PLOS ONE

RESEARCH ARTICLE

A framework for assessing the impact of accelerated approval

A. Lawrence Gould[1]*, Robert K. Campbell[2], John W. Loewy[3], Robert A. Beckman[4], Jyotirmoy Dey[5], Anja Schiel[6], Carl-Fredrik Burman[7], Joey Zhou[8], Zoran Antonijevic[9], Eva R. Miller[10], Rui Tang[11]

1 Methodology Research, BARDS, Merck & Co., Inc., Kenilworth, New Jersey, United States of America, 2 Molecular Pharmacology, Physiology and Biotechnology, Brown University, Providence, Rhode Island, United States of America, 3 DataForethought, Winchester, Massachusetts, United States of America, 4 Lombardi Comprehensive Cancer Center and Innovation Center for Biomedical Informatics, Georgetown University Medical Center, Washington, DC, United States of America, 5 Data and Statistical Sciences, AbbVie, North Chicago, Illinois, United States of America, 6 Department for Pharmacoeconomics, Norwegian Medicines Agency, Oslo, Norway, 7 Data Science & AI, AstraZeneca R&D, Gothenburg, Sweden, 8 Xcovery Pharmaceuticals, Palm Beach Gardens, Florida, United States of America, 9 Abond CRO, Allendale, Michigan, United States of America, 10 Independent Biostatistical Consultant, Middletown Twp, Pennsylvania, United States of America, 11 Methodology and Data Visualization, Biostatistics Department, Servier Pharmaceuticals US, Boston, Massachusetts, United States of America

* goulda@merck.com

Abstract

The FDA's Accelerated Approval program (AA) is a regulatory program to expedite availability of products to treat serious or life-threatening illnesses that lack effective treatment alternatives. Ideally, all of the many stakeholders such as patients, physicians, regulators, and health technology assessment [HTA] agencies that are affected by AA should benefit from it. In practice, however, there is intense debate over whether evidence supporting AA is sufficient to meet the needs of the stakeholders who collectively bring an approved product into routine clinical care. As AAs have become more common, it becomes essential to be able to determine their impact objectively and reproducibly in a way that provides for consistent evaluation of therapeutic decision alternatives. We describe the basic features of an approach for evaluating AA impact that accommodates stakeholder-specific views about potential benefits, risks, and costs. The approach is based on a formal decision-analytic framework combining predictive distributions for therapeutic outcomes (efficacy and safety) based on statistical models that incorporate findings from AA trials with stakeholder assessments of various actions that might be taken. The framework described here provides a starting point for communicating the value of a treatment granted AA in the context of what is important to various stakeholders.

Introduction

The FDA's Accelerated Approval program (AA) is one of four FDA programs to expedite availability of products that treat serious or life-threatening illnesses where effective treatments

do not exist or where currently available treatments are inadequately effective or unacceptably toxic (S1 Appendix). A product is a candidate for AA if it meaningfully affects a surrogate endpoint that is *reasonably likely* to predict clinical benefit or a clinical endpoint that can be measured earlier than irreversible morbidity or mortality (IMM) and is reasonably likely to predict a subsequent effect on IMM [1].

From 1992 through mid-2021, the FDA granted 269 AAs for 162 different treatments (drugs or drug combinations) [2]. Most were for infectious disease (24/162) and cancer (113/162). The creation of the AA pathway was influenced by the HIV epidemic and patient advocacy. Thirty HIV treatments received accelerated approved using surrogate endpoints and all fulfilled their post-approval study requirements [3]. The surrogate endpoints for efficacy against HIV proved to be highly predictive of clinical outcome, and were eventually adopted for traditional approval. These drugs and post-approval innovations in their use enable HIV-infected persons to have life expectancies close to the life expectancies of uninfected persons [4]. There have been more than 110 AAs for malignant hematology and oncology indications [5]. The emergence of targeted cancer drugs and immune checkpoint inhibitors was accompanied by increased use of AA in oncology, with 70 approvals from 2014–2018 [5]. Six oncology AAs granted before 2009 were withdrawn, with three of the drugs resubmitted and approved based on improved evidence supporting their use [6]. A seventh withdrawal was initiated in 2019 [7]. Almost all of the cancer AAs prior to 2016 have been converted to full approvals, while most from 2016–2018 remain to be converted.

All stakeholders (patients, physicians, sponsors of innovative treatments, reimbursement agencies, and regulators) ideally should benefit from AA. However, concern has been expressed about the value of accelerated approvals [8]. Some of the main points of conflict are over what has been measured in studies supporting accelerated approval, who has been assessed in these studies, and how long it will take for results to become available for conversion to traditional approval. AA trials provide less information than regulatory approval usually requires, and the short-term benefit/risk balance based on these trials may not reflect the true benefit/risk balance for the product. Decision-makers must rely on evidence that is less complete than what current practices provide to create cost-effectiveness assessments of new treatments [9]. Although the regulations concerning AA require sponsors to conduct clinical trials to establish efficacy based on clinical benefit, adherence in this regard has not been universal [10]. Issues affecting the completion of these requirements include equipoise concerns when a drug has been shown superior to standard care in the AA trials [11] and reluctance of patients to participate in a randomized confirmatory trial of a marketed drug available without the risks or inconvenience of research.

A common definition of the intuitively appealing term 'impact' that is used in discussions of the benefits and risks of AA can be elusive. There are many possibilities, depending on the needs and values of various stakeholders. For this discussion, impact refers to the value of suitable metrics reflecting subsequent actions that a stakeholder might take as a result of an evaluation of the information supporting AA. For example(s), a regulatory agency may decide to approve or not approve a drug via AA depending on the distribution of possible predicted metric values such as estimated probability of survival for more than 5 years regardless of promising early findings. An HTA could decide to put a drug approved via AA on its formulary for reimbursement (or not) depending on the distribution of possible predicted metric values such as whether the anticipated cost per quality-adjusted life year was within an acceptable range. A sponsor might decide whether or not to go forward with seeking AA given the metric value findings from trials that might be submitted in support seeking AA. Physicians and patients might decide to initiate treatment (or not) depending on possible predicted outcomes. "Values" in this context depends on what is important (and how important) to each

stakeholder. Stakeholder values would be elicited in any implementation. Different stakeholders will have different attributes that are important to them, and even may differ in how importantly they regard attributes that are shared with other stakeholders.

An evaluation of the impact of AA is essential for assessing its value as it applies to the stakeholders that it affects. Values for stakeholders can be unique or shared, as illustrated conceptually in Fig 1. A common process for evaluating the impact of AA can be helpful in integrating the assessments reflecting the values for the various stakeholders. These values can provide useful guidance for decisions that any stakeholder might consider.

A recent white paper from the Institute for Clinical and Economic Review [12] proposes strategies for strengthening the performance of the AA process by addressing a number of concerns that are relevant for stakeholder assessment once AA has been granted. These concerns include inconsistencies in the level of uncertainty deemed to qualify surrogate endpoints as reasonably likely to predict a clinically meaningful treatment effect; a lack of clarity over
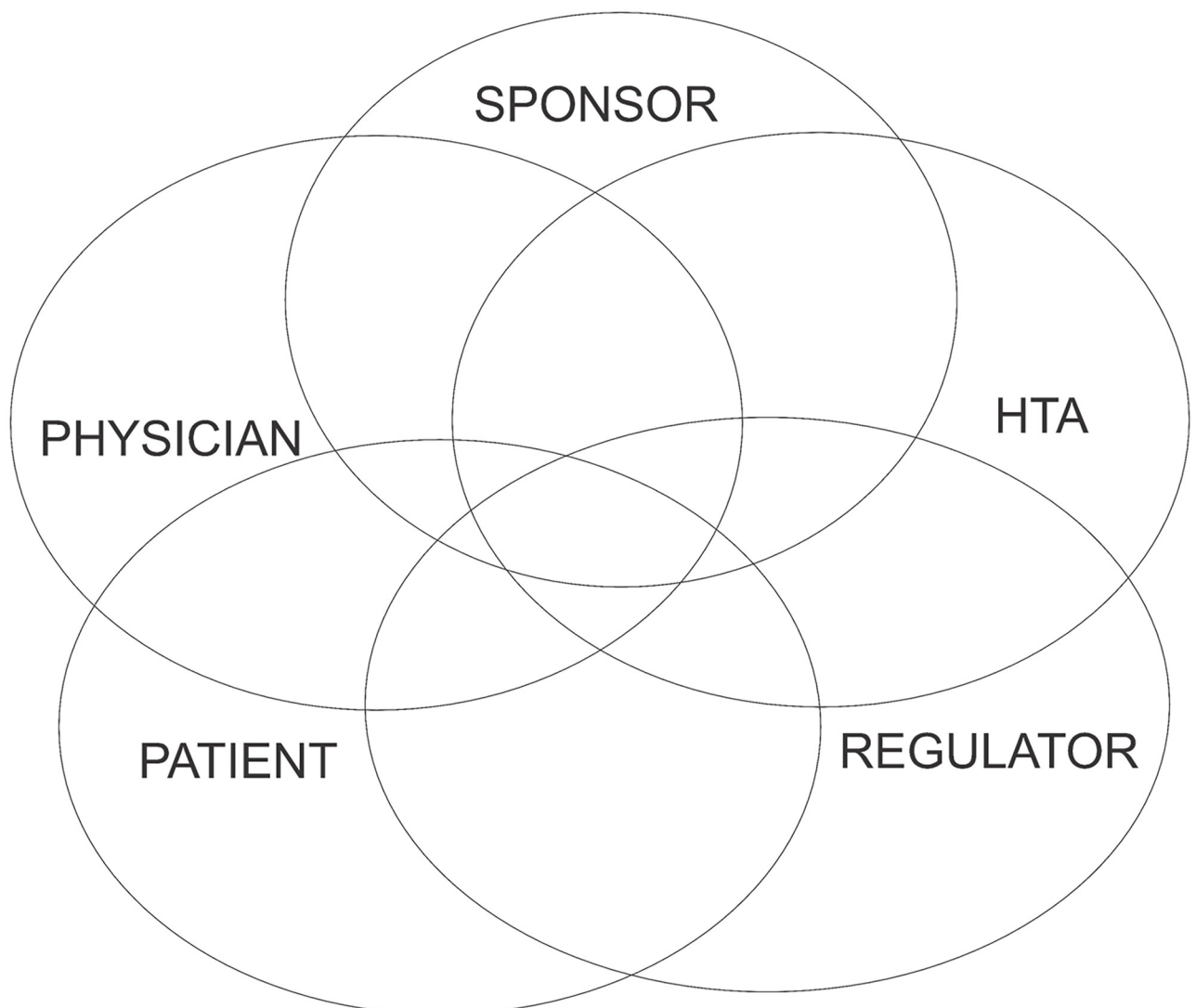

Fig 1. Unique and shared values of key AA program stakeholders.

https://doi.org/10.1371/journal.pone.0265712.g001

what magnitude of change in such outpoints justifies accelerated approval; and the high prices commanded by these products despite their relative lack of evidence.

We propose a common framework and, therefore, a common language, for addressing the challenges of decision making faced by various stakeholders inherent with the Accelerated Approval process that differs from decision science based methods that have been described previously. With this proposal, each stakeholder uses information from the AA trials to decide on a course of action that reflects stakeholder-specific values by an objective, transparent, and reproducible process that differs among stakeholders only with respect to the stakeholder-specific actions and values. This approach provides a way for stakeholders to understand the bases for each other's decisions and may be useful for informing their own decisions. The specificity of the values for different stakeholders means that the stakeholders could reach different action decisions, possibly at different points in time, so that there may be no immediate determination of 'consensus'. Consequently, the common framework approach described here differs from the consensus-based decision model for assessing health systems described by Xu *et al* [13] in which different statistical decision methods employed by stakeholders with a common objective are integrated to reach a consensus. It also differs from previous applications of decision science methods that focus on sponsor decisions about confirmatory or post-POC study designs, and how these design options could affect outcomes, utility, and future (post-approval) decisions by stakeholders [14–18].

The context differs as well from previous work in that not all stakeholders implement their decision processes contemporaneously. Accelerated Approval trials provide much drug-related (and competitive) evidence that can affect some stakeholder decisions. This evidence includes the study results, label content, FDA review documents (sometimes), and the current treatment landscape. Most importantly, since the drug has been approved to market (at least in the US), the decisions at issue are pushed from the sponsor and regulatory agencies to stakeholders such as patients, physicians, and health technology agencies that determine implementation in clinical care.

Method

Defining impact

This article outlines a process for defining 'impact' that is grounded in well-established decision theory. Although the process is objective and reproducible, its components may be objective or subjective, to reflect clinical reality, and the implementation details can be adapted to the needs of different stakeholders. For clarity of exposition, the description that follows focuses on the impact of AA on patients who receive the drug through access enabled by accelerated approval. Similar developments can be carried out for the other stakeholders. Decision theoretic concepts reflecting the values of different stakeholders have been applied to the design and analysis of clinical trials when only a subset of the patient population may be likely to benefit materially from the test drug [14, 17–19].

There are at least two key assumptions and two parts to the process. The first key assumption is that the expected <u>clinical</u> benefit of a treatment for a patient is a function of the short-term outcomes observed in the studies supporting AA, observable patient attributes, and, possibly, unobserved confounding factors. The assumption implies that the short-term benefits on biomarkers or surrogate endpoints demonstrated by these studies would predict reliably (in a probabilistic sense) the actual clinical benefits and risks for a patient undergoing long-term treatment. This could be questionable for various reasons, especially when there is a substantial lapse of time between the accelerated and the clinical endpoints. The anticipated clinical benefit generally will depend on the nature of the disease itself and the expectations of the

treatment, whether short-term or long-term, including the usually disease-specific therapeutic strategy.

The second key assumption, common to all clinical development programs, is that patients are partially or wholly exchangeable [20] so that the same model applies at least for important subgroups of the patients (e.g., men, women), certainly for those in the AA trials and ideally for all patients who might be treated.

The first part of the process entails estimation of the parameters of models relating clinical responses/outcomes resulting from the innovative treatment to the patient attributes and short term findings from AA studies, including assessments of the variability of the parameter estimates and the uncertainty about predicted clinical outcomes. Ideally, if (a) the general functional (mathematical) relationship between the attributes of a patient and the clinical response is known, (b) the exact (error-free) values of the parameters that characterize the functional relationship as it applies to a particular patient are known, and (c) there are no confounding factors then, at least in principle, it would be possible to determine the expected clinical outcome for any patient. In reality, the model and parameter values are not known exactly, although reasonable approximations may be achievable based on available data. However, even if one could determine the expected clinical outcome exactly for any patient, intrinsic variability among patients would lead to different actual responses of patients with the same expected response, so that parameter values still would need to be estimated.

The second part of the process is the use of the parameter estimates to predict the efficacy and toxicity outcome (clinical response) for a patient who is a candidate for treatment with a product that received AA. This prediction is subject to uncertainty due to intrinsic patient variability, the uncertainty associated with the estimates of the parameters used to determine the expected response, and the influence of possible confounders.

Evaluations of the impact of AA do not take place in an innovative vacuum. The overall benefit of a therapy in a given population will depend on how long it remains in use before it is replaced by superior therapeutic alternatives [19, 21]. The therapeutic alternatives at the time of accelerated approval will influence the initial impact of the AA, while subsequent approvals of other therapies for the same indication and patient population will impact the overall time course of this impact.

Impact on patients

The potential impact for a patient treated with a product receiving AA could be positive if the product conveys clinical benefit relative to the best available treatment option for the patient, or negative if it prevents the patient from receiving a better treatment or makes the patient's condition worse because of toxicity. Many other factors influence the impact on a patient, particularly those affecting quality of life such as discomfort, management of dosing, lost work hours, burdensome medical costs, and paperwork management. These can in principle be incorporated in the assessment of impact, although specifically how would depend on the particular circumstances. Not treating the patient with a truly effective product could have a negative impact because the patient would fail to achieve the product's benefit while avoiding potential toxicities of alternative treatments. Ideally, AA provides information for estimating the probabilities corresponding to possible therapeutic outcomes for a patient that can be refined as more information accumulates. Eichler *et al* [22] describe how this might be carried out.

An unintended negative impact of granting AA could be the failure of an ongoing trial of a potentially effective agent because of patient reluctance to be randomized so that future patients may not receive the drug because a reimbursement body (e.g. IQWiG [Institut für

Qualität und Wirtschaftlichkeit im Gesundheitswesen] in Germany) could conclude that insufficient evidence of benefit had been shown. Consequently, it could be more beneficial to refuse an AA application if full approval could be granted shortly afterwards when longer term data are available.

A key objective of this article is the proposal to assess the impact of AA in an objective, transparent, quantitative, and reproducible way by means of well-known statistical decision principles. S2 Appendix briefly outlines the statistical decision process used here; detailed descriptions of the theory and principles are available in the literature [23–26]. Realistic models could incorporate more detailed specification of what might be meant by 'outcome', effects of possible selection bias based on promising early results, the choice of endpoints, consideration of the influence of possible differences between the populations in the AA trials, in the confirmatory trials. and in general clinical practice. At best, caution is required because the findings from one trial may not be replicated in a second trial or in a general clinical population [27–29].

Considerations for assessing impact

The efficacy and toxicity outcomes of a therapy are important considerations for assessing the impact of AA, but not the only ones; different stakeholders may focus on different issues such as benefit/risk or cost considerations. What follows focuses for simplicity of exposition on efficacy and toxicity issues. Suppose for example, that the short-term efficacy information from the AA trials defines a predictive distribution of a relevant clinical response (e.g., progression-free survival past 6 months), and that a similar distribution of clinical response can be determined for a previous therapy. Fig 2 displays some possibilities for the distributions of the clinical response for the new and previous therapies, including the possibility that the distribution of clinical responses for the new therapy may be reflect greater variability than for the previous therapy because of less information about the clinical responses for the new therapy.

The orderings of the clinical response distributions can be expressed by various metrics. Basing the orderings of efficacy (and toxicity) on the predictive distributions of the outcomes reflects uncertainty both about the actual AA outcomes and the basis for inferring the likely clinical outcomes based on the AA results.

Suppose that the efficacy information provided in support of AA uses patient and treatment attributes to map the predicted efficacy outcomes (E = [E_1, E_2, E_3]) to the categories/ conclusions illustrated in Fig 2. These categories refer to the desired efficacy outcomes, not to the metrics used to support AA. The mapping depends on the indication, the disease, clinical judgment, and the expectations of therapy.
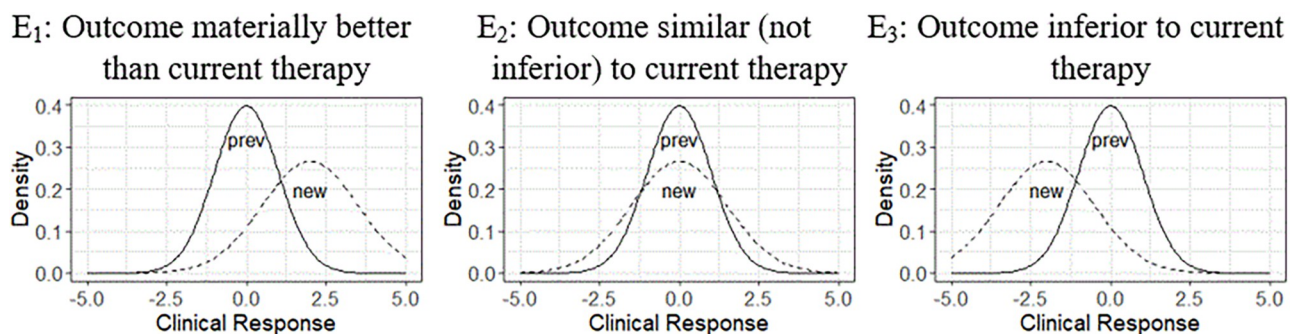

Fig 2. Orderings of clinical response for a new and a previous therapy.

Table 1. Predicted toxicity outcome.

T =		
	T_1:	Less toxic than current therapy
	T_2:	About as toxic as current therapy
	T_3:	More toxic than current therapy, but manageable
	T_4:	Intolerable toxicity

The predicted toxicity findings similarly can be assigned to various categories as in Table 1. 'Better', 'outcome', and 'toxicity' can mean different things to different people/stakeholders. We focus here on their meaning from the patient's point of view. Suppose that probabilities associated with the various efficacy and toxicity categories can be related to predictive information about a patient who could be, but has not yet been, treated with the product, and the condition to be treated. The prediction depends on the attributes of the candidate patient and the parameters of the model that relates these attributes (including treatment) to the anticipated clinical outcome. The predicted outcome is subject to uncertainty about the true values of the parameters, including the accuracy of the prediction of clinical outcome based on surrogate measures, and intrinsic variation among patients with identical attributes. This simple approach can be extended in various ways including incorporating time into the evaluation of the effect of AA.

These outcome categories can be combined into clinically meaningful categories describing predicted therapeutic outcomes for patients that combine efficacy and toxicity. These categories ordinarily will be defined in numerical terms such as expected duration of progression-free survival so that, for instance, a better outcome could be a lower confidence bound on the hazard ratio for the new treatment relative to the standard that exceeded 1.5 (i.e., 50% improvement in survival). What follows assumes that there are appropriate definitions of categories of efficacy and toxicity. Let y_P denote the predictive observations made on a candidate for treatment with the product in question. The surrogate endpoint and biomarker values obtained in the AA trials may be useful for determining predictive distributions of therapeutic outcomes and, consequently, estimates of the probabilities associated with each of the Efficacy and Toxicity outcomes if the underlying, possibly disease-specific, assumptions that are required can be satisfied. Table 2 lists the therapeutic outcome categories, the corresponding true probabilities, and the probabilities estimated from the AA findings. The potentially 12 categories corresponding to the possible combinations of efficacy and toxicity outcomes are collapsed into 6 because materially less efficacy, intolerable toxicity, or worse toxicity without materially better efficacy all identify unacceptable clinical outcomes that do not need to be considered separately.

The estimated probability of the last category $[w_6(y_P)]$ may be low because any approval, let alone AA, seems unlikely in that case. However, reality may be different if the intolerable toxicity is a rare event unobserved at the time of AA. The probabilities associated with the

Table 2. True and estimated probabilities associated with the therapeutic outcomes in Fig 2 and Table 1 for a patient with attributes y_P.

Efficacy relative to current therapy		True probabilities				AA estimated probabilities			
		Toxicity relative to current therapy				Toxicity relative to current therapy			
		Less	Similar	Worse	Intolerable	Less	Similar	Worse	Intolerable
		T_1	T_2	T_3	T_4	T_1	T_2	T_3	T_4
Materially better	E_1	$W_1(y_P)$	$W_2(y_P)$	$W_3(y_P)$	$W_6(y_P)$	$w_1(y_P)$	$w_2(y_P)$	$w_3(y_P)$	$w_6(y_P)$
Similar	E_2	$W_4(y_P)$	$W_5(y_P)$			$w_4(y_P)$	$w_5(y_P)$		
Inferior	E_3								

combined outcomes generally will depend on both the efficacy and toxicity outcomes. Efficacy and toxicity may reflect a common mechanism of action, and both are often correlated with pharmacokinetic exposure.

Statistical considerations

Details of how the statistical issues should be addressed are outside the scope of this paper because they will be situation-specific. Nonetheless, sources of uncertainty should be kept in mind when using short-term AA findings based on biomarkers and surrogate measures to predict clinical outcomes. Factors affecting the clinical outcome relevance and predictability for patients include

1. Early trial sampling variance, e.g., the standard error of the early endpoints from AA trials.

2. Selection bias that occurs because AA ordinarily would be granted only when early findings are strong. Some drugs of modest efficacy could by chance provide an early finding strong enough to justify AA, while some drugs of greater efficacy could not provide sufficiently strong early findings, also by chance, and thereby not be granted AA.

3. The correlation between the observations that justify AA and actual clinical outcomes–that is, how well the early findings predict clinical outcome.

4. How well the population represented in the AA trials represents the general population for whom the drug may be prescribed and the circumstances of its clinical use.

5. The consistency of a patient's predicted clinical outcome from treatment that reflects the variability of the attributes used to predict the outcome and the randomness of the patient's outcome given a set of attribute values.

6. The variability and sensitivity of the measures of utility to the prior specifications of uncertainty and the effect of the amount of information used for their computation.

7. Decision makers have unequal access to information. Sponsors and FDA have access to details of individual patients in the drug development programs, while other stakeholders (patients, prescribers, HTAs, and payers) do not have this information.

8. Small sample sizes: AA trials tend to include fewer patients than conventional regulatory trials. Consequently, their findings are subject to more uncertainty and may present an overly optimistic picture of treatment effect.

   For example, as illustrated in Fig 3, an analysis of 24 month survival in 41 trials of treatments for recurrent glioblastoma found that the smallest trials had the highest survival rates while the larger trials tended to have lower survival rates [30].

Determining the values of the consequences of stakeholder actions

The discussion to this point, summarized in Table 2, describes a way to conceptualize the probabilities associated with the clinical outcomes based on the information provided in the AA trials. However, these probabilities do not by themselves provide sufficient guidance for determining the impact of these trials. This impact depends on how information from the trials drives decisions about possible treatment actions. These decisions also depend on the values the stakeholder attaches to the therapeutic outcome possibilities in Table 2. A numerical measure that reflects these probabilities and values provides a quantitative way to assess impact. This approach is outlined below, and a hypothetical example is presented in the next section.
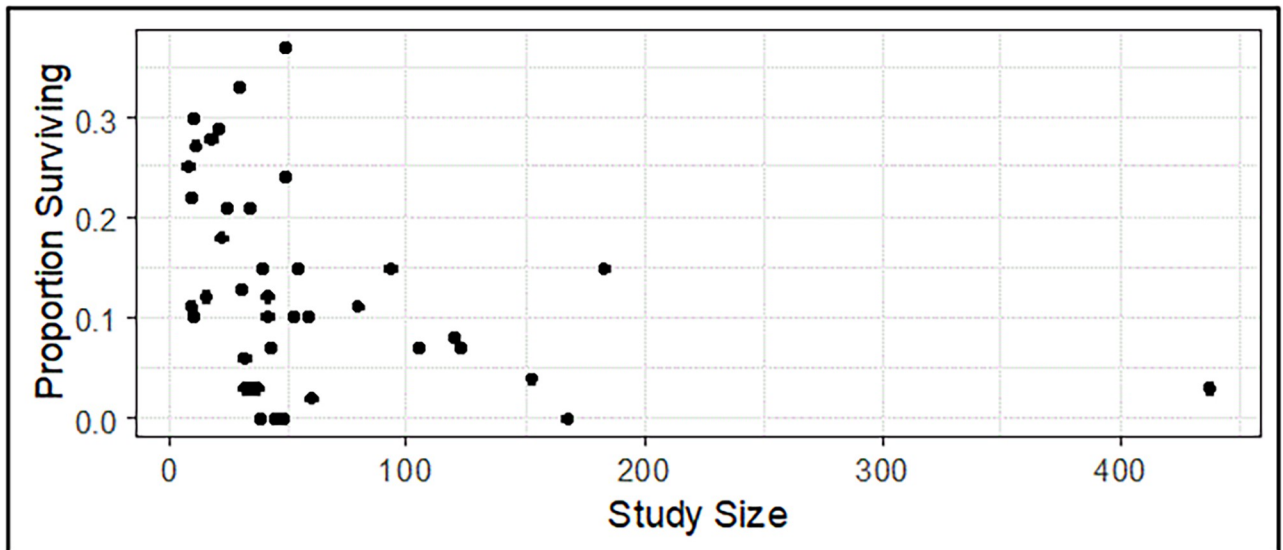
Fig 3. Proportion of patients surviving 24 months or longer in 41 trials of treatments for recurrent glioblastoma [30].

https://doi.org/10.1371/journal.pone.0265712.g003

At its most basic level, the decision process amounts to generating an action about treatment, e.g.,

$$a_1 = \text{``Refuse treatment''}, \quad a_2 = \text{``Accept treatment''}, \quad a_3 = \text{``Delay/defer treatment''},$$

The third possibility, "Delay/defer treatment" allows for some hedging, e.g., to obtain further information about the anticipated effect of treatment by consultation with other specialists, or to consider further the implications about the uncertainty of the predicted outcome due to the limited information from the AA trials. The action taken depends on the information that efficacy and toxicity findings from the AA trials provide and the anticipated relevance for the individual patient. For the present, suppose that the action taken depends on the estimated probabilities for the patient of the various clinical therapeutic outcomes, which may be all that the physician and patient know from the AA findings [31, 32].

Suppose that a stakeholder can assign a value ('utility' in decision analysis terms, or desirability) to the consequence (therapeutic outcome) of any action that the stakeholder might take. The term "value" is used in what follows to indicate a measure of desirability of an action/outcome combination from the standpoint of the stakeholder. The possible therapeutic outcomes reflecting efficacy and toxicity from Table 2 can be considered in the context of the actions from the preceding paragraph ("accept treatment", "defer/delay treatment", and "refuse treatment"). The values corresponding to the action/outcome combinations can be expressed as in Table 3, where entry v_{ij} refers to the value that the stakeholder attaches to the therapeutic outcome x_j when action a_i is taken. The actual v_{ij} values depend on who is

Table 3. Typical value table.

Action	Therapeutic Outcome			
	x_1	x_2	. . .	x_6
a_1	v_{11}	v_{12}	. . .	v_{16}
a_2	v_{21}	v_{22}	. . .	v_{26}
a_3	v_{31}	v_{32}	. . .	v_{36}

https://doi.org/10.1371/journal.pone.0265712.t003

specifying the values (e.g., physician, patient) and on the patient's status. Different stakeholders will have different value tables.

Given estimates of the probabilities of the therapeutic outcomes that might be realized as a consequence of taking an action, the value (utility, desirability) of any action can be calculated as the expectation of the values over the set of probabilities corresponding to the therapeutic outcomes.

The possible decisions/actions and outcomes can be defined and agreed to by all stakeholders, though they may not regard these as equally important. Similarly, all of the probabilities can be defined, though stakeholders may not agree on how close these are to the true probabilities (left panel in Table 2). The values, which drive the decisions are likely to be more stakeholder-dependent, and is the focus of the following paragraphs.

The expected value of each action reflects the action-dependent values that a stakeholder (e.g., patient) assigns to each therapeutic outcome, and the stakeholder's assessment of the likelihood of the outcomes corresponding to the therapeutic options.

Given true and estimated probabilities of the outcomes in Table 2 and the values in Table 3, the $\underline{\text{true values}}$ corresponding to the various actions are

$$V_{\text{True}}(a_i | W_1, W_2, W_3, W_4, W_5, W_6) \\ = W_1 v_{i1} + W_2 v_{i2} + W_3 v_{i3} + W_4 v_{i4} + W_5 v_{i5} + W_6 v_{i6}, \ i = 1, 2, 3$$

and the $\underline{\text{estimated values}}$ corresponding to the various actions are given by

$$V_{\text{est}}(a_i | w_1, w_2, w_3, w_4, w_5, w_6) = w_1 v_{i1} + w_2 v_{i2} + w_3 v_{i3} + w_4 v_{i4} + w_5 v_{i5} + w_6 v_{i6}, \ i = 1, 2, 3$$

In practice, a patient might choose the action that gave the largest estimated expected value.

$$A^* = \{a_i | V_{\text{est}}(a_i) = \max(V_{\text{est}}(a_1), V_{\text{est}}(a_2), V_{\text{est}}(a_3))\}$$

The best choice arguably would be the action a_i that maximized the expected value based on the unknown true probabilities of the various outcomes,

$$V_{\text{True}}(a_i | W_1, W_2, W_3, W_4, W_5, W_6) = W_1 v_{i1} + W_2 v_{i2} + W_3 v_{i3} + W_4 v_{i4} + W_5 v_{i5} + W_6 v_{i6}$$

[which could be quite different from $V_{\text{est}}(a_i | w_1, w_2, w_3, w_4, w_5, w_6)$], so that

$$A_{\text{opt}} = \{a_i | V_{\text{true}}(a_i) = \max(V_{\text{true}}(a_1), V_{\text{true}}(a_2), V_{\text{true}}(a_3))\}$$

The optimal action A_{opt} depends on the true probabilities. The value of the "best" course of action that the patient or physician conceivably could take based on the AA finding is $V(A_{\text{opt}})$, the metric value associated with A_{opt}. Unfortunately, this value is not known because the true clinical outcome probabilities on which it is based ($W_1 - W_6$) are unknown.

We focus on actions that the evidence in hand suggests provides the best value. What constitutes a "best" value for a stakeholder depends on the stakeholder's value table. Although, for example, a patient and a physician should have the same outcome probabilities based on the findings from the AA trial(s) and the true "State of Nature", they may have different value tables (Table 3) and so may assess the impact of the AA trials differently. Consequently, their best choice actions may differ, and will need to be reconciled so that a decision about the therapeutic course can be made. Framing the process as just described should facilitate reaching a decision because the focus is concentrated on the values, which can be the basis for discussion.

All stakeholders do similar calculations, but with different value definitions because of differing perspectives. This process described here focuses on the integration of the value matrix differences, which could be subjective, and not on the conclusions from the AA findings,

which are not subjective (or at least are less subjective). This issue has been considered in the literature [33, 34].

The different value arrays of the various stakeholders represent the values they assign to the possible actions they might take. As noted in the introduction, these potential actions will differ among stakeholders. This circumstance arises in the design of confirmatory trials based on biomarkers and when different subgroups of the general population might have different clinical outcomes from a particular therapy. Decision theoretic principles can be useful in addressing the issues that arise in these contexts [14, 17, 18].

The willingness of different stakeholders to accept uncertainty might not be aligned even though seemingly the same evidence is assessed. Reaching consensus on the value of a new course of treatment is a non-trivial task that requires careful attention to the underlying issues [6, 8]. A process that requires stakeholders to explicitly state probabilities and expectations may facilitate better recognition and understanding of these conflicts. Having a common process (and metric) should help discussions of policy and strategy with respect to AA.

For example, HTAs and payers often require more complex studies that allow a robust prediction of the future likely use of a new drug in the context of diverse national healthcare systems. AA development programs should consider the possibility that the shortcomings of the AA process could hamper HTA assessments. This means that discussion and dialogue with HTA/Payer organizations as well as regulators aimed at prior consensus or agreement should be undertaken before embarking on an AA development program.

Illustrative example

The following simple example illustrates the basic ideas of the proposed method. Its application in practice could be appreciably more complex.

Findings from an AA trial

Suppose that Accelerated Approval is based on the findings from a single-arm trial of 50 or perhaps 100 patients carried out in a homogeneous patient population to evaluate the effect of combining a new treatment with "standard of care" (SoC) in the treatment of a rare cancer for which there is no effective treatment. The AA findings may be based on whether a surrogate endpoint such as a biomarker demonstrates positive finding known not to occur with SoC. Let P denote the true, unknown probability that a patient receiving the new treatment demonstrates a positive surrogate endpoint effect (SE). The trial will be considered a success if the outcome of the trial demonstrates that P is likely to exceed 20%. Toxicity also may be an issue, so an estimate of the value of Q, the probability that a patient receiving the new treatment experiences a severe toxic event, also will be obtained.

The analysis of the AA trial could be carried out using a conventional frequentist or a Bayesian approach. The Bayesian approach will be used for the purpose of illustration; conventional frequentist calculations yield similar results. Suppose, *a priori*, that one expects P to be about 10% and Q to be about 20%. These are pessimistic assumptions because the success of the trial depends on demonstrating that P is larger than 0.20. For the purpose of analysis, the prior distributions of P and Q are taken to be beta distributions with parameters (1, 9) for P and (2, 8) for Q; the expected values of these distributions are 0.10 and 0.20, respectively. Since the trial is relatively small, the effect of sampling variability of the estimates of P and Q based on the outcome of the trial needs to be taken into account, e.g., by credible (or confidence) intervals.

Suppose first that N = 50, 15 of the 50 patients had a positive SE effect, and none had a severe toxicity event. This means that the posterior distribution of P is a beta distribution with

parameters $(1+15, 9+35) = (16, 44)$. Consequently, the posterior expected value of P is 16/60 = 0.27, a 95% credible interval for P is (0.16, 0.38), and the posterior probability that P > 20% is 0.88. This could be regarded as a successful outcome if being 88% 'sure' that P > 20% defines 'success'. The posterior distribution of Q is a beta distribution with parameters (2, 58) so that the posterior expected value of Q is 2/60 = 0.03 and a 95% credible interval for Q is (0.004, 0.09).

If N = 100 with 30 of the patients having a positive SE effect and none having a severe toxicity event, then the posterior distribution of P is a beta distribution with parameters (31, 79) so that the expected value of P is 31/110 = 0.28, a 95% credible interval for P is (0.20, 0.37), and the posterior probability that P > 20% is 0.98, strongly suggesting that the trial had a 'successful' outcome. The posterior distribution of Q is a beta distribution with parameters (2, 108) so that the expected value of Q is 2/110 = 0.018 and a 95% credible interval for Q is (0.002, 0.05). That is, given the trial outcome, the expected risk of toxicity is about 2%, not 20% as assumed at the outset.

Predicting clinical outcome

Suppose that this finding is sufficiently promising, and the need for therapy sufficiently dire, so that AA may be granted for adding the new treatment to SoC. A key issue at this point is the reliability of the clinical outcome predictions based on the trial findings, as outlined in Table 2. The SE may or may not turn out to be a good surrogate for a tangible clinical benefit (CB) such as surviving for 6 months or for 5 years for patients with this cancer. Different stakeholders may attach quite different values to the importance of the SE as a predictor of the probability of demonstrating CB.

In reality, demonstrating a positive SE does not guarantee CB nor does failure to demonstrate a positive SE rule out the possibility of CB. Suppose for the sake of illustration that a patient demonstrating a positive SE has a 70% chance of demonstrating CB with treatment, i.e., Prob(CB|SE) = 0.7. Suppose also that a patient who does not demonstrate a positive SE has a 100q% chance of achieving CB with treatment, i.e., Prob(CB|no SE) = q. CB, being the best outcome, corresponds to efficacy outcome E_1 in Fig 2. If P(SE) = 0.28, then the probability that a treated patient achieves CB is

$$P(E_1) = Prob(SE) \times Prob(CB|SE) + (1 - Prob(SE)) \times Prob(CB|no\ SE) = 0.196 + 0.72q$$

We assume for illustration that achieving CB without demonstrating SE is arguably unlikely if the surrogate effect is based on a biomarker that has been demonstrated to have experimental validity. This means that for this example, the value of q is likely to be small, say less than 0.15, so that a patient who does not demonstrate a positive SE effect with treatment has somewhere between a 0% and a 15% chance of demonstrating CB. With this assumption, the estimated probability of achieving CB, $\hat{P}(E_1)$, lies between 0.20 and 0.30. For illustration, assume q = 0.05, so that $\hat{P}(E_1) = 0.23$ when Prob(SE) = 0.28.

Inferiority of clinical effect is irrelevant in this scenario because there are no effective treatments. Consequently, $\hat{P}(E_2) = 0.77$ and $P(E_3) = \hat{P}(E_3) = 0$. With respect to toxicity, adding a treatment with a potentially toxic effect cannot decrease the probability of severe toxicity, so T_1 cannot occur i.e., $P(T_1) = 0$. The expected toxicity rate based on the AA trial is 3%, so $\hat{P}(T_4) = 0.03$. Consequently, $\hat{P}(T_2\ or\ T_3) = 0.97$. For illustration, suppose that $\hat{P}(T_2) = \hat{P}(T_3) = 0.485$.

Again, even though it is not necessarily realistic, nor is it assumed for Table 2, suppose for simplicity of exposition that efficacy and toxicity outcomes are independent. Then the probabilities in Table 2 can be summarized in Table 4 for the various cases. The expected values of P

Table 4. Predictive probabilities of therapeutic outcomes as a function of anticipated values of P (probability of positive SE) and Q (probability of severe toxicity) when 15 out of 50 patients in a trial demonstrate SE and no patients demonstrate severe toxicity.

				Predicted Therapeutic Outcome					
			Efficacy	Better	Better	Better	Similar	Similar	
	P	Q	$p(E_1)$	Toxicity	Less	Similar	Manageable	Less	Similar	Other
Expected	0.27	0.03	0.23		0	0.11	0.11	0	0.38	0.41
Worst Case	0.16	0.09	0.15		0	0.07	0.07	0	0.41	0.44
Best Case	0.38	0.004	0.30		0	0.14	0.14	0	0.34	0.37

The values corresponding to a sample of 100 patients with 30 demonstrating SE are nearly the same. $P(E_1)$ = estimated probability of clinical benefit (CB).

https://doi.org/10.1371/journal.pone.0265712.t004

Table 5. Illustrative scenarios for values of P = Prob(SE) and Q = Prob(Toxicity).

	True				E	Better	Better	Better	Similar	Similar	
Scenario	P	Q	Efficacy	Toxicity	T	Less	Similar	Manageable	Less	Similar	Other
1	0.4	0.03	Better	Same		0	0.11	0.11	0	0.38	0.41
2	0.4	0.1	Better	Worse		0	0.10	0.10	0	0.35	0.45
3	0.3	0.03	Same	Same		0	0.08	0.08	0	0.40	0.43
4	0.3	0.1	Same	Worse		0	0.08	0.08	0	0.37	0.47
5	0.1	0.03	Worse	Same		0	0.03	0.03	0	0.46	0.49
6	0.1	0.1	Worse	Worse		0	0.03	0.03	0	0.43	0.53

https://doi.org/10.1371/journal.pone.0265712.t005

and Q are as described above. The Worst Case values (pessimistic assessment) consist of the lower bound for P and the upper bound for Q, while the Best Case values (optimistic assessment) consist of the upper bound for P and the lower bound for Q.

The entries in Table 4 correspond to the estimated probabilities in Table 2 and assumptions about the predictability of CB from SE, and illustrate the effects of sampling variability and AA trial sample size. The true probabilities of the various therapeutic outcomes are needed to assess the impact of AA. Table 5 provides the true predicted probabilities of the therapeutic outcomes for each of six illustrative scenarios exploring the effect of various values of the probabilities of SE and a toxic effect.

Values corresponding to possible actions

In this example, "similar efficacy" is the same as "no efficacy" because there are no effective treatments. Likewise, "refusing treatment" is the same as 'deferring/delaying treatment". Table 6 provides two sets of values for the entries in Table 4. These two value sets could correspond to the values for two different patients or to a patient and the physician guiding the patient's treatment.

Accepting treatment when the efficacy is not "better" has a negative value because there is a potential economic issue due to the cost of the new medication. If the "True Outcome" (State of Nature) is "Better Efficacy and Less Toxicity", then the patient (and the physician) might set the value of accepting treatment (a_2) at 100 or maybe 1000 because it's good to get the better treatment. There is, of course, no necessity for the patient and the physician to have the same value tables.

If, on the other hand, the "True Outcome" is anything other than at least similar efficacy and no worse than manageable toxicity, then the patient might set the value of accepting treatment at -50, because getting a worse or very intolerable treatment would be unacceptable. The

Table 6. Value table for example.

Value Set	Action	Efficacy Toxicity	Better Less	Better Similar	Better Manageable	Similar Less	Similar Similar	Other Outcomes
					True Outcome			
1	a_1	Refuse	--	-100	-100	--	0	50
	a_2	Accept	--	100	100	--	-50	-50
	a_3	Defer/Delay	--	-100	-100	--	0	50
2	a_1	Refuse	--	-1000	-1000	--	0	100
	a_2	Accept	--	1000	1000	--	-50	-500
	a_3	Defer/Delay	--	-1000	-1000	--	0	100

Table 7. Expected utilities (values) corresponding to alternative courses of action as a function of the outcome of the AA trial and the unknown true probabilities of achieving a positive SE effect or experiencing severe toxicity when N = 50.

(P, Q values)	Value Set 1 Refuse	Value Set 1 Accept	Value Set 1 Defer/Delay	Value Set 2 Refuse	Value Set 2 Accept	Value Set 2 Defer/Delay
	Action (Decision)			Action (Decision)		
Expected	7.5	-28.1	7.5	-101.9	-96.0	-101.9
Worst Case	17.1	-37.8	17.1	-31.1	-191.3	-31.1
Best Case	-1.4	-18.2	-1.4	-172.3	-5.7	-172.3
True Scenario 1	-1.4	-17.4	-1.4	-176.6	-4.7	-176.6
True Scenario 2	2.3	-19.8	2.3	-156.7	-40.5	-156.7
True Scenario 3	5.4	-25.6	5.4	-119.6	-74.0	-119.6
True Scenario 4	8.6	-27.3	8.6	-103.8	-104.7	-103.8
True Scenario 5	19.0	-41.9	19.0	-5.5	-212.5	-5.5
True Scenario 6	21.2	-42.4	21.2	2.1	-233.2	2.1

The values for N = 100 are similar.

value of refusing treatment (a_1) if the treatment offers better efficacy and less toxicity might be -100 because it might be tragic to miss an important therapeutic opportunity. Finally, refusing or deferring treatment might imply that the value should be 100 or 1000 because it would be highly desirable to avoid intolerable toxicity or opportunity loss from ineffective treatments.

Table 7 combines the values (2 sets) in Table 6 with the estimated and true therapeutic outcome probabilities in Tables 4 and 5 to give the estimated and true values corresponding to the various scenarios, choices of bounds, and sample sizes. This table provides the basis for assessing how a patient (or physician) might prefer a course of action.

The values of P and Q in any of the cases or under any of the True Scenarios are not under the control of the patient or physician (or, in general, of any stakeholder). The stakeholders control the values associated with each decision/therapeutic outcome combination, and these values are crucial for identifying appropriate decisions. A striking finding from the Value Set 1 part of Table 7 is that accepting treatment would not be the best course of action essentially regardless of the values of P and Q assumed here, so that the treatment is unlikely to be useful or acceptable to the patient. This could mean, in evaluating the treatment for AA, that AA might not be appropriate at all. However, increasing the value of a CB (Value Set 2) implies that accepting treatment could lead to the "largest" utilities and therefore be a preferred

decision under the "Expected" and "Best Case" values of P and Q or True Scenarios 1–3 even if a CB is unlikely.

That is not to say, of course, that if Value Sets 1 and 2 corresponded to the values of a patient and a physician, reflecting what each thought was important, that the action that the physician would recommend would be the same as the action that the patient's analysis revealed (or the same as what the analysis of any other stakeholder, e.g., an HTA) revealed. Resolving the differences would require discussion and negotiation, but ideally could be simplified because all of the stakeholders were using the same metric.

Of course, decision makers do not know which scenario actually applies in any real situation. One way to resolve the uncertainty as to which scenario to accept and, therefore, which treatment option to take, would be to use historical information from the AA trial(s) or other sources if available to estimate probabilities corresponding to the true scenarios. The confidence one might have in the reliability of the decision certainly would depend on the strength of the historical evidence, including the size of the AA trial(s), as indicated in the Statistical considerations section above.

Discussion

In a 21st century clinical development paradigm, the AA process represents an avenue for bringing potentially life-saving new medical treatments to patients sooner. However, drugs approved via the AA approach may turn out not to provide actual clinical benefit. AA s generally are based on fewer, smaller, or shorter clinical trials than are used to support conventional regulatory approval, so that there will be less information about the occurrence of rare or delayed adverse events that may emerge over time and affect the label [35]. The uncertainty about the actual clinical value of a new treatment receiving regulatory approval based on surrogate endpoints is the risk that must be accepted in order to have expedited availability of new products that patients can use as prescribed by their medical providers. Product performance may become increasingly more difficult to evaluate as a result of a provision in section 3022 of the 21st Century Cures Act that requires the FDA to consider lower standards of evidence, including 'real world evidence".

Under AA, health system stakeholders need to be able to communicate with each other using standardized metrics so that the health value of the new treatment that has just been approved can continue to be evaluated for its meaningfulness as the treatment sponsor continues clinical development activities. The principles described above are an attempt to formulate a process for the evaluation. Such a process is needed and important because all stakeholders (patients, physicians, sponsors of the medical treatment being developed, regulatory authorities and payers) need to understand how the new product performs for patients for whom the product is indicated.

The framework proposed here differs from decision science based methods that have been described previously in that stakeholders decide on a course of action that reflects stakeholder-specific values by an objective, transparent, and reproducible process that differs among stakeholders only with respect to the stakeholder-specific actions and values. Stakeholders could reach different action decisions, possibly at different points in time, so that there may be no immediate determination of 'consensus' (as opposed to the decision model described by Xu *et al* [13]). It also has a different aim from methods that focus on sponsor decisions about confirmatory or post-POC study designs, and how these design options could affect outcomes, utility, and future (post-approval) decisions by stakeholders. The context differs as well from previous work in that not all stakeholders implement their decision processes

contemporaneously because the AA process pushes important decisions from the sponsor and regulatory agencies to stakeholders that determine implementation in clinical care.

Although the proposed process is general, its implementation provides a starting point on which the health systems stakeholders can focus while preparing their own internal systems for capturing the relevant information necessary in a paradigm where the product is now available to the public based on AA. The current regulatory paradigm has been evolving to address criticisms that it has been holding up the ability of potentially lifesaving treatments to reach patients for whom effective treatment alternatives do not exist. The ideas presented here describe a process for facilitating the effectiveness of the paradigm as it evolves by providing value added communication among all stakeholders.

This description of the process for assessing the impact of AA has omitted many implementation details in the interest of clarity and brevity that we anticipate addressing in future articles. For example, the interpretation of the outcomes of AA trials and their use to attach probabilities to predicted outcomes depends on the attributes and design of the AA trials such as whether the trials had 1 or 2 arms, the nature of the controls, if any (e.g., active, standard of care, placebo/no control), the nature of the response (dichotomous, discrete, continuous), etc. Additional details include the models used to obtain predictive distributions for clinical outcomes based on the outcomes of AA trials and historical information, models for regulator-sponsor interdependence, continuous outcomes, and the effect of dependence of efficacy and toxicity outcomes. Moreover, the biology of the underlying disease may be influential in ascertaining the true impact of AA [36, 37].

Greater uncertainty about perceived treatment effectiveness may increase the hesitancy of payers to provide reimbursement to patients choosing to receive new treatments. Sponsors may accept greater uncertainty than would be optimal from a public health perspective [17, 18]. Payers will often use the concept of number needed to treat (NNT) or number needed to harm (NNH) to evaluate this uncertainty to make reimbursement decisions [38, 39]. Other considerations also are important, such as the potential Incremental Cost Effectiveness Ratio of a new treatment compared with a "best" alternative treatment based on key efficacy variables and the cost of approved screening tests for targeted therapies.

Supporting information

S1 Appendix. Accelerated approval [1, 35, 40, 41].
(DOCX)

S2 Appendix. Elements of a statistical decision framework.
(DOCX)

S1 File.
(ZIP)

Acknowledgments

Author Contributions

Conceptualization: A. Lawrence Gould, Robert K. Campbell, Robert A. Beckman, Anja Schiel, Carl-Fredrik Burman, Zoran Antonijevic.

Investigation: A. Lawrence Gould, John W. Loewy.

Methodology: A. Lawrence Gould.

Project administration: Robert K. Campbell.

Visualization: A. Lawrence Gould, Robert K. Campbell, John W. Loewy, Robert A. Beckman, Jyotirmoy Dey, Zoran Antonijevic, Eva R. Miller.

Writing – original draft: A. Lawrence Gould, Robert K. Campbell, Zoran Antonijevic.

Writing – review & editing: A. Lawrence Gould, Robert K. Campbell, John W. Loewy, Robert A. Beckman, Jyotirmoy Dey, Anja Schiel, Carl-Fredrik Burman, Joey Zhou, Zoran Antonijevic, Eva R. Miller, Rui Tang.

References

1. US Food and Drug Administration. Guidance for Industry: Expedited Programs for Serious Conditions —Drugs and Biologics 2017 [https://www.fda.gov/downloads/Drugs/ GuidanceComplianceRegulatoryInformation/Guidances/UCM358301.pdf.

2. US Food and Drug Administration. CDER Drug and Biologic Acclerated Approvals Based on a Surrogate Endpoint (June 30, 2021)2021 October 8, 2021. https://www.fda.gov/media/151146/download.

3. Woodcock J. Expediting drug development for serious illness: Tradeoffs between patient access and certainty. Clinical Trials. 2018; 15(3):230–4. https://doi.org/10.1177/1740774518770656 PMID: 29871508

4. Yoshimura K. Current status of HIV/AIDS in the ART era. Journal of Infection and Chemotherapy. 2017; 23(1):12–6. https://doi.org/10.1016/j.jiac.2016.10.002 PMID: 27825722

5. Beaver JA, Howie LJ, Pelosof L, Kim T, Liu JZ, Goldberg KB, et al. A 25-Year Experience of US Food and Drug Administration Accelerated Approval of Malignant Hematology and Oncology Drugs and Biologics A Review. JAMA Oncology. 2018; 4(6):849–56. https://doi.org/10.1001/jamaoncol.2017.5618 PMID: 29494733

6. Califf RM. Expedited and facilitated drug evaluations and evidence of benefit and risk: The cup is half-full. Clinical Trials. 2018; 15(3):235–9. https://doi.org/10.1177/1740774518771347 PMID: 29871505

7. Anon. ANNOUNCE prompts questions over the Accelerated Approval process. Nature Reviews Clinical Oncology. 2019; 16(August 2019):459.

8. Ladanie A, Speich B, Naudet F, Agarwal A, Pereira TV, Sclafani F, et al. The comparative effectiveness of innovative treatments for cancer (CEIT-Cancer) project: Rationale and design of the database and the collection of evidence available at approval of novel drugs. Trials. 2018;19. https://doi.org/10.1186/ s13063-018-2877-z PMID: 30231912

9. Woolacott N, Corbett M, Jones-Diette J, Hodgson R. Methodological challenges for the evaluation of clinical effectiveness in the context of accelerated regulatory approval: an overview. Journal of Clinical Epidemiology. 2017; 90:108–18. https://doi.org/10.1016/j.jclinepi.2017.07.002 PMID: 28709997

10. Wallach JD, Ross JS, Naci H. The US Food and Drug Administration's expedited approval programs: Evidentiary standards, regulatory trade-offs, and potential improvements. Clinical Trials. 2018; 15 (3):219–29. https://doi.org/10.1177/1740774518770648 PMID: 29871509

11. Laage T, Loewy JW, Menon S, Miller ER, Pulkstenis E, Kan-Dobrosky N, et al. Ethical considerations in adaptive design clinical trials. Therapeutic Innovation & Regulatory Science. 2017; 51(2):190–9. https:// doi.org/10.1177/2168479016667766 PMID: 30231727

12. Kaltenboeck A, Mehlman A, Pearson SD. Strengthening the accelerated approval pathway: An analysis of potential policy reforms and their impact on uncertainty, access, innovation, and costs. Institute for Clinical and Economic Review White Paper April 26, 2021. URL https://icer.org/assessment/fda- accelerated-approval-pathway/

13. Xu Y, Lai KK, Leung WKJ. A consensus-based decision model for assessing the health systems. Plos One. 2020; 15(8). https://doi.org/10.1371/journal.pone.0237892 PMID: 32810150

14. Beckman RA, Burman CF, Chen C, Jobjornsson S, Konig F, Stallard N, et al. Decision analysis from the perspectives of single and multiple stakeholders. In: Antionijevic Z, Beckman RA, editors. Platform Trials in Drug Development: Umbrella Trials and Basket Trials. Boca Raton, FL: Chapman & Hall/CRC; 2018. p. 141–52.

15. Miller F, Burman CF. A decision theoretical modeling for Phase III investments and drug licensing. Journal of Biopharmaceutical Statistics. 2018; 28(4):698–721. https://doi.org/10.1080/10543406.2017.1377729 PMID: 28920757

16. Ondra T, Dmitrienko A, Friede T, Graf A, Miller F, Stallard N, et al. Methods for identification and confirmation of targeted subgroups in clinical trials: A systematic review. Journal of Biopharmaceutical Statistics. 2016; 26(1):99–119. https://doi.org/10.1080/10543406.2015.1092034 PMID: 26378339

17. Ondra T, Jobjornsson S, Beckman RA, Burman CF, Konig F, Stallard N, et al. Optimizing trial designs for targeted therapies. Plos One. 2016; 11(9). https://doi.org/10.1371/journal.pone.0163726 PMID: 27684573

18. Ondra T, Jobjornsson S, Beckman RA, Burman CF, Konig F, Stallard N, et al. Optimized adaptive enrichment designs. Statistical Methods in Medical Research. 2019; 28(7):2096–111. https://doi.org/10.1177/0962280217747312 PMID: 29254436

19. Pearce M, Hee SW, Madan J, Posch M, Day S, Miller F, et al. Value of information methods to design a clinical trial in a small population to optimise a health economic utility function. Bmc Medical Research Methodology. 2018;18.

20. Greenland S, Draper D. Exchangeability. In: Armitage P, Colton T, editors. Encyclopedia of Biostatistics. Chichester: Wiley; 1998. p. 1426–7.

21. Raju GK, Gurumurthi K, Domike R, Theoret MR, Pazdur R, Woodcock J. Using a Benefit-Risk Analysis Approach to Capture Regulatory Decision Making: Melanoma. Clinical Pharmacology & Therapeutics. 2019; 106(1):123–35. https://doi.org/10.1002/cpt.1461 PMID: 30993685

22. Eichler HG, Oye K, Baird LG, Abadie E, Brown J, Drum CL, et al. Adaptive Licensing: Taking the next step in the evolution of drug approval. Clinical Pharmacology & Therapeutics. 2012; 91(3):426–37. https://doi.org/10.1038/clpt.2011.345 PMID: 22336591

23. Berger JO. Statistical Decision Theory and Bayesian Analysis. 2 ed. New York: Springer; 1985.

24. DeGroot MH. Optimal Statistical Decisions. New York: Wiley; 2004.

25. Parmigiani G, Inoue LYT. Decision Theory. Chichester: Wiley; 2009.

26. Pratt JW, Raiffa H, Schlaifer R. Introduction to Statistical Decision Theory. Cambridge, MA: MIT Press; 1995.

27. Ioannidis JPA. Why most published research findings are false. PLoS Medicine. 2005; 2(8):e124. https://doi.org/10.1371/journal.pmed.0020124 PMID: 16060722

28. Ioannidis JPA. Why most clinical research is not useful. PLoS Medicine. 2016; 13(6).

29. Pocock SJ, Stone GW. The primary outcome is positive—Is that good enough? New England Journal of Medicine. 2016; 375(10):971–9. https://doi.org/10.1056/NEJMra1601511 PMID: 27602669

30. Chiocca EA, Nassiri F, Wang J, Peruzzi P, Zadeh G. Viral and other therapies for recurrent glioblatoma: is a 24-month durable response unusual? Neuro-Oncology. 2019; 21(1).

31. Marks HM, Coleman ME. Scientific data and theories for salmonellosis dose-response assessment. Human and Ecological Risk Assessment. 2017; 23(8):1857–76.

32. Small MJ, Güvenç Ü, DeKay ML. When can scientific studies promote consensus among conflicting stakeholders? Risk Analysis. 2014; 34(11):1978–04. https://doi.org/10.1111/risa.12237 PMID: 24954376

33. Rao S, Beckman RA, Riazi S, Yabar CS, Boca SM, Marshall JL, et al. Quantification and expert evaluation of evidence for chemopredictive biomarkers to personalize cancer treatment. Oncotarget. 2017; 8(23):37923–34. https://doi.org/10.18632/oncotarget.13544 PMID: 27888622

34. Sharma V, Fong A, Beckman RA, Rao S, Boca SM, McGarvey PB, et al. Eye-tracking study to enhance usability of molecular diagnostics reports in cancer precision medicine. Journal of Clinical Oncology Precision Oncology. 2018;2. https://doi.org/10.1200/PO.17.00296 PMID: 35135129

35. Mostaghim SR, Gagne JJ, Kesselheim AS. Safety related label changes for new drugs after approval in the US through expedited regulatory pathways: retrospective cohort study. Bmj-British Medical Journal. 2017; 358:j3837. https://doi.org/10.1136/bmj.j3837 PMID: 28882831

36. Kelly RJ, Hillner BE, Smith TJ. Cost effectiveness of crizotinib for anaplastic lymphoma kinase-positive, non-small-cell lung cancer: Who is going to blink at the cost? Journal of Clinical Oncology. 2014; 32(10):983–5. https://doi.org/10.1200/JCO.2013.54.6002 PMID: 24567437

37. Rusthoven CG, Doebele RC. Management of brain metastaces in ALK-positive non-small-cell lung cancer. Journal of Clinical Oncology. 2016; 34(24):2814–1819. https://doi.org/10.1200/JCO.2016.67.2410 PMID: 27298405

38. Hutton JL. Misleading statistics: The problems surrounding number to treat and number needed to harm. Pharmaceutical Medicine. 2010; 24(3):145–9.

39. Laupacis A, Sackett DL, Roberts RS. An assessment of clinically useful measures of the consequences of treatment. New England Journal of Medicine. 1988; 318(26):1728–33. https://doi.org/10.1056/NEJM198806303182605 PMID: 3374545

40. Edwards KT. The role of patient participation in drug approvals: Lessons from the accelerated approval of Eteplirsen. Food and Drug Law Journal. 2017; 72(3):406–50.

41. Naci H, Smalley KR, Kesselheim AS. Characteristics of Preapproval and Postapproval Studies for Drugs Granted Accelerated Approval by the US Food and Drug Administration. JAMA-Journal of the American Medical Association. 2017; 318(7):626–36. https://doi.org/10.1001/jama.2017.9415 PMID: 28810023