT WILL THESE TRENDS IN THE NATURE OF MEDICAL TREATMENTS MEAN FOR CLINICAL KNOWLEDGE GENERATION?

Building on our observations and expectations of 21st century drug treatments, we foresee at least three broad scenarios applicable to the broad spectrum of drugs.

The first scenario is not too dissimilar from the old statin paradigm: a small to medium size effect is to be demonstrated for consistent drug-target interaction across a large target population. A topical example is afforded by the urgent search for coronavirus disease 2019 treatments and vaccines: numbers of patients

available for enrollment in RCTs were high, at least during the initial peak of the pandemic, the treatment target (i.e., the virus), was the same for all patients (at least for those treatments that directly targeted viral infection and replication, and discounting viral mutations) and effect sizes were expected to be small. Moreover, the disease did not lend itself to evidence generation by way of non-RCT methods because effect modifiers were poorly understood and construction of external control arms was not a viable option in light of the fast evolution in disease management, resulting in "drift" in clinical outcomes.[19] In this case, large well planned RCTs, ideally multi-arm platform trials were doable and seemed like the only useful way to generate robust evidence. Note that even in this scenario, RWE has, and always had, an established place to complement RCTs for disease epidemiology, detection, and evaluation of safety issues, studying special populations and observing on-market effectiveness for products with RCT demonstrated efficacy.

The second scenario, at the other extreme end of the spectrum, will be a growing number of drugs for ultra-rare diseases, very small subpopulations, or individualized treatments where adequately powered RCTs are impossible to conduct.

These drug-indication pairs require a rethink of the traditional evidence generation paradigm, characterized by clinical trials conducted in a small sample of selected patients with a hope that the trial patients will be more or less representative of the much larger target population for whom the conclusions are drawn; in this traditional scenario, the majority of patients treated after marketing approval contribute little or nothing to knowledge generation, the research setting is separate from the treatment setting.

By contrast, with some niche products, the study population and the target population will become near-identical, as almost every new patient treated will necessarily become a trial patient throughout the life span of the drug (i.e., before and after an initial marketing authorization). Hence, the research and treatment settings will increasingly overlap. The good news is, when the trial population is the target population, concerns over representativeness or external validity of results become a moot point.

However, the bad news is, when RCTs are not doable as a consequence of small populations or the inability to ethically or practically create a (randomized) concurrent control population, what type of evidence is left to assess these new drugs? Options are necessarily limited and include the below types of study:

- Evaluation of disease trends before and after treatment (sometimes referred to as interrupted time series). Note that this quasi-experimental analysis, where a patient serves as their own control, is not feasible in neonate or very young children, a scenario that is becoming more frequent with gene therapies. In a few rare disease conditions, it may be possible to conduct double-blind, (placebo-) controlled crossover study design where each individual patient also serves as their own control.[20]
- Running a (small) single-arm trial and compare the single-arm trial data with external controls.[21] With the growth of information available in disease registries and from e-health records, it may be possible to match individual patients receiving an experimental therapy to one or more comparable patients (e.g., with the same mutation). Such external controls may come from different regions and, in many cases, control patients may have died by the time their data are used. However, external controls will by their very nature not completely match the patients under investigation, so the information generated by external controls in this setting is naturally limited. In addition, even with rich, high-quality data sets, not all relevant confounders may be available or known and cannot be matched.
- Extrapolation based on experience with broadly similar treatment modalities, *in vitro* experiments, animal models of efficacy, and pharmacokinetic/pharmacodynamic modeling, as well as modeling of disease progression based on natural history data.[20]

The third scenario sits between the first two and can be conceptualized as a combination of the other two scenarios, in that RCTs are doable for some patient subgroups but not for others.

The third scenario is illustrated by the example of CF: the most prevalent CF causing mutation, F508del, is sufficiently prevalent to enable the conduct of adequately powered RCTs within a reasonable time frame. Two confirmatory RCTs, comparing lumacaftor, a CFTR corrector, in combination with ivacaftor, a CFTR potentiator, with placebo on top of standard of care, successfully enrolled a total of 1,108 patients in 12 months.[22] Yet, for some patients with ultra-rare CFTR mutations, RCT information will not become available in the foreseeable future. For some of those patients the decision to treat or not to treat with the combination (or a component drug) might be based solely on an *in vitro* test of change in CFTR-mediated chloride transport, combined with extrapolation to clinical study results—even though the magnitude of the change in chloride transport is not correlated with the magnitude of clinical response for individual mutations.[23] Although this level of "evidence" may be deemed unsatisfactory, it is the best available basis for treatment decisions. The treatment outcome in these patients will need to be closely monitored and, effectively, the study population and the target population are becoming identical until enough experience has been gathered by way of RWE to inform treatment decisions for future generations of patients.

Whereas rare diseases may appear as the obvious examples of conditions requiring different study designs for different subpopulations, we anticipate the majority of cases in this scenario will likely emerge in the field of oncology, as a result of biomarker-driven drug development. For many types of malignancies, there will be different actionable biomarkers, some sufficiently prevalent to allow for adequately powered RCTs, perhaps in the form of multi-arm comparisons. However, the prevalence of other potentially actionable stratification biomarkers will be extremely low, presenting drug developers and decision makers with a situation analogous to that described above for rare mutations in CF. For example, the activating E17K mutation in AKT1 (v-akt murine thymoma viral oncogene homologue 1) kinase has a central role in one of the most frequently activated proliferation and survival pathways in cancer; it is found in around 3% of cases of breast cancer but is even less common in other solid tumors making a traditional anatomy-based stratified development impossible.[24–26]

The above examples illustrate what we argue will be the inevitable future of evidence generation for many drug-indication pairs: a

combination of randomized and nonrandomized methods, drawing on a variety of data sources, with data being prospectively or retrospectively collected. In most cases, patients will need to be followed up for prolonged periods of time, especially postauthorization, to empirically confirm or refute *a priori* assumptions of (comparative) efficacy and safety.

Figure 1 conceptualizes the complex matrix of evidence generation defined by the following required:

- breadth of information, which refers to (sub-)populations characterized by different mutations, disease severity, comorbidity, or phenotypic patient factors,
- depth of information, which refers to different efficacy and safety end points, and
- context of information, which refers to a range of comparator treatments, and treatment combinations.

The figure also highlights that the information from RCTs can only answer a minuscule fraction of the near-infinite number of questions about subpopulations, interactions, treatment settings, effects, etc., that are relevant to patients and healthcare professionals at the point of care. Although this limitation, along with the issue of external validity of RCT results, was true even for blockbuster-type drugs, it will become the dominant issue with 21st century type drugs.

If we wish to move from "satisficing" strategies ("do we have at least minimal information to be satisfied that the drug works on average in at least one group?") to optimal evidence generation strategies ("how can we collect the best possible information to refine the understanding of biological effects in different contexts?"), then the use of RWE and other sources of information complementing RCT information has to become the norm.

As RWE is becoming part of the picture, there are more questions to be answered in addition to the analytics approach, including limitations of real-world data availability (especially from specialist settings),[27] quality, and completeness, comparability of outcomes collected in the real-world vs. more rigorously assessed or adjudicated data collected in RCTs, and matching of outcomes in RCTs, such as quality of life data, that may not be assessed in routine clinical practice.

Fortuitously, there are drivers in the healthcare systems that may help address these issues, at least to some degree. These include the demand from payers and health systems to raise cost-effectiveness of treatments, demand from patients to get access to innovative treatments faster, the growing need to understand real-world/long-term effectiveness and safety, and a mindset shift that realizes

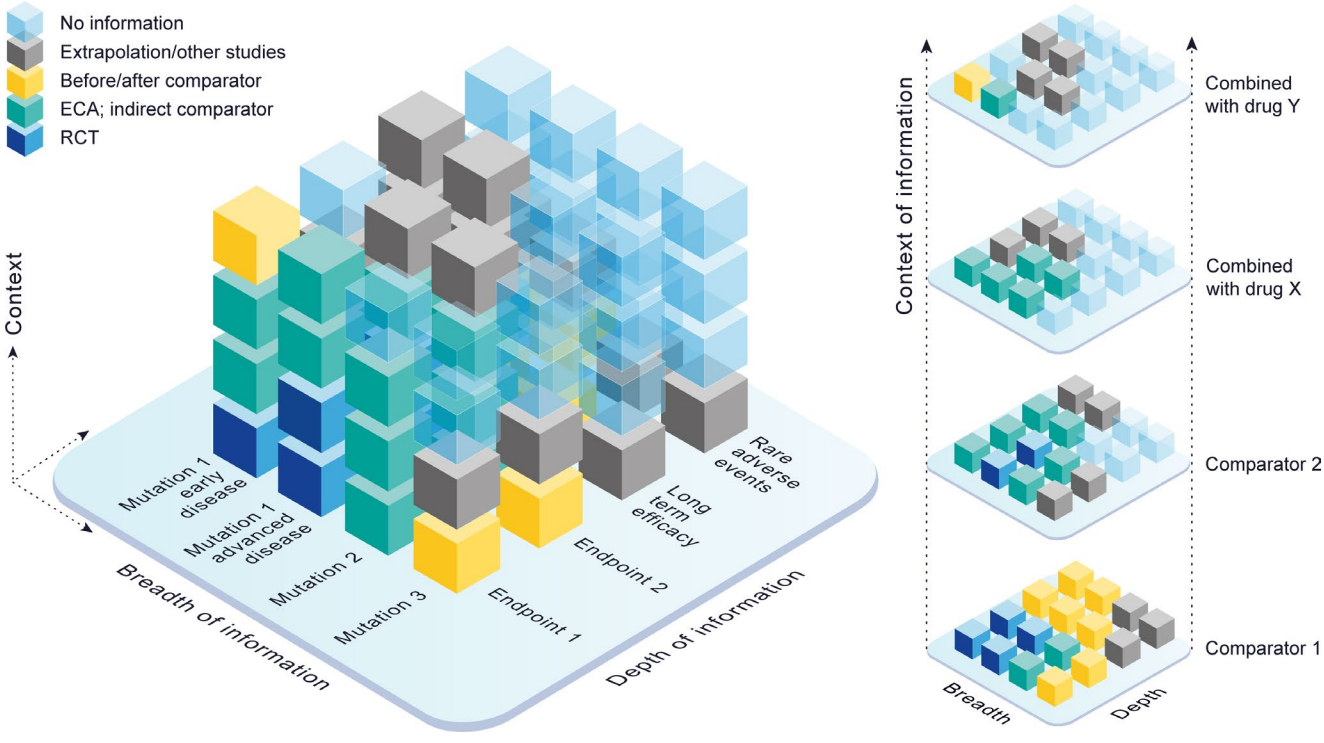

Figure 1 The complex matrix of research questions and methods. The graphic conceptualizes the complexities of research questions associated with a hypothetical drug treatment intended for a disease condition caused by different mutations in individual patients. Complexity is defined along three axes: the x-axis depicts breadth of information (i.e., different patient subgroups, based on mutation, phenotype, or disease stage), the z-axis depicts depth of information (i.e., different types of efficacy or safety end points of interest), and the y-axis depicts context of information (i.e., different comparators or treatment combinations). Each cell in the three-dimensional matrix represents an item of information that may be relevant for a particular decision maker and/or patient subgroup. Different study types (symbolized by different colors) will be required to generate the information, given the appropriateness of methods for different research questions as well as practical constraints on evidence generation. Note that for some research questions, there will be no data and information available at all, at least at the time of market launch. See main text for real-life examples that fit the schematic. ECA, external control arm; RCT, randomized controlled trial.

even RCTs have their own limitations. These drivers are gradually bringing about an increasing availability and emphasis on improved quality of real-world data, in turn, leading to growing use and acceptance of RWE along with RCTs.

Finally, the need for additional sources of (non-RCT) evidence and better methods for evidence synthesis is coming to the fore in the increasingly common situation of detecting small effects in small target populations. It is easy to detect large effects, like cure or complete and durable disappearance of symptoms, either by way of randomized or even nonrandomized methods, but small effects are tricky to demonstrate. Yet, some beneficial drug effects may be described as small in terms of evidence generation but may still be relevant in the eyes of patients. Sufficiently large randomized trials, which could reliably assess such effects are not feasible, due to a scarcity of patients. On the other hand, nonrandomized real-world studies may be doable (e.g., by relying on external control groups from various sources), as discussed above. However, the potential biases can be appreciable, so the results often cannot be trusted when the benefits of a treatment are only moderate.[1] In this situation, all options available to decision makers look unattractive. Should strict evidence standards be maintained, potentially denying patients a relevant benefit "because there is no robust evidence"? Or should evidence standards be relaxed, giving patients the benefit of the doubt, but at the risk of doing more harm than good?

There is, in our view, no single obvious solution to address the uncertainty resulting from unavoidably scarce evidence. However, to avoid potentially inconsistent one-off decisions we submit that regulatory (and other decision) frameworks will need to evolve to encompass four elements when decisions are taken under these circumstances:

- An agreement on the level of evidence needed at the point of initial authorization and market launch, including an agreement on what evidence can be derived from RCTs and/or (nonrandomized) real-world studies.
- Rigorous on-market follow-up of clinical outcomes in, ideally, the majority of patients, not only of safety but also effectiveness, with a view to progressively maximizing the understanding of benefits and harms. Where applicable, the postmarketing evidence generation plan should address the continuum of evolving RCT and nonrandomized RWE evidentiary value over time.
- Transparent, explicit, and retraceable synthesis of the available evidence including its limitations and trade-offs in the decision, and
- a regulatory (and Health Technology Assessment or payers') framework that allows frequent assessment and communication of new information and refinement of regulatory and other decisions as knowledge accumulates over time.

Besides a comprehensive decision framework, we need to address another question: how can we synthesize the information from different sources and analytic methods in a meaningful way? We have discussed how, with 21st century drugs, the need for more explicit methods of evidence synthesis is becoming apparent but there is no ready-to-use methodology that would lend itself to the many facets of the evidence matrix (**Figure 1**). It is tempting to speculate that advances in meta-analytical techniques, bioinformatics, causal inference, and artificial intelligence, building on artificial intelligence advances from other fields, can be harnessed to achieve the goal in a more explicit and systematic way in future. Similarly, regulators and other decision makers attempt to look at the "totality of evidence." Yet, this is most often done by implicitly weighing strength and importance of evidence from different study types. Methods like multicriteria decision analysis would be particularly important for communicating regulatory (or Health Technology Assessment or payers) decisions about complex quality, nonclinical, and clinical data from multiple sources and with many uncertainties; such methods are available and have been explored in the drug regulatory and other contexts but there is more work to be done to establish their value and role.[28]

CONCLUSION

The nature of new drugs coming to market and the conundrum of demonstrating moderate effects in very small populations dictates new ways of generating decision-quality evidence. Nostalgia for the "magic" of large RCTs and prolonging the debate about "RCTs vs. RWE" are unhelpful. Whereas fully acknowledging the need to improve the feasibility of RCTs,[1] we need to embrace the inevitable change and explore sound, explicit methods of synthesizing randomized and nonrandomized data, including the time after market introduction.

SUPPORTING INFORMATION

Supplementary information accompanies this paper on the *Clinical Pharmacology & Therapeutics* website (www.cpt-journal.com).

FUNDING

No funding was received for this work.

CONFLICT OF INTEREST

The authors declared no competing interests for this work.

DISCLAIMER

The views expressed in this article are the personal views of the author(s) and may not be understood or quoted as being made on behalf of or reflecting the position of the regulatory agency/agencies or organizations with which the author(s) is/are employed/affiliated.

As an Associate Editor of *Clinical Pharmacology and Therapeutics*, Spiros Vamvakas was not involved in the review or decision process for this paper.

1. Collins, R., Bowman, L., Landray, M. & Peta, R. The magic of randomization versus the myth of real-world evidence. *N. Engl. J. Med.* **382**, 674–678 (2020).
2. Trusheim, M.R., Berndt, E.R. & Douglas, F.L. Stratified medicine: strategic and economic implications of combining drugs and clinical biomarkers. *Nat. Rev. Drug Discov.* **6**, 287–293 (2007).

3. Alliance for Regenerative Medicines website: ARM Annual Report & Sector Year in Review <https://www.alliancerm.org/sector-report/2019-annual-report/> (2019).
4. Statement from FDA Commissioner Scott Gottlieb, M.D. and Peter Marks, M.D., Ph.D., Director of the Center for Biologics Evaluation and Research on new policies to advance development of safe and effective cell and gene therapies <http://bit.ly/2tH1DLX> (2019).
5. Chasman, D.I., Posada, D., Subrahmanyan, L., Cook, N.R., Stanton, V.P. Jr & Ridker, P.M. Pharmacogenetic study of statin therapy and cholesterol reduction. *JAMA* **291**, 2821–2827 (2004).
6. Kitzmiller, J.P., Mikulik, E.B., Dauki, A.M., Murkherjee, C. & Luzum, J.A. Pharmacogenomics of statins: understanding susceptibility to adverse effects. *Pharmgenomics Pers. Med.* **9**, 97–106 (2016).
7. Tambuyzer, E. *et al*. Therapies for rare diseases: therapeutic modalities, progress and challenges ahead. *Nat. Rev. Drug Discov.* **19**, 93–111 (2020).
8. EMA website: Kymriah EPAR – Public Assessment Report <https://www.ema.europa.eu/en/documents/assessment-report/kymriah-epar-public-assessment-report_en.pdf>.
9. EMA website: Yescarta EPAR – Public Assessment Report <https://www.ema.europa.eu/en/documents/assessment-report/yescarta-epar-public-assessment-report_en.pdf>.
10. Neelapu, S.S. CAR-T efficacy: is conditioning the key? *Blood* **133**, 1799–1800 (2019).
11. Majzner, R.G. & Mackall, C.L. Clinical lessons learned from the first leg of the CAR T cell journey. *Nat. Med.* **25**, 1341–1355 (2019).
12. Kim, J. *et al*. Patient-customized oligonucleotide therapy for a rare genetic disease. *N. Engl. J. Med.* **381**, 1644–1652 (2019).
13. Woodcock, J. & Marks, P. Drug regulation in the era of individualized therapies. *N. Engl. J. Med.* **381**, 1678–1680 (2019).
14. EMA webstite: Zynteglo EPAR – Public Assessment Report <https://www.ema.europa.eu/en/documents/assessment-report/zynteglo-epar-public-assessment-report_en.pdf>.
15. Eichler, H.G. *et al*. Data rich, information poor: can we use electronic health records to create a learning healthcare system for pharmaceuticals? *Clin. Pharmacol. Ther.* **105**, 912–922 (2019).
16. EMA website: Qualification opinion on Cellular therapy module of the European Society for Blood & Marrow Transplantation (EBMT) Registry <https://www.ema.europa.eu/en/documents/scientific-guideline/qualification-opinion-cellular-therapy-module-european-society-blood-marrow-transplantation-ebmt_en.pdf>.
17. European Cystic Fibrosis Society Patient Registry website <https://www.ecfs.eu/ecfspr>.
18. Cystic Fibrosis Foundation website: Annual report 2018 <https://www.cff.org/Research/Researcher-Resources/Patient-Registry/2018-Patient-Registry-Annual-Data-Report.pdf>.
19. Eichler, H.G. *et al*. "Threshold-crossing": a useful way to establish the counterfactual in clinical trials? *Clin. Pharmacol. Ther.* **100**, 699–712 (2016).
20. Mulberg, A.E. *et al*. Regulatory strategies for rare diseases under current global regulatory statutes: a discussion with stakeholders. *Orphanet. J. Rare Dis.* **14**, 36 (2019).
21. Carrigan, G. *et al*. Using electronic health records to derive control arms for early phase single-arm lung cancer trials: proof-of-concept in randomized controlled trials. *Clin. Pharmacol. Ther.* **107**, 369–377 (2020).
22. Wainwright, C.E. *et al*. Lumacaftor-ivacaftor in patients with cystic fibrosis homozygous for Phe508del CFTR. *N. Engl. J. Med.* **373**, 220–231 (2015).
23. FDA website. Highlights of prescribing information <https://www.accessdata.fda.gov/drugsatfda_docs/label/2018/210491lbl.pdf>.
24. Carpten, J.D. *et al*. A transforming mutation in the pleckstrin homology domain of AKT1 in cancer. *Nature* **448**, 439–444 (2007).
25. Bleeker, F.E. *et al*. AKT1(E17K) in human solid tumours. *Oncogene* **27**, 5648–5650 (2008).
26. Hyman, D.M., Taylor, B.S. & Baselga, J. Implementing genome-driven oncology. *Cell* **168**, 584–599 (2017).
27. Flynn, R. *et al*. Ability of primary care health databases to assess medicinal products discussed by the European Union Pharmacovigilance Risk Assessment Committee. *Clin. Pharmacol. Ther.* **107**, 957–965 (2020).
28. Tervonen, T., Angelis, A., Hockley, K., Pignatti, F. & Phillips, L.D. Quantifying preferences in drug benefit-risk decisions. *Clin. Pharmacol. Ther.* **106**, 955–959 (2019).