A Maturity Model for the Scientific Review of Clinical Trial Designs and Their Informativeness

Shawn Dolley *, Thea Norman , Doug McNair , Dan Hartman

Posted Date: 10 April 2023

doi: 10.20944/preprints202304.0147.v1

Keywords: clinical trial; informativeness; design review; trial methods maturity model

Article

A Maturity Model for the Scientific Review of Clinical Trial Designs and Their Informativeness

Dolley S [1,*], Norman T [2], McNair D [2] and Hartman D [2]

[1] Open Global Health, Arlington, Virginia, USA
[2] The Bill & Melinda Gates Foundation, Seattle, Washington, USA
* Correspondence: shawn.dolley@gmail.com

Abstract: Many clinical trials end uninformatively. Informativeness, in the context of clinical trials, defines whether a study's results definitively answer its research questions with meaningful next steps. One subset of these trials are those focused on global health set in low-resource settings. Global health clinical trials benefitting people in low-resource settings are funded primarily by a limited number of large foundations, pharmaceutical firms ("industry"), and national governments. While clinical trial protocols are required to go through reviews in regulatory and ethical domains, outside of industry-funded trials, funders rarely require focused scientific design reviews. There are no documented standards and processes, or even best practices, for funders to perform scientific reviews after the funding commitment. Considering the investment in and standardization of ethical and regulatory reviews, and the prevalence of studies ending without clarity or never finishing, it may be that scientific reviews of trial designs with a focus on informativeness offer the best chance for improved outcomes and return on investments in clinical trials. A maturity model is a helpful tool for knowledge transfer to help grow capabilities in a new area, or for those looking to perform a self-assessment in an existing area. Such a model is offered for scientific design reviews of clinical trial protocols: a valuable and often-neglected governance step for funders or sponsors, among others. This maturity model includes 11 process areas and 5 maturity levels. Each of the 55 process area levels is populated with descriptions on a continuum toward an optimal state to improve trial protocols in the area of risk of failure. This tool allows for prescriptive guidance on next investments to improve attributes of post-funding reviews of trials, with a focus on informativeness.

Keywords: clinical trial; informativeness; design review; trial methods maturity model

Assessing quality in global health clinical trials

In addition to industry, hundreds of global health clinical trials (GHCT) are funded annually by private foundations, governments, and consortia. A meaningful number of these trials end without being published or without trustworthy results [1,2,3]. A query of the clinicaltrials.gov registry, in April 2022, found 644 active or pending phase 1, 2, or 3 clinical trials with at least one site in sub-Saharan Africa. Industry—either alone or as leader of a funding group—funded just over half the trials. The US government funded just over 5% of trials. The remaining 40% of trials were funded by private foundations, with some contribution from other governments or organizations. These GHCT had plans to enroll 2,792,554 participants (human research subjects). Before a trial begins, industry routinely performs scientific or methodological reviews on trial protocols to identify flaws in design. There is no direct evidence that most others do. It is imaginable that upward of 50% of GHCT begin with no dedicated scientific design review. This may account for the large difference in informativeness between industry and non-industry clinical trials (CT) found recently [4].

In the lifecycle of a clinical trial, there are two phases prior to the trial's start and participant recruitment. First is a phase when the trial has not procured a funding commitment (pre-funding), and then the second is a post-funding phase. The dominant approach used by government funders to decide if a research study will be funded is peer-review. Peer-review "has become the main mechanism for distributing resources from governments and philanthropic funders" when it comes to health research [5]. While peer-review for pre-funding decisions is well established, it continues to

 doi:10.20944/preprints202304.0147.v1

evolve, and not necessarily in a scientific direction. For example, a large faction of stakeholders believe peer-review ought to change to only assess the investigator, not the proposed project, or include a lottery [5,6]. One systematic review found that, in pre-funding peer-review, comments on research design represented 2%, methodology 4%, and methodological details 5%, respectively, of total comments [7]. During pre-funding, these reviewers also needed to comment on dozens of other factors [7]. This dynamic—along with the sometimes large time gap between pre-funding and trial inception and the design changes therein—makes peer review inadequate for scientific design review.

In the post-funding phase, there are two other types of review that focus on elements outside of trial design. These reviews and related concepts are described in Table 1. The two reviews that happen completely or primarily in post-funding and before participant recruitment begins are regulatory and ethical [8]. The regulatory and ethics review domains are relatively mature and well-developed.

Table 1. Types of reviews for clinical trials.

Review Type	Definition
Pre-funding peer review	Researchers, academics, and scientists—recruited by funders to volunteer their efforts—assess applications from their peers that solicit funding for clinical studies. This evaluation, performed by experts external to the funder, is "used to decide whether studies will be funded." [5,9,10]
Regulatory review	"An investigation of the proposed clinical trial protocol to assess whether the trial will expose subjects to unnecessary risks, qualifications of investigators, commitments to obtain informed consent, intent to both gain approval from an independent review board and comply with drug regulations." [11] Alternately, "assessment and validation of the submitted regulatory documents, assessment of study protocol, scientific evidence to show the product is safe for the study, the risk and benefits of patients' participation in the study; qualification of the study team, commitment to follow regulatory and GCP guidelines and to protect participants, and ensuring ethical clearance has been achieved." (communication, Dr. Beno Yakubu, Nigeria Agency for Food & Drug Administration, April 16, 2022)
Ethical / bioethical review	Joint consideration by expert group members on a clinical trial's ability to follow best practices in seven areas: "respect for subjects, informed consent, independent review, favorable risk-benefit ratio, fair subject selection, social and clinical value, and scientific validity." [12] [The process must assess risk-benefit in areas such as participant consent, confidentiality, data security, minimizing harms to participants; in return for patient risk, the study must conform to scientific principles, and take into account the existing body of evidence, and make a contribution to generalizable knowledge." [13,14,15]
Scientific design review	An evaluation focused on study protocol details that estimates and describes risks of not achieving statistically sound and meaningful results—such a review must evaluate biostatistical methods, question formulation, inclusion and exclusion criteria, commonality of endpoints, site selection, prudence of prevalence and effect assumptions, and more, without focus on generic good clinical practice, ethics, and regulatory topics that will be covered by others. A research unit wrote "[our] protocol review monitoring committee is the scientific review panel, responsible for ensuring the scientific merit and rigor of the protocol, while the IRB ensures that the study is ethical and safe." [16, 17]

Ethical and regulatory reviewers have bits in the definitions of their review purpose that cross into consideration of clinical trial design methods. For example, "some guidance to ethics reviewers make clear that their remit includes ensuring a trial is based on scientific principles." [16] However, "it is clear that scientific assessments are a source of confusion for some ethics committees…ethics committee members revealed that they often had doubts about whether scientific validity is within their purview."[13] Because the main focus of an ethics review is *not* assessing optimal trial methods,

 doi:10.20944/preprints202304.0147.v1

"ethicists entering a review may be concerned about whether they have "the scientific literacy necessary to read and understand a protocol". [13] One African expert described this as part of the confusion with "overlapping mandates and common goals" amongst sponsors, regulators, and ethicists [8]. Regulators and ethicists in low- and middle-income countries are often not trained in the scientific disciplines necessary to evaluate trial design risk—such as biostatistics and pharmacokinetics. Members of institutional review boards (IRBs) seeking to deliver on their primary purpose—delivering an International Council for Harmonisation (ICH) E6, E8, E9, and Good Clinical Practice (GCP) guideline-supported participant protection review—and members of regulatory boards seeking to deliver on safety and participant protection may, justifiably, take only a secondary look at a trial's statistical details. A cursory assessment of methods by an ethics committee may be necessary for them, but it may not be sufficient for funders. Likewise in the regulatory realm: the review of a protocol post-funding will include only targeted scientific assessment, since, for regulators, the focus on safety and similar matters crowds out efforts to identify more optimal approaches in trial design.

This state of affairs leaves an opportunity gap for scientific review of GHCT designs post-funding. These design reviews are either not happening or are not happening systematically. Industry performs scientific design reviews; it may or may not be coincidental that industry-funded trials were more likely to be informative during COVID than those trials funded by others [18]. The cancer clinical research community—influenced by the US government—has adopted the post-funding scientific review of trial designs, but no other research community has followed [19]. An Alzheimer's disease consortium stated "Although there is scant literature on this topic, design flaws that could have been corrected before trial initiation via scientific review are common." [17] Rigorous trial design reviews, especially ones keyed to informativeness as described below, happen sparsely in the world of non-industry funders. Standards do not exist.

Informativeness

Informativeness is a characterization of a clinical trial that indicates the study will achieve its recruitment, statistical power, and other design goals, resulting in credibly answering its research questions. An informative trial "provides robust clinical insight and a solid on-ramp to either the next phase of development, a policy change, a new standard of care, or the decision not to progress further." [20] Conversely, an uninformative trial is "one that provides results that are not of meaningful use for a patient, clinician, researcher, or policy maker."[21] Across a number of stakeholders working to identify design practices associated with uninformativeness, there is consensus on a core set of failures. These include principal investigators (PIs) being unrealistic or overly optimistic in their ability to set and achieve feasible and appropriate sample sizes, failure to embrace trial designs that constitute "solid science", and non-use of evidence-based disease burden and effect rates [18,22,23,24]. "Studies that failed to influence policy change or a confident next step in a go/no-go decision were associated with factors such as lack of use of common endpoints, lack of conservatism in effect estimates, not using biostatistical simulation to derive sample sizes, using unduly restrictive inclusion criteria, and avoiding use of innovative trial designs." [20] Qualities that drive informativeness are almost all identifiable early and are defined during the design phase of the trial. Eleven of Zarin et al's twelve "red flags" for uninformativeness can be identified before a trial begins recruiting [25]. A multi-stakeholder working group of experts led by the Experimental Cancer Medical Centres made recommendations on how to improve clinical trials. Seven of the group's ten consensus recommendations could or must be planned and addressed during the design phase of a trial [26]. Because likelihood of informativeness is cemented from a PI's design work and design choices, post-funding scientific design reviews have high potential to identify risks of uninformative outcomes and suggest fixes before the trial is frozen and cannot be changed.

A maturity model for scientific design reviews of clinical trials

The practice of performing repeated, consistent, high quality, managed scientific design reviews of GHCT requires business processes. A maturity model is "a tool that helps assess the current

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 April 2023 doi:10.20944/preprints202304.0147.v1

effectiveness of a person or group and supports figuring out what capabilities they need to acquire next in order to improve their performance." [27] "Maturity models are recognized as tools for demonstrating the gradual and systematic improvement of an organization's general skills or processes." [28] As an organization desires to implement clinical trial scientific design/methodology reviews, or improve existing reviews, a maturity model can help to improve quality and capacity.

There are a number of flavors of maturity models. A suitable model for presenting a maturity model is the Object Management Group Business Process Maturity Model (BPMM-OMG) [29]. Maturity Levels (ML) are displayed on the Y-axis and are "well-defined evolutionary plateaus toward achieving a mature…process." [30] The ML titles specific to BPMM-ORG and their fixed definitions are shown in Table 2. These levels act as ratings, scores, or grades for key parts of a review process.

Table 2. Maturity Levels (BPMM-ORG) (BPMM-ORG)[1].

Maturity Code	Maturity Levels	Maturity Level Definition
ML5	Innovating	Wherein both proactive and opportunistic improvement actions seek innovations that can close gaps between the organization's current capability & the capability required to achieve its business objectives.
ML4	Predictable	Wherein the capabilities enabled by standard processes are exploited & provided back into work units. Process performance is managed statistically through the workflow to understand & control variation so that process outcomes can be predicted from intermediate states.
ML3	Standardized	Wherein common, standard processes are synthesized from best practices identified in the work groups & tailoring guidelines are provided supporting different business needs. Standard processes provide an economy of scale & a foundation for learning from common measures & experience.
ML2	Managed	Wherein management stabilizes the work within local work units to ensure that it can be performed in a repeatable way that satisfies the workgroup's primary commitments. However, work units performing similar tasks may use different procedures.
ML1	Initial	Wherein business processes are performed in inconsistent sometimes ad hoc ways with results that are difficult to predict.

The X-axes in maturity models represent groups of capabilities, often called Process Areas (PA). PA's are a "cluster of related activities that achieve a set of goals for a capability." [30] In order to create a usable maturity model, users must carefully select the range of capacity and efforts—the cluster of related activities: in order to evaluate a scientific design review practice, the process areas must be identified and organized. These PA descriptions for scientific design reviews are shown in Table 3. In each 'cell', or capability cluster at a particular level of maturity, the contents include examples of mastery at that level.

Table 3. Process Areas for performing scientific design reviews of clinical trials.

Process Area	Process Area Definition
Informativeness-centric	A scientific design review where the main focus is on identifying and reducing risks that the trial will end without definitively answering its research question. An informativeness-centric review leaves as secondary any design concerns tied exclusively to regulatory, bioethics, and clinical operations topics. Focus is on evidence-based drivers of informativeness, such as sample size methods, use of local up-to-date epidemiological data

[1] Maturity Level definitions here are taken directly from BPMM-ORG.

 doi:10.20944/preprints202304.0147.v1

	as input variables, conservative effect estimates, use of biostatistical simulation, and use of common endpoints.
Breadth of review expertise	Every trial has a variety of attributes that might make it distinctive. These attributes may appear across a range of trial elements, such as the intervention, stage of the disease, trial phase, trial site(s), or design characteristics. Breadth of review expertise means the expert review panel includes, for most or all unique attributes, a reviewer who has implemented, provided oversight for, designed, or critiqued that attribute in the past. This represents how complete the application of reviewer expertise to all details of a trial can be. This is often correlated with more, rather than fewer, reviewer individuals on a panel.
Depth of reviewer expertise	Depth of expertise means the review panel includes, for key attributes of a trial and its design, reviewers who have designed and implemented, participated in, or provided oversight for related trials. The reviewer is known to others in the field as being a well-known or famous resource or author on intricacies, advanced methods, or the corpus of work in a specific topic; typically this requires decades of experience.
Iterative	There are multiple rounds of analysis, edits, and collaboration in the review. Each expert sub-panel or working group iterates its findings and recommendations. Sub-panels consolidate and submit their review to a higher-level panel, which iterates with the sub-panels and within itself. The higher-level panel iterates the review with the PI. The iterations ensure each critique and recommendation has been refined, prioritized, and understood.
Information-enhanced	There is a wide variety of information beyond the protocol that could indicate the risk level and riskiest attributes of a trial's design. This information, if curated, and put in the hands of reviewers, makes for a richer review. An information-enhanced review means one where the protocol is accompanied by information incremental to the protocol requested by reviewers and sourced from PI or internally, that provides risk insights to the reviewer.
Solution-oriented	Solution-oriented means reviews ought to focus on solutions to multiple stakeholders'—but especially PIs'—challenges as well as the challenges inherent in design attributes. Solutions offered ought to be specific, timely, feasible, and informativeness-forward. Solutions could include links to other experts, additional funding such as trial planning grants, data, or other resources.
Software-enabled	Software-enabled scientific design review means all relevant portions of the process that can be reliably enhanced with technology would be. This ranges from basic mechanics such as scheduling, communication, and secure document sharing, all the way to the use of artificial intelligence for prediction and data mining. Software could be used to support other process areas, such as measuring time spans, or for scouring registries, databases, historical protocols, and publications toward information enhancement.
Collaborative	A collaborative review process is one that is increasingly communicative within and across stakeholder groups. This communication and collaboration could be flexible enough to adjust to changes in context. Collaborations could range from enabling quick scheduling & correspondence to partnering more deeply in-person, telephonically, or with other real-time engagement. Reviewers speaking to PIs about protocol review findings and recommendations is a crux of collaboration.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 April 2023 doi:10.20944/preprints202304.0147.v1

Rich in data & analytics	A review program rich in data & analytics is one that collects, cleans, curates, and enriches information about all parts of reviews, and uses analysis and visualization to communicate more richly with stakeholders, answer questions, aid in actionable decision-making, as well alerting to trends and finding opportunities.
Reliability and quality	The platform and approaches to delivering reviews perform their intended function. Reviews and the mechanics of delivering them are dependable. The team and platform sustains a level of quality over time. There is an increasingly lower number of fails, and defined approaches to fix failed reviews. Stakeholders perceive quality and value in reviews. There is a consistency of delivery over time; costs are maintainable.
Time appropriate	The review approach takes into account time sensitivities of disease urgency and current context, and needs of the funders, PIs, and other stakeholders. Each segment in the review process may consider, relative to different facets of time, the attributes of sustainability, routineness, elasticity, rigidity, or fragility. Organizations are precise around trade-offs related to timing and deliberate in their application of time aids and boundaries.

Once a maturity model variant is selected and the topic-specific PAs are populated, users can plot the maturity levels on the Y-axis and the PA on the X-axis. In the case of an 11-item PA list and 5-level maturity ladder, the 55-cell tool is difficult to represent in one image. As a result, the following tables show the PAs individually, with descriptions of the maturity of process aspects at each level.

Table 5. Process Area 1, Informativeness-centric An informative trial includes a hypothesis that addresses an important and unresolved scientific, medical, or policy question; is designed to provide meaningful evidence related to this question; must have a realistic plan for recruiting sufficient participants; must be conducted and analyzed in a scientifically valid manner; and reports methods and results accurately, completely, and promptly. [21] An alternate definition is that an informative trial is designed to have the best chance to complete on time, answer its research questions definitively, and effect policy change or a regulatory process, through special commitment to a) siting the trial based on epidemiology and impact rather than convenience, b) completing a statistical analysis plan concurrently with the trial protocol, c) using accepted endpoints and conservative effect and prevalence/incidence estimates, and d) utilizing contemporary techniques, such as statistical simulation, innovative trial designs, and software to monitor recruitment.

Maturity Code	Maturity Levels	Capabilities, Informativeness-centric review
ML5	Innovating	Review organization is efficiently applying artificial intelligence, digital innovation, and novel data to each review to surface informativeness risk. New research, based on studies that end informatively, is being performed to identify true drivers of informativeness. Innovations are invested in that make drivers to informativeness more clearly identified by reviewers. Improvements to continuing education for reviewers occur. Uncovering new drivers to uninformativeness happens through convenings of peers, talks from visionaries and researchers, and other engagement. Methods of innovative, effective continuing education for expert reviewers who are not pure trial methodologists are explored. Because of the pace of technology improvement and new inventions in the space, not all innovations can be adequately captured and described here.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 April 2023 doi:10.20944/preprints202304.0147.v1

ML4	Predictable	Informativeness information in the public domain specific to the design to be reviewed is collected and provided to reviewers along with the review materials: review-specific 'benchmarks' of similar trials crafted just in time. Statistics on historical coverage of informativeness review variables are collected and provided to reviewers to ensure more complete reviews. As each review progresses, at interim points early indicators of trending areas of concern amongst reviewers, and/or other-generated risk scores are provided to all reviewers.
ML3	Standardized	The review organization selects the most appropriate definition of informativeness. A list of common drivers of uninformativeness is made, referred to, and used to educate and orient reviewers. This list is sourced empirically and reinforced with clinical study exemplars. The primacy of informativeness over good clinical practice, regulatory, ethical, or other types of recommendations is mentioned often and debated. The large majority of review recommendations tie to informativeness rather than GCP, regulatory, reducing bias, or other areas of interest. All stakeholders beyond reviewers are prompted to understand informativeness and its drivers.
ML2	Managed	Across reviewers, there is a mix of education level on uninformativeness and what causes it. There is variation in review output based on varying levels of trial experience and other variables. Some reviews have documented recommendations with rich meta-information, others have limited context, others have no documented output other than a discussion. Informativeness is specified as the most important quality of a trial, and the purpose of the review only sometimes focuses on informativeness. Sometimes reviewers ignore the informativeness focus and default to their own focus preferences (regulatory, ethics, clinops, equity, other). Reviews may miss obvious PI shortcuts that increase risk of uninformativeness.
ML1	Initial	Focus on informativeness varies often. Many reviewers believe the word informativeness refers to the traditional definition, rather than being a term of art tied to trial design rigor. Review comments optimize the study design for items unrelated to informativeness, such as safety, ethics, or clinical operations (clinops). There is no clear, written statement that the review is about informativeness. There is no discussion or education about the fact that a trial could fail to be informative... PIs can avoid being reviewed if desired. Presence of specific reviewer roles decides whether certain informativeness levers are reviewed.

Table 6. Process Area 2, Breadth of review expertise.

Maturity Code	Maturity Levels	Capabilities, Breadth of review expertise
ML5	Innovating	Unique approaches are developed to gain expertise when an expert cannot be sourced. For example, digital mining of prior art and summation through natural language processing, or pre-automated blitz surveys of a relevant PI community at large, could assemble expertise. Artificial-intelligence-powered recruiting techniques could find new experts. Automated identification of best practices

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 April 2023 doi:10.20944/preprints202304.0147.v1

		in a range of expert areas from mining protocols, trials, and past reviews could be applied in lieu of a human. Because of the pace of technology improvement and new inventions in the space, not all innovations can be adequately captured and described here.
ML4	Predictable	Experience with spikes in review demand informs when to build in redundancy of SMEs, curators, and managers, to have multiple review teams available with full breadth of review expertise. Insight into the ability of doing run-time recruiting of expertise beyond the breadth of the current portfolio of experts is known. Approaches exist for how to find, stage, and staff expertise for rare or complex interventions, diseases, and trial types when they map to specific review timelines, and requests are all predictable and routinized. Information on trial sites most commonly reviewed is used to identify the ideal geographic background of experts from low-resource settings. Data exists to present the elasticity of value of expertise breadth, toward understanding how many experts in diverse but applicable niches is the right amount per review.
ML3	Standardized	A core set of minimum expertise coverage across expert roles is set and is attained consistently, or the review is not performed. The 'mix' of experts on a single review team and the relative ratio of headcount is established. Some policy around run-time requests from stakeholders for specific expertise is set. Policies on funding expert reviewers, ranges of rates, and levels of effort and definition of scope within segments of a review are established. Business rules exist to know when a review team will include an expert for the trial site geographies; an expert specializing in gender equity or other equity considerations relevant to the trial; specialty experts; an implementation research expert; an expert tied to the specific intervention type, if it is unique; a disease-specific expert; and the number of experts in core disciplines such as biostatistics and pharmacokinetics/pharmacodynamics.
ML2	Managed	Specific, repeated areas of needed expertise, or “expert roles”, have been identified, such as biostatisticians, epidemiologists, or pharmaco-kineticists. Attempts are made, but not always achieved, to deliver minimum “expertise coverage” within and across reviews. Experts have differences in agendas, amount of time to give to the review, number of prior reviews, or processes. Specialized or “custom-to-trial” expert needs identified as aspirational. There is consensus for what the minimum, core, required set of breadth of topical expertise is in a review panel.
ML1	Initial	Most reviews do not have reviewers with experience in special or secondary trial elements such as disease, intervention type, site geography, and participant attributes. Representation of expertise for a particular trial element, or “expertise coverage” has sparsity within and across reviews and is not predictable. For example, there may be an odd time when a review panel includes no expert for biostatistics or for pharmacokinetics. There is no consensus on what the core, required set of breadth of expertise in reviewers is.

Table 7. Process Area 3, Depth of reviewer expertise Terms included: Depth: “In subject matter expertise, the ever-increasing ability to identify, characterize, and extemporize on a myriad number of discrete facts, concepts and ideas related to a topic area, using or creating interconnected networks

 doi:10.20944/preprints202304.0147.v1

of meaning and significance in the process." [31] Here, depth of reviewer expertise relates to the skill of any single reviewer. An expert reviewer with a very high depth of expertise would be considered a known figure in their field, a recognized author, or sought after speaker in that niche. Breadth of review expertise, on the other hand, refers to the range covered by all reviewers on a review panel. Topgrading: "Topgrading simply means proactively seeking out and employing the most talented people available, redeploying those of lesser ability…employing only A players [top talent]." [32].

Maturity Code	Maturity Levels	Capabilities, Depth of reviewer expertise
ML5	Innovating	Automated or innovative methods for "topgrading" occur to continue to increase the level of expertise among those already with expert status. Artificial intelligence may be used to create "digital twins" of deep experts, and tested for superior review recommendations. Data science may be used to predict timelines of turnover among expert reviewers toward ensuring no gaps in coverage. Prediction of future new study types and domains identify new types of experts to be recruited in advance. Because of the pace of technology improvement and new inventions in the space, not all innovations can be adequately captured and described here.
ML4	Predictable	Insights from historical review recommendations are used to predict optimal remedial or *de novo* education or experience to further deepen expertise. Analysis of review recommendations over time may identify reviewers who are outliers within expertise area-study type combinations, toward quality improvement. A 'bench' of future review capacity may be built, consisting of high-quality experts pre-trained in the review process, in order to not have a decline in review quality in situations of existing reviewer turnover. Unique reviewers may, during or between reviews, create original content, resources, data, or tools that reinforce their review feedback, and provide this content to the review team.
ML3	Standardized	A minimum level of depth of expertise is set and practiced. Financial minimums, boundaries, resources, and methods are set to ensure the review process can afford deep experts. Multiple defined sources, registries or catalogs of experts are established. Standard recruiting and onboarding content or tools may be developed to ensure a high quality of new expert reviewers. Documentation exists to describe to reviewers what is expected of them, toward their understanding that particular parts of their expertise may or may not be needed, to further ensure they are spending their review bandwidth in their area of expertise.
ML2	Managed	Approaches for recruiting expert reviewers are practiced and improved. Breadth of expertise in a review team is sacrificed if needed to ensure minimum depth of expertise standards. An understanding of current and future financial budgets for procuring depth of expertise, if necessary, are understood. At times, adding multiple experts in a domain who review in serial, the second reviewer using the recommendations of the first, may be used as a way to increase depth of expertise. Depth is not uniform across reviews.
ML1	Initial	Based on the definition of 'expert' herein, some members of some review teams are experts, and some are not. In order to achieve breadth of expertise, some sacrifice in depth of expertise is made.

 doi:10.20944/preprints202304.0147.v1

	Financial and time investments needed to "source and hire" deep experts may not be affordable yet.

Table 8. Process Area 4, Iterative Terms included: Tranche: One of the parts into which an investment is divided. In this case, the "investment" of reviewer panelists' time is not done in a single sitting, with all reviewers critiquing simultaneously. Instead, a smaller review panel critiques together, and their feedback is fed to a second reviewer panel who does their critique (the second "tranche"). Multiple tranches of review, if managed well, may lead to the creation of a wider set of recommendation topics as well as enabling deeper considerations on initially-identified recommendations. Reviewer panel: An assembled team of experts tasked to read and critique a clinical trial protocol and associated documents on behalf of a funder (where the team is not part of the PI group, nor has a conflict of interest), who makes their comments and recommendations around the same time, and often represents a wide range of technical competencies rather than all being expert in one discipline.

Maturity Code	Maturity Levels	Capabilities, Iterative review
ML5	Innovating	Artificial intelligence could analyze historical data to identify and cue optimal orders of communication, duration and type of communication, rounds of communication within and across stakeholder groups, to lead to feedback eliciting more protocol changes. Particular order of iterative review across reviewer roles could be optimized with higher order processing. For example, it may be that implementation research expertise must be used to identify possible study change recommendations first, then a biostatistician needs to review those recommendations to identify strength or risk before the whole recommendation is delivered, as opposed to biostatistical review first. Because of the pace of technology improvement and new inventions in the space, not all innovations can be adequately captured and described here.
ML4	Predictable	Processes for inter-reviewer debate are streamlined and tuned based on prior experiential data. Management begins to predict optimal composition of later tranches/reviewer panels based on initial tranche/reviewer panel composition, especially by certain trial context. Identification of opportunities to add an early tranche of concept review during early, pre-protocol, conversational stages is sought.
ML3	Standardized	The cadence across review teams for communication, clarification, document, and feedback delivery is well-established and maintained. There is standardization of role types of reviewers in the first, second, and further tranches. Schedules and processes for members of each tranche are standardized and communicated. Consistent achievement of no duplication of expertise in any one tranche, with minimal sparsity, is attained, for powering efficient iteration. For example, a human 'curator' who convenes meetings with each tranche of review, and otherwise removes duplicate comments and manages information flow (i.e., a "compiler") may emerge. A culture of respect and equality is sought between different tranches of reviewers or reviewer subgroups.
ML2	Managed	More than one tranche of review panel exists, deliberately. The internal team managing reviews encourages reviewers to consider and discuss recommendations or commentary from other reviewers within and across tranches, or have meetings to do such

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 April 2023 doi:10.20944/preprints202304.0147.v1

		consideration directly. Availability of documents and information does not follow a standard timeline. There are no consistent means for the second tranche of reviewers to gain clarification on commentary from initial review tranche. Processes are not worked out on final decision-making power across reviewers.
ML1	Initial	An individual reviewer's recommendations are often not visible to other reviewers. No iterative tranches of reviews with regard to communication with the PI: review comments go directly to PI from reviewer panel. Occasionally a cycling of comments will occur within a single tranche of review, and individual feedback is adjusted.

Table 9. Process Area 5, Information-enhanced Terms included: Information: "Here, "information-as-thing", the attributive use of "information" to denote things regarded as informative; varieties include data, text, document, and objects that can be stored and retrieved—as opposed to processes or knowledge." [33] Specifically, that when reviewers are enabled with precise, efficient, rich, and actionable information at the right time, there is a likelihood of a better review. An information-enhanced review is one where reviewers have the benefit of additional, just-in-time, documents that are original and completed both by the PI—in the form of a fixed request for specific information beyond the protocol—and the reviewer organization—in the form of aggregated relevant contextual variables specific to the trial to be reviewed.

Maturity Code	Maturity Levels	Capabilities, Information-enhanced review
ML5	Innovating	External data, both around similar studies and the review trial, are used to predict uninformativeness and made available to reviewers. External information, such as libraries of study protocols, are used to identify opportunities for review recommendations. Information is synthesized, either from past reviews, current recommendations, or other sources to help with individual reviews or an aggregate level of reviews. Data science is used to evaluate novel data, real-world data, and other information to identify if there are reliable associations with variables of interest, as well as associations between review feedback and informative outcomes in reviewed trials. Because of the pace of technology improvement and new inventions in the space, not all innovations can be adequately captured and described here.
ML4	Predictable	Accompanying, enriching data, such as completed statistical analysis plans, target product profiles, target policy profiles, participant communication plans, dissemination plans, intermediate PI responses, or additional information requested by reviewers are collected and provided back to reviewers reliably. Protocol changes made are used to predict improved profiles of information compiled. Analytics on the effect or likelihood of uninformativeness associated with missing ancillary data, benchmarks, SAPs, and other information are performed.
ML3	Standardized	Consistent and standard information requests for additional information beyond the protocol by reviewing organization, and delivery by the PI of the same, enable improved review recommendations and faster reviews. The review organization has designed a fixed template of required or requested information to be compiled by the PI just prior to the review, to bring further clarity to certain key aspects of the protocol. Prior reviews identify core

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 April 2023 doi:10.20944/preprints202304.0147.v1

		information of reviewer interest that, when collected from similar, recent trials, becomes a 'benchmark' against which to validate the trial being reviewed; these benchmarks are created for and given to review panels.
ML2	Managed	Trial protocols are often complete and usually arrive on time. Requests are made to the PI to provide new content outside of the protocol to aid reviewers; such requests are not always fulfilled. Financial or other resources may exist to support PI's or reviewers compiling their own review-enriching information, within certain specified bounds. Definitions of minimum requirements of study protocols exist. Alternate types of reviews may be implemented for situations where minimum criteria of information in a review are not met, such as when only a draft study synopsis is available.
ML1	Initial	Profiles of information that reviewers would like to receive along with a protocol are discussed and requested. Reviews sometimes include a complete study protocol, and sometimes rely only on a study synopsis or largely incomplete protocol. Rarely does enhancing or enriching information—either incrementally solicited from the PI at the time of the review, or developed by the funder specific to the review—accompany a protocol.

Table 10. Process Area 6, Solution-oriented. Terms included: Solution: Implementation of a technique or method—enabled by adjustments by or with people, process, information, or investment—to improve capability or delivery in a situation. Here, a solution is a change in some attribute of the trial, trial design, or trial design method that increases the likelihood that the trial will end informatively, in the opinion of the reviewers. While communication may be a solution in itself or part of a solution, communication is handled separately, as Process Area 8, "Collaborative reviews".

Maturity Code	Maturity Levels	Capabilities, Solution-oriented review
ML5	Innovating	Customization or optimization of review process to grantee and study type may occur, toward the review approach most likely to lead to a solution implementation for the review issues found. Opportunistic incremental funding for review recommendation-driven changes is routinized and data-driven. Incentives for grantees to adjust to informativeness recommendations may be tested. Understanding informativeness 'returns' tied to changes in protocols might occur. Optimization of aggressiveness of review recommendation or a mix of trial planning grant investments could be developed. Because of the pace of technology improvement and new inventions in the space, not all innovations can be adequately captured and described here.
ML4	Predictable	Networks of subject matter experts (SME) are made available, and in some cases, funded to aid in solutions. Review recommendation metadata that is consistently tagged is used in analytics to help reviewers improve. If solutions to trial risks of uninformativeness come with financial support, such as trial planning grants, historical outcomes of same may be used to predict improvements in current solution-centric grants. Feasibility, cost, and appropriateness of solutions recommended are matched with trial attributes and become indicative and actionable. Variations in review

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 April 2023 doi:10.20944/preprints202304.0147.v1

		recommended solutions for the same issues identified in reviews across time are identified.
ML3	Standardized	Review feedback delivered in consistent format for every review, in a timeframe that is actionable. Feedback includes prescriptive or action-centric wording, often with an example solution. Types, sizes, and formats of follow-on assistance to grantees tied to feedback begin to take shape and be templatized or packaged. Other PI-focused 'solutions' may be aided, incentivized, or subsidized with trial planning grants or other support. Considerations for variation across trials or PI teams may be made (e.g., language translation for reviewers or PIs designing solutions & responses).
ML2	Managed	There is some oversight and management of feedback, so it is not cryptic, sparse, or otherwise non-actionable. Feedback recommendations are translatable into solution-centric action by the PI. Review feedback likely to arrive sooner or within a timeframe that corrective action can be taken or attempted. In some cases, feedback can be packaged by PIs into requests for financial assistance to solve the design opportunity. The review recommendation suggestions presence, practicality, and detail may vary based on who is on the review panel, or who is compiling recommendations for the PI.
ML1	Initial	Review feedback is delivered as insights or facts, and only occasionally is delivered in writing with recommended solutions with examples. Feedback is not proofread and edited, and therefore delivered in different 'voices', perhaps without context, further ensuring difficulty for PI in translating feedback to solution concepts. Recommended actions are often missing. No option for iterative dialogue with reviewers is offered, nor is question & answer available.

Table 11. Process Area 7, Software-enabled.

Maturity Code	Maturity Levels	Capabilities, Software-enabled review
ML5	Innovating	Previously unleveraged software types are experimented with to increase review speed, efficiency, and quality. The need for additional language support may be identified and addressed. Additional and perhaps wider groups of stakeholders are addressed who could, through software, be enabled with transparency or other services. Alternate software-driven environments for reviews to take place, such as the metaverse, may be plausible. Because of the pace of technology improvement and new inventions in the space, not all innovations can be adequately captured and described here.
ML4	Predictable	User feedback is collected by all stakeholder types on experiences with standard software tools used, and the results cue actions to improve the experience. The internal review managing team optimizes software used to reach improved level of predictability, speed, efficiency, and quality in all aspects. Service levels are established, including for software outages. Software-specific dictionaries, users guides, common definitions of elements, and metadata are established. Usage information, user complaints, and

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 April 2023 doi:10.20944/preprints202304.0147.v1

		lack of scalability to new capacity building all are collected toward knowing when to move to different software.
ML3	Standardized	An internal collaboration platform is used for those managing reviews and reviewers and for storing review documents. One software is selected for data curation and analytics. A software-centric solution may be rolled out to PIs being reviewed for document management and other communication. Software may be tested and settled on for specific needs of the review management team (e.g., project management or customer relationship management) or, in advanced cases, for the reviewers (e.g., statistical simulation or real-world evidence analysis of epidemiology data). Templates frozen and always used for recording and delivering review feedback and recommendations to PIs. Training and analytic software has been through a selection process, identified, and is in various stages of implementation. Cloud storage is likely in use.
ML2	Managed	Software is used to track review statistics and review recommendations. Different software may be used across review teams simultaneously. Turnover in software tools happens over time as the most feasible tools are settled on. Privacy, confidentiality, and cybersecurity is ensured. Collaboration tools may be crude or sophisticated. Software is usually not cloud-based. Templates experimented with for delivering review feedback and recommendations.
ML1	Initial	No enterprise-class or collaborative software is used in some of these reviews. Widespread, standard, but single user-centric software such as Microsoft Word or Microsoft Excel may be used frequently. Version control or lost data issues can occur. Some reviews include all recommendations tracked and stored historically.

Table 12. Process Area 8, Collaborative. Terms included: Collaboration: "A process where key stakeholders constructively explore different aspects of a problem domain and search for solutions, leading to joint decision making about the potential future of that domain, while using agreed-upon rules and a temporary structure." [34]. There are limits to collaboration; limitations could be legal, organizational, and functional. In fact, one attribute of success in a review system might be an "arm's length" distance or element of independence between reviewers and authors of a protocol. What seems clear is that *outside* the areas where independence and reserve are required, increased collaboration makes the review better.

Maturity Code	**Maturity Levels**	**Capabilities, Collaborative review**
ML5	Innovating	New collaboration technologies and software are piloted for potential addition to the program. Deeper slicing and dicing of recommendation uptake data tied to communication and collaboration information occurs and is acted upon. New types of communication and collaboration are tested. "Joint reviews" with external stakeholder reviewers are enabled. Because of the pace of technology improvement and new inventions in the space, not all innovations can be adequately captured and described here.
ML4	Predictable	The most productive approaches to the review process are identified through data analysis of uptake of review recommendations. Collaboration, where appropriate, begins to be

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 April 2023 doi:10.20944/preprints202304.0147.v1

		customized based on PI attributes. Uptake of change management, communication-centric process improvement or implementation programs happens, such as the AIM Change Management Methodology [35], to aid in successful collaboration. Collection of data on collaboration effectiveness, and use of that data, occurs. As collaboration moderates protocol changes, insights on collaboration are driven into other process areas. Collaboration learnings drive choices in enhancing software, information, and other program areas.
ML3	Standardized	Communication of schedules and milestone activities is proactively delivered in standard ways across all review stakeholders. Human collaboration "curators" or relationship managers facilitate easier engagement amongst parties. Service levels for communication exist. Reviewers see others' comments and have collaboration mechanisms and defined time-based cues to collaborate. Standards for collaboration and communication between reviewers and PIs are established and seen as an important facet of collaboration. Collaboration with non-reviewer, non-PI stakeholders is defined, if any exist. Reviewers speak to PIs after delivering their written recommendations and receiving PI's written responses to achieve maximum possible collaboration.
ML2	Managed	Communication from the internal team managing reviews to reviewers and grantees is ensured, if inconsistent or delayed at times. Communication between stakeholders is routed through the hub of an internal team managing the review process. Episodes of routine, repeating communication are identified, and best practices are identified. Collaboration among reviewers is specified and prompted. Some reviews include reviewer-to-reviewer collaboration and reviewer-to-PI collaboration.
ML1	Initial	Collaboration between the internal team managing the review process and grantees may be sporadic, irregular, and not proactive. Reviewers will generally not be collaborating or even communicating with each other, nor with grantees. Sporadic bursts of collaboration may occur.

Table 13. Process Area 9, Rich in data & analytics. Terms included: Visualization: "Data visualization are based on qualitative or quantitative data and result in an image, graphical depiction, or chart, that is representative of the raw data, readable by viewers and supports exploration, examination, and communication of the data." [36].

Maturity Code	Maturity Levels	Capabilities, review rich in data & analytics
ML5	Innovating	Artificial intelligence is used to make predictions about future results in all areas of the review process. Report libraries are documented with service levels for publish/subscribe and metadata for all user types. Opportunities for bespoke analytics and visualizations are offered to executives with dedicated reporting experts available to deliver same. Fast turnaround is enabled for situations when executive decision-making has need of actionable intelligence. Data sharing with peer partner organizations may occur. Because of the pace of technology improvement and new inventions in the space, not all innovations can be adequately captured and described here.

ML4	Predictable	External or novel data may be added to the internal trial and review data to enrich potential analytics. Information on review mechanics (personnel involved, multiple types of dates during the process, durations, and other operational facts) are used to create actionable insights. Information about trends in recommendations and outliers are provided to decision makers and reviewers. A wider set of stakeholders are, in an automated way, given access to visualizations. Historical data is used to drive improvements regularly (e.g., every quarter).
ML3	Standardized	A wider set of data is collected in a commercial off the shelf software, using common definitions. No documented data dictionary exists. Multiple stakeholders have been able to request specific data fields be captured in the future. A regular yet limited set of visualizations is produced that includes some review metadata. It may take some time to collect, curate, and cleanse data, but these processes are routinized. These analytics are not offered to all parties; they include tables and graphs. Most questions answered relate to 'slicing and dicing' data on the output—often counts—of reviews in aggregate. Answering the question "What happened…" about a review from the distant past is possible without using an oral tradition.
ML2	Managed	A core set of trial elements, trial attributes, review components, and review metadata are all defined as being worthy of capture. Experimentation with data formats and approaches for capture and storage are made. Data sparsity exists. Depth of data varies. Formats are not consistent. Storage is in a mix of locations. Custom visualizations might be created.
ML1	Initial	Some data is captured during and after the review, in different formats. No standard definitions are set for trial attributes[2] or for review metadata[3] that will be stored. Review comments or recommendations are usually recorded as they are delivered. No visualizations are offered to any stakeholder. Questions are not solicited from stakeholders that would require analysis of historical review metadata or attributes.

Table 14. Process Area 10, Reliability and quality[4]. Terms included: Reliability Growth: "Reliability growth is defined as the positive improvement in a reliability metric or parameter of a product (e.g., a system) over a period of time due to changes in the product's design and/or the manufacturing process." [37]. Quality Control: "While reliability is concerned with the performance

[2] Attributes are details, characteristics, features, or tags describing some unique facet or quality. Here, trial attributes might include the trial phase, trial site location(s), cost to the funder, duration, or sample size.

[3] Metadata are "data about data", and in this case refer to statistics or detail about the data collected about reviews. These might be the date review documents were delivered to reviewers, the number of reviewers in each review panel tranche, how many meetings occurred with the PI, or a 'success store' poll result from reviewers.

[4] If an organization lacks resources sufficient to promptly attend to the Reliability and Quality process area, a best practice would be for them to document the process area as "deferred", accompanied by appropriate review and approval. The hazard analysis or "failure mode, effects, and criticality analysis" records for their organization's program would show that the reliability risk is not currently mitigated.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 April 2023 doi:10.20944/preprints202304.0147.v1

of a product over its lifetime, quality control is concerned with performance of a product at a point in time."[38].

Maturity Code	Maturity Levels	Capabilities, Reliability and quality
ML5	Innovating	Methods for increasingly frequent monitoring and measurement of the program 'system' are experimented with, including the use of Bayesian statistical approaches. Attempts could be made to apply external quality improvement through unified "mash-up" reviews with other panels currently reviewing the same trial. Novel applications of expert opinion, the Dempster-Shafer theory, or fuzzy rules could be attempted [39]. Because of the pace of technology improvement and new inventions, not all innovations can be adequately captured and described here.
ML4	Predictable	Processes are documented for handling complaints, so that the information makes its way to program systems, definitions, and mechanics toward increasing reliability growth. Over time, an increasing frequency occurs of assessing data on review quality and review impact, and driving those assessment to review practitioners and into system improvements. Multiple definitions of review impact are documented, tied to single and multiple groups of stakeholders, and assessments and efforts for reliability growth and quality are made.
ML3	Standardized	A level of minimum quality guides delivery of all reviews. Failures are identified quickly, stakeholders are notified, and the situation is remediated. Costs per review occur within tight or predictable bounds, and growth or decreases in costs are predicted. All program members have defined and maintained and continuing minimum training in program-specific knowledge required, such as change management programs and domain knowledge (e.g., informativeness). Routine enhancements for reliability growth or quality that are suggested, tested, and implemented are defined and documented. Defined frequency of data collection, review tests, surveys, and other techniques to measure reliability and quality will be standardized. Reliability is high enough to "clone" a review team: insert a new independent, separate review team and onboard them to perform quality reviews alongside an incumbent team.
ML2	Managed	The attributes of successful, acceptable, and failed reviews are defined, agreed, and documented, but no tests or scoring are applied systematically. Data is collected, perhaps arbitrarily, post-review, from various stakeholders, toward improved information about reliability and quality. Leadership, communication, and consistency aim to routinize the techniques to deliver high-availability and high-maintainability reviews. Minimum and maximum per-review cost amounts are set and tracked. Implementation-wise, scenarios are identified that strain quality, reliability, cost, and other key factors. Change management, data systems, domain knowledge (e.g., informativeness), and other resources that benefit members are applied inconsistently.
ML1	Initial	"The first prototypes produced during the development of a new complex system will contain design, manufacturing and/or engineering deficiencies." [35] The concept of a "quality review" is

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 April 2023 doi:10.20944/preprints202304.0147.v1

	being defined by program members or management, but is not set. Reaching consensus on whether a 'failed' review has occurred is difficult. Techniques have been identified to deliver reviews which include reasonable or high availability of a review to a client, customer, or stakeholder. Techniques for maintaining consistency are envisioned. Techniques for addressing change management challenges are experimented with.

Table 15. Process Area 11, Time appropriate .

Maturity Code	Maturity Levels	Capabilities, Time appropriate
ML5	Innovating	Multiple stakeholders are solicited about the utility of fast reviews, and whether they are appropriate and valuable. Feedback continues to be collected about the verity of a PI or other stakeholder needing a review completed 'immediately'. Information is collected or estimated regarding cost/benefit of faster reviews relative to loss of informativeness due to lower quality recommendations. Faster review times are attempted or assessed by techniques such as pre-staging packaged recommendations by disease-trial phase combination, or by utilizing checklists for expert reviewers to avoid free-flowing protocol interpretation. Methods for activating internal staff and expert reviewers in emergencies that decrease burnout are tested.
ML4	Predictable	Data is collected about review and review stage durations systematically. That data is analyzed statistically to identify causal factors of what drives time overages. If data is available on review quality, it can be an input to the trade-off between duration and quality. Routinized or automated communication is delivered to reviewers, PIs, or other stakeholders when they are approaching or exceeding deadlines. Remediation is applied for reviewers who repeatedly cannot deliver on time. Identification early of PIs who likely will not be able to deliver documents on time or will try and change or delay information-exchange meetings is performed. Steps are taken to proactively address these issues. Management of external or contextual factors that decrease predictability are addressed.
ML3	Standardized	The time extent of the review—the duration—has tight boundaries. Expectations are high that the boundaries of review duration are not exceeded. While causal factors of reviews going overtime are not collected systematically and statistically evaluated, work is applied to gut feel sensibilities of what makes reviews go 'long'. Documented guidelines of who and what is included in each review stage, and how long it can take, are provided in advance per review, as well as in reviewer onboarding. Flexibility, in response to unexpected events, is bounded. Durations are measured in real-time, and communication is high. Repeated patterns of contextual variation may lead to multiple review offerings with different durations. (e.g., expedited and routine).
ML2	Managed	Stages of reviews are identified and defined. Estimates of review duration (e.g., minimum, maximum, median, and mean durations) are considered for the entire review but not the stages within a review span. A single review's scope or definition of output is

		defined and linked to a typical duration; these expectations are shared with PIs and other stakeholders. Attempts are made to not let reviews exceed a certain number of weeks. It is more likely that reviews go long due to exigencies than because of trial-specific variation. Fixed time "policies" are offered regardless of pressures from certain stakeholders, however exceptions are often made to these time-centric policies.
ML1	Initial	Duration of reviews cannot be predicted. Variation across reviews in duration is high or meaningful. Reviews requiring speed invariably have loss of review quality that could include lower quality trial recommendations, less engagement with PIs, less collaboration, or less focus on informativeness.

References

1. Zheutlin AR, Niforatos J, Stulberg E, Sussman J. Research Waste in Randomized Clinical Trials: a Cross-Sectional Analysis. Journal of General Internal Medicine. 2020 Oct;35(10):3105-7.
2. Carlisle B, Kimmelman J, Ramsay T, MacKinnon N. Unsuccessful trial accrual and human subjects protections: An empirical analysis of recently closed trials. *Clin Trials*. 2015;12(1):77–83. doi:10.1177/1740774514558307
3. Williams RJ, Tse T, DiPiazza K, Zarin DA. Terminated Trials in the ClinicalTrials.gov Results Database: Evaluation of Availability of Primary Outcome Data and Reasons for Termination. *PLoS ONE*. 2015;10(5). doi:10.1371/journal.pone.0127242
4. Hutchinson N, Moyer H, Zarin DA, Kimmelman J. The proportion of randomized controlled trials that inform clinical practice. Elife. 2022 Aug 17;11:e79491.
5. Bendiscioli S. The troubles with peer review for allocating research funding: Funders need to experiment with versions of peer review and decision-making. EMBO reports. 2019 Dec 5;20(12):e49472.
6. Guthrie S, Ghiga I, Wooding S. What do we know about grant peer review in the health sciences? F1000Research. 2018;6:1335.
7. Hug SE, Aeschbach M. Criteria for assessing grant applications: A systematic review. Palgrave Communications. 2020 Mar 10;6(1):1-5.
8. Noor RA. Health research oversight in Africa. Acta tropica. 2009 Nov 1;112:S63-70.
9. Recio-Saucedo A, Crane K, Meadmore K, Fackrell K, Church H, Fraser S, Blatch-Jones A. What works for peer review and decision-making in research funding: a realist synthesis. Research integrity and peer review. 2022 Dec;7(1):1-28.
10. Turner S, Bull A, Chinnery F, Hinks J, Mcardle N, Moran R, Payne H, Guegan EW, Worswick L, Wyatt JC. Evaluation of stakeholder views on peer review of NIHR applications for funding: a qualitative study. BMJ open. 2018 Dec 1;8(12):e022548.
11. Investigational New Drug (IND) Application. United States Food and Drug Administration website. Last reviewed February 24, 2021. Accessed April 15, 2022. https://www.fda.gov/drugs/types-applications/investigational-new-drug-ind-application
12. "Ethics in Clinical Research". National Institutes of Health Clinical Center website. Updated October 21, 2021. Accessed January 12, 2023. https://clinicalcenter.nih.gov/recruit/ethics.html
13. Binik A, Hey SP. A framework for assessing scientific merit in ethical review of clinical research. Ethics & human research. 2019 Mar;41(2):2-13.
14. Emanuel EJ, Wendler D, Grady C. What makes clinical research ethical? Jama. 2000 May 24;283(20):2701-11.
15. Mooney-Somers J, Olsen A. Ethical review and qualitative research competence: Guidance for reviewers and applicants. Research Ethics. 2017 Jul;13(3-4):128-38.
16. Williams E, Brown TJ, Griffith P, Rahimi A, Oilepo R, Hammers H, et al. Improving the time to activation of new clinical trials at a National Cancer Institute–designated comprehensive cancer center. JCO Oncology Practice. 2020 Apr;16(4):e324-32.
17. Knopman D, Alford E, Tate K, Long M, Khachaturian AS. Patients come from populations and populations contain patients. A two-stage scientific and ethics review: The next adaptation for single institutional review boards. Alzheimer's & Dementia. 2017 Aug 1;13(8):940-6.
18. Hutchinson N, Klas K, Carlisle BG, Kimmelman J, Waligora M. How Informative were Early SARS-CoV-2 Treatment and Prevention Trials? A longitudinal cohort analysis of trials registered on clinicaltrials. gov. Plos one. 2022 Jan 21;17(1):e0262114.

19. Ning N, Yan J, Dietrich MF, Xie XJ, Gerber DE. Institutional scientific review of cancer clinical research protocols: a unique requirement that affects activation timelines. Journal of oncology practice. 2017 Dec;13(12):e982-91.
20. Hartman D, Heaton P, Cammack N, Hudson I, Dolley S, Netsi E, Norman T, Mundel T. Clinical trials in the pandemic age: What is fit for purpose? Gates Open Research. 2020;4.
21. Zarin DA, Goodman SN, Kimmelman J. Harms from uninformative clinical trials. Jama. 2019 Sep 3;322(9):813-4.
22. Abrams D, Montesi SB, Moore SK, Manson DK, Klipper KM, Case MA, Brodie D, Beitler JR. Powering bias and clinically important treatment effects in randomized trials of critical illness. Critical care medicine. 2020 Dec 1;48(12):1710-9.
23. Benjamin DM, Hey SP, MacPherson A, Hachem Y, Smith KS, Zhang SX, Wong S, Dolter S, Mandel DR, Kimmelman J. Principal investigators over-optimistically forecast scientific and operational outcomes for clinical trials. PloS one. 2022 Feb 8;17(2):e0262862.
24. Rosala-Hallas A, Bhangu A, Blazeby J, Bowman L, Clarke M, Lang T, Nasser M, Siegfried N, Soares-Weiser K, Sydes MR, Wang D. Global health trials methodological research agenda: results from a priority setting exercise. Trials. 2018 Dec;19(1):1-8.
25. Zarin DA, Goodman SN, Kimmelman J. eTable: Conditions for Trial Uninformativeness. Harms from uninformative clinical trials. Jama. 2019 Sep 3;322(9):813-4.
26. Blagden SP, Billingham L, Brown LC, Buckland SW, Cooper AM, Ellis S, Fisher W, Hughes H, Keatley DA, Maignen FM, Morozov A. Effective delivery of Complex Innovative Design (CID) cancer trials—A consensus statement. British Journal of Cancer. 2020 Feb;122(4):473-82.
27. Fowler M. Maturity Model. Martinfowler.com website. August 24, 2014. Accessed July 25, 2022. https://martinfowler.com/bliki/MaturityModel.html
28. Santos-Neto JB, Costa AP. Enterprise maturity models: a systematic literature review. Enterprise Information Systems. 2019 May 28;13(5):719-69.
29. OMG Standards Development Organization. Object Management Group website. Accessed April 4, 2022. https://www.omg.org/
30. Paulk MC, Curtis B, Chrissis MB, Weber CV. Capability maturity model, version 1.1. IEEE software. 1993 Jul;10(4):18-27.
31. Wineburg S. Beyond "breadth and depth": Subject matter knowledge and assessment. Theory into practice. 1997 Sep 1;36(4):255-61.
32. Smart BD, Smart GH. Topgrading the organization. Directors & Boards. 1997 Mar 22;21:22-9.
33. Buckland MK. Information as thing. Journal of the American Society for information science. 1991 Jun;42(5):351-60.
34. Wood DJ, Gray B. Toward a comprehensive theory of collaboration. The Journal of applied behavioral science. 1991 Jun;27(2):139-62.
35. The AIM Change Management Methodology. Implementation Management Associates website. Accessed April 10, 2022. https://www.imaworldwide.com/aim-change-management-methodology
36. Azzam T, Evergreen S, Germuth AA, Kistler SJ. Data visualization and evaluation. New Directions for Evaluation. 2013 Sep;2013(139):7-32.
37. RGA Overview. ReliaWiki website. Last updated December 11, 2015. Accessed May 14, 2022. https://www.reliawiki.com/index.php/RGA_Overview
38. Introduction to Life Data Analysis. ReliaWiki website. Last updated December 16, 2015. Accessed May 14, 2022. https://www.reliawiki.com/index.php/Introduction_to_Life_Data_Analysis
39. Gupta G, Ghasemian H, Janvekar AA. A novel failure mode effect and criticality analysis (FMECA) using fuzzy rule-based method: A case study of industrial centrifugal pump. Engineering Failure Analysis. 2021 May 1;123:105305.