

# A Proposal for Integrated Efficacy-to-Effectiveness (E2E) Clinical Trials

HP Selker<sup>1,2</sup>, KA Oye<sup>3</sup>, H-G Eichler<sup>4</sup>, NL Stockbridge<sup>5</sup>, CR Mehta<sup>6</sup>, KI Kaitin<sup>1,7</sup>, NE McElwee<sup>8</sup>, PK Honig<sup>9</sup>, JK Erban<sup>1,10</sup> and RB D'Agostino<sup>11,12</sup>

We propose an “efficacy-to-effectiveness” (E2E) clinical trial design, in which an effectiveness trial would commence seamlessly upon completion of the efficacy trial. *Efficacy trials* use inclusion/exclusion criteria to produce relatively homogeneous samples of participants with the target condition, conducted in settings that foster adherence to rigorous clinical protocols. *Effectiveness trials* use inclusion/exclusion criteria that generate heterogeneous samples that are more similar to the general patient spectrum, conducted in more varied settings, with protocols that approximate typical clinical care. In E2E trials, results from the efficacy trial component would be used to design the effectiveness trial component, to confirm and/or discern associations between clinical characteristics and treatment effects in typical care, and potentially to test new hypotheses. An E2E approach may improve the evidentiary basis for selecting treatments, expand understanding of the effectiveness of treatments in subgroups with particular clinical features, and foster incorporation of effectiveness information into regulatory processes.

This article proposes an “efficacy-to-effectiveness” (E2E) approach to clinical trials that includes a seamless transition from efficacy to effectiveness trials. Our proposal is structured to capture the benefits of an integrated E2E approach without delaying regulatory review and approval. This approach offers logistical advantages by minimizing start-up costs by using the existing efficacy trial infrastructure. More important, it offers the opportunity for improved understanding of how a treatment will work in practice in patients with a typical range of characteristics, comorbidities, and treatment adherence, such as those seen in more usual clinical settings.

In addition to these medical benefits, facilitation of early effectiveness results could have a role in the availability and use of treatments. Payers point out that efficacy trials are necessary but not sufficient to inform decisions about payment coverage and reimbursement. Lack of clear demonstration of comparative clinical and economic value in usual care settings has increasingly resulted in outright denial of coverage, cost pass-through to patients through copayments (in the United States), and formulary restrictions of use. These conditions provide strong incentives for pharmaceutical companies

with novel and effective treatments to provide comparative effectiveness data as close in time as possible to marketing approval.

An E2E approach does face hurdles. In generating data, the potential benefits to all stakeholders will have to be balanced along with the incremental cost of such a design and the potential perceived dilution of treatment effect due to use in a more heterogeneous population. Acceptance by sponsors will require addressing the technical challenge of broadening inclusion criteria to increase heterogeneity while retaining sufficient statistical power in the face of greater variability. Acceptance also will require demonstrating the net benefits of the E2E effectiveness components to sponsors, regulators, payers, patients, and providers. Indeed, we believe that the benefits of an E2E approach will in many instances outweigh the risks and value for all stakeholders.

## CURRENT TRIALS SEPARATE EFFICACY AND EFFECTIVENESS

Both efficacy and effectiveness trials provide useful information for health-care decision makers, patients, clinicians, health plans, other payers, and regulators.

<sup>1</sup>Tufts Clinical and Translational Science Institute, Tufts University, Boston, Massachusetts, USA; <sup>2</sup>Institute for Clinical Research and Health Policy Studies, Tufts Medical Center, Boston, Massachusetts, USA; <sup>3</sup>Massachusetts Institute of Technology Political Science and Engineering Systems, Cambridge, Massachusetts, USA; <sup>4</sup>European Medicines Agency, London, UK; <sup>5</sup>US Food and Drug Administration Center for Drug Evaluation and Research, Silver Spring, Maryland, USA; <sup>6</sup>Harvard School of Public Health, Boston, Massachusetts, USA; <sup>7</sup>Tufts Center for the Study of Drug Development, Boston, Massachusetts, USA; <sup>8</sup>Merck & Co., Whitehouse Station, New Jersey, USA; <sup>9</sup>Collegeville, Pennsylvania, USA; <sup>10</sup>Tufts Cancer Center, Boston, Massachusetts, USA; <sup>11</sup>Department of Mathematics and Statistics, Boston University, Boston, Massachusetts, USA; <sup>12</sup>Harvard Clinical Research Institute, Boston, Massachusetts, USA. Correspondence: HP Selker ([hselker@tuftsmedicalcenter.org](mailto:hselker@tuftsmedicalcenter.org))

Received 20 June 2013; accepted 29 August 2013; advance online publication 4 December 2013. doi:10.1038/clpt.2013.177

- *Efficacy trials* are designed to establish the existence of a treatment effect under optimal conditions. To this end, efficacy trials use exclusion criteria that yield relatively homogeneous samples of participants, and are conducted in settings that foster adherence to carefully specified clinical protocols, often for shorter periods than those ultimately used in general patient care. Efficacy trials are used to secure regulatory approval with labeling of appropriate conditions of use.
- *Effectiveness trials* are designed to establish whether treatment effects identified in efficacy trials carry over to more typical use of drugs in practice. To this end, effectiveness trials use broader inclusion and narrower exclusion criteria, which produces heterogeneous samples that represent a usual patient spectrum; they are conducted in more varied clinical settings that approximate usual practice conditions, and they use treatment protocols that approximate typical clinical standards of care and durations of treatment.

In current practice, efficacy and effectiveness trials are conducted separately, with effectiveness trials typically done years later, if at all.<sup>1–4</sup> Therefore, approval and introduction of new drugs are currently based on results from optimized patient samples and ideal conditions, without data on effectiveness in usual clinical care that could be generalizable to the broader population and conditions in which the new medication ultimately will be used.

The path to postapproval effectiveness trials is a road not often taken. In a typical drug development program, a promising drug candidate is identified, early-phase safety and dosing studies are conducted, and targeted efficacy trials are done. If these succeed, marketing approval and publication may follow. Once a drug is approved, developers are unlikely to support an effectiveness trial. The reasons for this may be that with more heterogeneous populations and poorer signal-to-noise ratios, effectiveness trials often show lesser effects than the efficacy trials on which licensing decisions were based. Furthermore, more toxicity may be found when used with more typical regimens and in general populations.<sup>5</sup> If independent funding is found for a postapproval effectiveness trial, those conducting it may or may not have the involvement of the original investigators, may not have access to information that would inform understanding of the heterogeneity of effects in various subgroups, and may not have access to efficacy trial data that could generate predictive models to facilitate translation into widespread use. As an alternative to this uncertain and disjointed process, we suggest the integrated continuous translational research process described below.

## PROPOSED APPROACH

### Elements of integrated E2E trials

With the E2E approach, as planned from the outset, potentially including patients and other stakeholders, an effectiveness trial starts immediately after the efficacy trial component is completed. (Here, we use as context trials in which the comparator is a placebo rather than an alternative active agent, as is usually

the case in efficacy trials, but the E2E method also should be applicable to comparative efficacy and effectiveness.) The transition could be handled in a variety of ways; one could be that an external data safety monitoring board (DSMB) would review the data and analytic methods of the efficacy trial and then ratify the results, using go/no-go criteria prespecified by investigators, the sponsor, and the responsible regulatory agency. Approval by the DSMB would trigger the start of the regulatory approval process and the transition from the efficacy trial to an effectiveness trial. At this point, the composition of the DSMB for the effectiveness component would be determined, potentially retaining some members and adding others. Although a pause between the efficacy and effectiveness components might occur as sites are added and entry criteria and data collection are adjusted, because this transition was planned with criteria prespecified from the outset, it should be short and efficient. Then, as screening of populations for the effectiveness trial with less restrictive criteria ensues, some patients who would have been “screen failures” for the efficacy trial component might become appropriate subjects for the effectiveness trial component. Thereby, these results will be more applicable to a general population than the results of the efficacy trial component (**Supplementary Appendix** online).

The use of protocols more typical of usual care, with more diverse participants, and in a wider range of clinical settings may be expected to alter, perhaps more often diminish, the net (average) treatment effect. Heterogeneity of the treatment effect across the study sample may mute average effectiveness. However, this heterogeneity will also offer opportunities to detect how patient characteristics affect treatment effects. Divergence of individual and subgroup treatment effects from the overall average result will occur in both directions. In an overall negative trial, individuals or groups may be detected who benefited from treatment, whereas in an overall positive trial, individuals or groups may be detected for whom benefits did not outweigh risks.<sup>6,7</sup> These differences also could significantly change the cost–benefit analysis for use of a treatment.<sup>8</sup> It would not be warranted to use such results to alter regulatory approval that was based on efficacy data, but such insights from the effectiveness component could inform care and reimbursement decisions. Such information about the heterogeneity of treatment effects from the effectiveness trial component could be represented in multivariable predictive models, described further below, to aid treatment decisions in the care of individuals. In addition to these dimensions, in some cases, hypotheses generated by findings in the efficacy trial component about relationships between clinical characteristics and treatment effects could be prespecified and formally tested in the effectiveness portion. In such cases, the effectiveness component would include targeted data collection to test such hypotheses.

The effectiveness portion will differ from the efficacy portion in its specification of participants, treatment settings, and the protocol, but such differences will need to be compatible with the request for approval by the licensing authority on the basis of the efficacy trial component. Enrollment may be by primary-care physicians or community-based specialists rather than by special study sites, care will be more typical of

community settings, and protocol adjustments may include reductions in detailed data collection (and possibly the addition of assessments to test new hypotheses and heterogeneity of treatment effects), but the use of the drug would not be beyond the label indications being sought. Before initiation of effectiveness trials, regulatory agencies and the DSMB review the prespecified transition and the differences between effectiveness and efficacy trials, including relaxed entry criteria and longer duration, to confirm that the plan is still appropriate.

### Statistical framework

Both efficacy and effectiveness components of E2E studies will be randomized controlled clinical trials, with sample size based on prespecified  $\alpha$  and statistical power. Changes, including new hypotheses, could be incorporated into the effectiveness portion before its start, in a manner consistent with the adaptive clinical trial criteria of the US Food and Drug Administration. Although the E2E transition is not framed as an adaptive sequence, there may be circumstances in which adaptive and other features might be incorporated into the efficacy and/or effectiveness portions. Although the data from the efficacy portion might be used to formulate appropriate hypotheses for testing in the effectiveness portion, in the analysis of trial results, data from the two stages will not be combined. This facilitates interpretability of the inference and ensures control of type 1 error for each stage separately.

When data from the efficacy trial are used to formulate questions to investigate in the effectiveness portion, the hypotheses will fall into two categories. The first consists of a small number of primary or coprimary hypotheses intended for formal testing, controlling for type 1 error with appropriate adaptive or multiplicity adjustments. The second includes a larger number of secondary hypotheses, intended for exploratory investigation only, without control of type 1 error. Many hypotheses will test for interactions between treatments and preselected subgroups. To test these hypotheses, participant characteristics in the effectiveness portion must include subject attributes with which treatment effects may interact. In some settings, the efficiency of effectiveness trials could be improved by introducing adaptive interim analyses with decision rules for dropping subgroups that do not appear to benefit from the new treatment.<sup>9,10</sup> For primary hypotheses, subgroups, decision rules, and statistical methods for hypothesis testing would have to be prespecified to maintain the validity of conclusions. For secondary hypotheses, the objective would be to identify potential treatment-modifying influences to inform future hypothesis testing.

For primary hypotheses, because the effectiveness portion will have more diversity in participants and settings, standard deviations will generally be larger and overall differences smaller than in the efficacy portion. Thus, to detect a given effect, larger sample sizes will be needed. In cases in which the confirmation of effects seen in the efficacy trial is the primary objective, a larger sample size will be warranted. In other cases in which the objective is hypothesis generation or the development of predictive models, larger sample sizes may not be needed.

### E2E-derived clinical predictive instruments as decision aids

Separate from these analyses of E2E trial results, but still representing treatment effects, data from both trial components could be used in multivariable patient-specific predictive models of treatment benefits. Examples include “clinical predictive instruments,” logistic regressions that generate 0–100% predictions of diagnoses and/or outcomes as decision support in real-time care. The independent variables in these models include clinical and biological features and how patient characteristics and treatments interact, with the dependent variables being the clinical outcomes of interest, and the ultimate goal being the improvement of patient-specific treatment decisions.

An example of this approach is the thrombolytic predictive instrument (TPI), used for assessing patient-specific benefit of coronary reperfusion therapy for ST elevation myocardial infarction.<sup>11</sup> Created using data from multiple trials and registries of patients with ST elevation myocardial infarction, the TPI’s multivariable logistic regression models include key patient characteristics predictive of key clinical outcomes, and interaction variables to capture treatment effects and predict 30-day mortality, 1-year mortality, cardiac arrest, major bleeding, and stroke. When the patient’s individualized predictions of treatment benefits are printed on text headers of electrocardiograms, clinician use of reperfusion is improved for patients with less obvious ST elevation myocardial infarctions, for women with ST elevation myocardial infarction, for patients in whom reperfusion typically is underused, and for patients seen in settings with limited on-site relevant expertise.<sup>12</sup> Thereby, clinical trial data based on heterogeneous samples can be used to support patient-personalized treatment by representing the heterogeneity of treatment effect among diverse patients. Another example of a predictive model based on trial data from heterogeneous populations is the Stroke-TPI, which identifies patients who may benefit from thrombolytic therapy for stroke.<sup>13</sup> Another example predicts the probability that a patient receiving breast cancer chemotherapy will benefit from bone marrow growth factor treatment to counter suppression of white blood cells by chemotherapy.<sup>14</sup> These examples show how clinical trial data, especially with the heterogeneity of effectiveness trials, can be used to support individualized treatment.

As electronic health records become more ubiquitous, such predictive models based on E2E trial data could become electronic health record–based clinical decision support.<sup>15</sup> Thereby, individual treatment decisions could inform clinicians and patients about likely treatment outcomes on the basis of characteristics acquired from the electronic health record. In addition to indicating when one treatment option is more likely to benefit the patient than another, it also could help identify situations in which there is strong uncertainty regarding the best treatment for a patient. This is analogous to the justification for conducting a randomized clinical trial when there is “clinical equipoise” between treatment options. When equipoise is based on predictions of model-based predictive instruments, this has been termed “mathematical equipoise.”<sup>15</sup> (e.g., the TPI has been modified to predict a patient’s risk for 30-day mortality if treated and if not treated (i) with thrombolytic therapy and

(ii) by an alternative type of reperfusion, namely, percutaneous coronary intervention). By its simultaneous predictions of both treatments' outcomes (in addition to that from no treatment), this "percutaneous coronary intervention–TPI" indicates when one form of reperfusion has a better predicted outcome than the other, so it can immediately be offered to the patient.<sup>15,16</sup> It also can indicate when there is essentially equivalence—equipoise—between the two forms of reperfusion. From a research perspective, where such equivalence exists, randomization into treatment groups may be justified; from a clinical perspective, this could foster informed conversations between clinicians and patients on factors such as risk, uncertainty, preferences, and choices.

## DISCUSSION

We propose an integrated E2E trial approach with the following features:

- The first portion will be a standard randomized controlled efficacy trial, conducted with relatively homogeneous populations, powered to detect evidence of superiority over placebo or active control.
- The second portion will be an effectiveness trial, conducted in usual care settings using a more heterogeneous sample that is representative of the expected patient spectrum.
- If it is intended that a primary effect seen in the efficacy portion is to be confirmed in the effectiveness portion, the sample size of the effectiveness component will need to be increased because net average differences between treated and untreated subjects will probably be smaller and variability will probably be larger.
- For the intent of generating hypotheses on possible specific effects based on subpopulations or circumstances, samples will need to include sufficient numbers of subjects with those characteristics that may be associated with treatment effects, to permit detection of interaction effects.
- To enable development of predictive models for use in clinical settings, effectiveness trials should include heterogeneous populations and analytic approaches that allow detection of interactions between patient characteristics and treatment effects.
- To improve the E2E approach to trials, adaptive designs could be used within both the efficacy and effectiveness portions for interim assessments and modifications.

Although the E2E concept is simple, some practical barriers to implementation require consideration, including the following questions, which are addressed in the next section:

- Why would a company that has obtained, or anticipates obtaining, a positive result from an efficacy trial support an E2E effectiveness trial with the attendant possibility of a null result?
- What benefits would the E2E sequence provide for the public, for regulators, and for payers?

- What process issues need to be addressed in the implementation of E2E trials?
- What drugs and/or indications would be suitable for first applications of E2E?

### Why would a drug manufacturer support an E2E trial?

After an efficacy trial has shown significant treatment effects and licensing has been granted, manufacturers are reluctant to support follow-up effectiveness trials. If an effectiveness trial were to reveal reductions in average effects or in specific subgroups, a manufacturer might face reductions in use, limits on the scope of reimbursements, or even a reversal of regulatory approval. To create incentives for manufacturers to participate in E2E trials, a balanced solution that provides benefits to manufacturers as well as patients, payers, and providers is needed. Various aspects of this problem deserve attention:

1. Effectiveness trials conducted in usual care settings with heterogeneous populations will typically show less therapeutic effect and more adverse effects than efficacy trials. To the extent appropriate, established drug approvals based on efficacy results should not be at risk for reversal if average effects decline in effectiveness trials.
2. The exploration of heterogeneity of treatment effects may identify special benefits and/or support development of predictive instruments that facilitate optimal use. A formal avenue for incorporating such insights into marketing could improve use of treatments while creating incentives for manufacturers. Moreover, as market forces in medical care turn from fee-for-service/product to payment for measured outcomes, along with other entities, manufacturers will become increasingly responsible for ultimate outcomes. In this context, the ability to identify potentially good vs. poor responders will be important.
3. Effectiveness studies based on more heterogeneous populations may provide justification for expansion of original eligibility criteria. If an efficacy trial is based on a narrow population that excludes potential beneficiaries, conducting a broader effectiveness trial may offer a potential justification for expansion of use without risking market approval.
4. Data on subgroup-specific benefits in heterogeneous populations in effectiveness trials may provide a basis for reexamining initial negative findings on promising drugs for which expected general effects did not materialize. In an environment that increasingly focuses on patient-optimized treatments and cost-effectiveness, this approach may become a helpful part of introducing and sustaining new treatments.

One proposal that could be attractive to manufacturers could be the combination of the E2E effectiveness trial component with the initial authorization stage of recent proposals for adaptive licensing;<sup>17</sup> this could be well suited to the development and assessment of drugs with heterogeneous effects. If an E2E approach were to be embedded in proposed adaptive approaches

to drug licensing, sponsors may be able to secure market access for drugs, potentially reinforced by payers, that provide benefits for specific subgroups but do not now pass the usual thresholds for approval. From the perspective of manufacturers, this creates an opportunity to salvage drugs that would be killed in the late stages of development under the current development paradigm.

Disincentives to manufacturers could be reduced in a number of ways. If there are few agents in the marketplace, and especially if understanding heterogeneity of treatment effects might be critical to the use and marketing of a drug, the E2E effectiveness trial portion might well be attractive. For example, in Alzheimer's disease and traumatic brain injury, there are no highly successful agents today. A broad effectiveness trial that collected indicators of heterogeneity that could be used for subsequent identification of therapeutic responders, (such as specific types of illnesses, comorbidities, biomarkers, and/or genetic information) could support access to effective treatment for targeted patient groups. This also would be the case in areas in which there are some less than fully satisfactory agents, such as in Parkinson's disease or metastatic cancer. Furthermore, explanatory factors for heterogeneity of effects, identifiable by biomarkers, genetics, or other characteristics, might be attractive to all involved.

#### **How would patients, providers, payers, and regulators benefit from E2E trials?**

Information on the optimal use of treatments is a public good, with significant value to patients, providers, payers, and regulators. As with many public goods, there is underprovision of information on the effects of treatments in actual use, particularly with respect to understanding of rare adverse events and the benefits and risks for large but vulnerable subgroups.

The use of larger and more varied samples in the E2E effectiveness portion may facilitate the detection of uncommon adverse effects. The larger numbers will help uncover rare events, and inclusion of a wider range of individuals, ages, and comorbidities may expose propensities for adverse effects. Conducting the effectiveness phase of the E2E trial as soon as feasible will increase the likelihood of early detection of rare adverse effects, to the benefit of the public and all other stakeholders. This could be a useful adjunct to the use of registries and monitoring networks that seek to detect such effects in general use.

Prelicensing efficacy trials often fail to provide useful information on the benefits and risks of treatment to the most vulnerable subgroups. The elderly, the frail, and those with multiple comorbidities are an increasing segment of target populations in actual care, but these groups are often excluded from populations in efficacy trials. Together with providers and regulators, both private and public payers have expressed an interest in extending the evidence base in trials to include such underrepresented groups to improve assessments of efficacy, safety, and effectiveness.

For examples of how E2E may enable better identification of a responder population, consider the use of estrogen receptor testing to direct hormone-based treatments and the use of immunocytochemistry and gene amplification studies to define

the worth of trastuzumab (Herceptin) to treat breast cancer. The use of a companion diagnostic test for either estrogen receptor or human epidermal growth factor receptor-2 (HER2) may help identify susceptible tumors. Trastuzumab was developed with a development plan to screen-in potential responders and screen-out expected nonresponders. Yet correlating estrogen receptor or HER2 positivity and hormone or trastuzumab responsiveness remains a complex process, reflecting variability in the testing process and other factors in the population that may influence outcomes.<sup>18–21</sup> The effectiveness portion of an E2E trial might provide an opportunity to better identify a responder population in this case.

Finally, there may be benefits to patients, providers, payers, and regulators—and to manufacturers—in aligning E2E with proposed adaptive approaches to drug licensing. The E2E approach is premised on gathering information on wider effectiveness as early as possible, before full licensing, but without delaying market authorization. Under current licensing pathways, binary approval or disapproval decisions are based largely on randomized efficacy trials with confounder-cleansed populations. Adaptive approaches to licensing are intended to create a more flexible option between approval and rejection of drugs. Eichler *et al.* have assessed the benefits of adaptive approaches to licensing with iterative phases of evidence gathering and reassessment over the life cycle of drugs.<sup>17</sup> During the initial authorization phase, patient access to drugs would be managed to limit exposure to risks while information on the efficacy and safety of drugs in use is accumulated. Regulators in Europe, Singapore, Canada, and elsewhere are now actively considering initiation of pilot adaptive licensing projects with initial authorization phases. If pilots are initiated, then initial authorization periods could potentially be structured to provide information on effectiveness along with efficacy and safety. This will require consideration of the extent to which E2E would provide data needed for licensing.

#### **What design issues must be addressed to implement E2E trials?**

Although there are many nuances, the key issues for efficacy trial design are generally well understood. By contrast, the more variable design and implementation aspects of effectiveness trials are less well understood and, in the context of the E2E approach, will require further research.

First, there is no agreement on how to secure generalizable results through wider inclusion and/or on basing study samples on standards of usual care. Protocols range from what is essentially an efficacy trial in usual care settings to the broader enrollments of “pragmatic trials.”<sup>22</sup> In the E2E effectiveness portion, to minimize participant exposure to unnecessary risk, treatment should be limited to subgroups wherein previous probabilities of effectiveness and risk are most promising. Trade-offs between these goals will vary from case to case and will require institutional review board, DSMB, and regulatory review.

Second, issues of patient access to E2E trials will require further analysis. Could regulators approve a label with phased or conditional limitations on access at stages of effectiveness trials?

Would this take the form of clinical practice guidelines, or would a more complex approach be required? More generally, if the effectiveness trial takes place before licensing, what degree of discretion in controlling patient access should be accorded to sponsors and regulators? And how can this approach be adjusted to respond to the presence or absence of viable alternative treatments, especially for life-threatening conditions?

Third, work is needed on whether public and private payers should be required, or permitted, to pay for care during an E2E trial. Before and after licensing, how should the costs of effectiveness trials be supported? As data emerge on safety, efficacy, and effectiveness of drugs in use, how should payer reimbursement obligations change? In what form must such information be presented to affect payer obligations? And then how could/should emerging E2E evidence be used to set coverage and pricing at the early, middle, and late stages of the process? To what extent could/should market exclusivity horizons be modified if E2E trial results affect the timing and coverage of reimbursement?

Fourth, linked to this is the question of how emerging E2E evidence should be used to affect provider practices during and after an E2E trial and how it might be used to enhance communications with patients about what is known and what remains uncertain. Postauthorization measures may include clinical practice guidelines, changes to labels, controls of off-label use, control of qualified prescribers or facilities, mandatory participation in registry or observational study, and/or other forms of enhanced monitoring of safety and efficacy.

Fifth, in designing an E2E trial, key will be deciding which data collected in the efficacy portion should continue to be collected in the effectiveness portion, and which data not collected in the efficacy portion should be added in the effectiveness portion, including potentially quality of life and economic data. This complexity will need to be addressed in order to identify subgroup-specific effects and to create multivariable predictive models (See **Supplementary Appendix** online for Example of Such a Trial).

Finally, the heart of the E2E approach rests on generation and analysis of data on the effectiveness, safety, and efficacy of drugs that may be undergoing sponsor-administered trials and/or may already be in use. Who should have the power to compel generation and analysis of data? Who should have the obligation to pay? Who should own data on safety, efficacy, and effectiveness of drugs in use? What adjustments or changes in intellectual property rights may be needed to secure access to data needed for evaluation of effectiveness? These issues are as yet unresolved.

At this juncture, it is not clear whether current statutes enable or impede action on the specific elements of E2E trials treated in this article. Oye *et al.* address the legal foundations of adaptive licensing with attention to reimbursement in initial authorization.<sup>23</sup> Parallel work on the legal foundations of E2E is now needed.

#### Potential candidate treatments for testing in E2E trials

Factors relating to detection of heterogeneity of treatment effects could influence the choice of therapeutic area and

agent, in addition to plans for E2E trials. Prime candidates include drugs for which heterogeneity in safety, efficacy, and effectiveness are anticipated within a diagnosis. Subgroups may be defined by biomarkers, genetics, clinical characteristics, the presence of other indications, and/or the presence of other drugs or supplements that may generate positive or negative interaction effects. Alternatively, the approach could be to use E2E trials of drugs that were expected to have effects in trials of homogeneous samples but wherein the effects did not materialize. Study in more heterogeneous samples may have signals of effects in subgroups.

For early E2E trials, and perhaps in general, we suggest that study end points not be focused on mortality. In considering the transition from the end of the efficacy trial to the start of the effectiveness component, it would be ethically problematic to conduct a placebo-controlled randomized trial of an agent that has been shown in an efficacy portion to significantly reduce mortality. Exceptions might be when the target population for the effectiveness portion is so much broader than the efficacy trial that equipoise is still credible or when the trial design includes an enrollment protocol such that only those for whom there was true equipoise would be randomized, and those for whom equipoise has been dislodged would get their preferred treatment.<sup>12</sup> However, at least in early E2E trials, we suggest using easily quantifiable outcomes other than mortality.

In addition, to learn from early E2E trials, we recommend that time lines for study end points be relatively short, with the potential of showing impact within a 1-year effectiveness trial component. We also suggest that early E2E trials be conducted in clinical domains not already heavily populated by successful agents. For example, doing a trial of a new proton inhibitor for gastroesophageal reflux disorder would fit the criteria above, but considering the large number of successful agents in this therapeutic area, the benefit in conducting such a trial would be minimal. Alternatively, for clinical areas with few good therapeutic options, such as Alzheimer's disease or various advanced metastatic cancers, the benefit of assessing a promising new agent in an E2E trial would be substantial.

#### CONCLUSION

We believe that the E2E trial approach has promise for improving the evidentiary basis for adopting treatments in practice, accelerating understandings of treatment effectiveness in subgroups with particular clinical features, and providing a framework for incorporating this information into the regulatory process. It also should allow investigators to refine the use of an agent. In addition, we see E2E trial data as a basis for generating multivariable predictive instruments that could be used as decision support for treatment of individual patients and that reflect the heterogeneity of treatment effects. Finally, we see it as promoting the public good of getting the right treatment to the right patient at the right time as early as possible.

**SUPPLEMENTARY MATERIAL** is linked to the online version of the paper at <http://www.nature.com/cpt>

## ACKNOWLEDGMENTS

The authors thank Aaron Kirby for facilitating and managing this investigation and report, Abdmoain Abudabrnh for very helpful background research for this project, and Muriel Powers for expert manuscript preparation. The project was supported by the National Center for Research Resources of the National Institutes of Health under grant UL1 RR025752 and the National Center for Advancing Translational Sciences under grant UL1 TR000073. The content is solely the responsibility of the authors and does not necessarily represent the official views of the European Medicines Agency, the US Food and Drug Administration, or the National Institutes of Health.

## CONFLICT OF INTEREST

J.K.E. is participating in accrual of subjects to clinical trial NSABP B47. P.K.H. is an associate editor for *Clinical Pharmacology & Therapeutics* but was not involved in the review or decision process for this article. The other authors declared no conflict of interest.

© 2013 American Society for Clinical Pharmacology and Therapeutics

- Grines, C.L. *et al.* A comparison of immediate angioplasty with thrombolytic therapy for acute myocardial infarction. The Primary Angioplasty in Myocardial Infarction Study Group. *N. Engl. J. Med.* **328**, 673–679 (1993).
- Zijlstra, F., de Boer, M.J., Hoorntje, J.C., Reiffers, S., Reiber, J.H. & Suryapranata, H. A comparison of immediate coronary angioplasty with intravenous streptokinase in acute myocardial infarction. *N. Engl. J. Med.* **328**, 680–684 (1993).
- The Global Use of Strategies to Open Occluded Coronary Arteries in Acute Coronary Syndromes (GUSTO IIb) Angioplasty Substudy Investigators. A clinical trial comparing primary coronary angioplasty with tissue plasminogen activator for acute myocardial infarction. *N. Engl. J. Med.* **336**, 1621–1628 (1997).
- Aversano, T. *et al.*; Atlantic Cardiovascular Patient Outcomes Research Team (C-PORT). Thrombolytic therapy vs primary percutaneous coronary intervention for myocardial infarction in patients presenting to hospitals without on-site cardiac surgery: a randomized controlled trial. *JAMA* **287**, 1943–1951 (2002).
- Hassett, M.J., O'Malley, A.J., Pakes, J.R., Newhouse, J.P. & Earle, C.C. Frequency and cost of chemotherapy-related serious adverse effects in a population sample of women with breast cancer. *J. Natl. Cancer Inst.* **98**, 1108–1117 (2006).
- Kent, D.M. & Hayward, R.A. Limitations of applying summary results of clinical trials to individual patients: the need for risk stratification. *JAMA* **298**, 1209–1212 (2007).
- Kent, D.M., Rothwell, P.M., Ioannidis, J.P., Altman, D.G. & Hayward, R.A. Assessing and reporting heterogeneity in treatment effects in clinical trials: a proposal. *Trials* **11**, 85 (2010).
- Kent, D.M., Vijan, S., Hayward, R.A., Griffith, J.L., Beshansky, J.R. & Selker, H.P. Tissue plasminogen activator was cost-effective compared to streptokinase in only selected patients with acute myocardial infarction. *J. Clin. Epidemiol.* **57**, 843–852 (2004).
- Mehta, C., Gao, P., Bhatt, D.L., Harrington, R.A., Skerjanec, S. & Ware, J.H. Optimizing trial design: sequential, adaptive, and enrichment strategies. *Circulation* **119**, 597–605 (2009).
- Orloff, J. *et al.* The future of drug development: advancing clinical trial design. *Nat. Rev. Drug Discov.* **8**, 949–957 (2009).
- Selker, H.P. *et al.* Patient-specific predictions of outcomes in myocardial infarction for real-time emergency use: a thrombolytic predictive instrument. *Ann. Intern. Med.* **127**, 538–556 (1997).
- Selker, H.P., Beshansky, J.R. & Griffith, J.L.; TPI Trial Investigators. Use of the electrocardiograph-based thrombolytic predictive instrument to assist thrombolytic and reperfusion therapy for acute myocardial infarction. A multicenter, randomized, controlled, clinical effectiveness trial. *Ann. Intern. Med.* **137**, 87–95 (2002).
- Kent, D.M., Selker, H.P., Ruthazer, R., Bluhmki, E. & Hacke, W. The stroke-thrombolytic predictive instrument: a predictive instrument for intravenous thrombolysis in acute ischemic stroke. *Stroke* **37**, 2957–2962 (2006).
- Savvides, P.S., Terrin, N.C., Erban, J. & Selker, H.P. Development and validation of a patient-specific predictive instrument for the need for dose reduction in chemotherapy for breast cancer: a potential decision aid for the use of myeloid growth factors. *Support Care Cancer* **11**, 313–320 (2003).
- Selker, H.P., Ruthazer, R., Terrin, N., Griffith, J.L., Concannon, T. & Kent, D.M. Random treatment assignment using mathematical equipoise for comparative effectiveness trials. *Clin. Transl. Sci.* **4**, 10–16 (2011).
- Kent, D.M. *et al.* Comparison of mortality benefit of immediate thrombolytic therapy versus delayed primary angioplasty for acute myocardial infarction. *Am. J. Cardiol.* **99**, 1384–1388 (2007).
- Eichler, H.G. *et al.* Adaptive licensing: taking the next step in the evolution of drug approval. *Clin. Pharmacol. Ther.* **91**, 426–437 (2012).
- Mollerup, J., Henriksen, U., Müller, S. & Schönau, A. Dual color chromogenic in situ hybridization for determination of HER2 status in breast cancer: a large comparative study to current state of the art fluorescence in situ hybridization. *BMC Clin. Pathol.* **12**, 3 (2012).
- Paik, S., Kim, C. & Wolmark, N. HER2 status and benefit from adjuvant trastuzumab in breast cancer. *N. Engl. J. Med.* **358**, 1409–1411 (2008).
- Dowsett, M. *et al.* Disease-free survival according to degree of HER2 amplification for patients treated with adjuvant chemotherapy with or without 1 year of trastuzumab: the HERA Trial. *J. Clin. Oncol.* **27**, 2962–2969 (2009).
- Wolff, A.C. *et al.*; American Society of Clinical Oncology; College of American Pathologists. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. *J. Clin. Oncol.* **25**, 118–145 (2007).
- Thorpe, K.E. *et al.* A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers. *J. Clin. Epidemiol.* **62**, 464–475 (2009).
- Oye, K. *et al.* Legal foundations of adaptive licensing. *Clin. Pharmacol. Ther.* **94**, 309–311 (2013).



This work is licensed under a Creative Commons Attribution-NonCommercial-No Derivative Works 3.0 License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/3.0/>