Highly Efficient Clinical Trials

White Paper

for principal investigators and trial implementers

Global health research opportunity

The Bill & Melinda Gates Foundation (BMGF) has funded many dozens of clinical trials and clinical studies. The investigators we have funded and their teams that implement these trials deserve tremendous recognition for the challenges they have faced. We have had amazing results. However, like any research funder, many of our studies end with results that either require further research or are not generalizable. Some of the trials we fund inevitably make the same mistakes others have made in the past. Unsuccessful research—that can neither affirm an effect or no effect—also occur in studies that evaluate behavior, compliance with treatment, food, vitamins, and other interventions. These weaknesses are applicable to clinical research across global health conditions, environments, as well as in crisis settings. These problems are not due a minority of sub-par researchers. We believe some of these weaknesses are due to use of traditional study designs, lack of collaboration and foresight, and behaviors learned from a history of working in less funded settings. Any funder will want make their research funding result in more health gains faster.

The largest opportunity may not be training more trialists and research centers to double or triple the number of studies to run concurrently. The biggest opportunity may be using modern designs, going beyond traditional collaborations based on social relationships, and using well-tested data and technology that can make research much more efficient and solve health problems now.

What are ‘highly-efficient clinical trials’?

Our experience has shown the best clinical research a) ends with insights as quickly as is scientifically sound, b) occurs with multiple treatment arms evaluated and across multiple geographies, c) has generalizable and actionable results, and d) engages the local community toward a non-disruptive and ideally continuous presence. To this end, we believe highly efficient clinical trials are designed with precise knowledge of burden of disease; are placed where burden of disease is adequate to answer the question; use clinical trial simulation; are open to adaptation in design; employ interim evaluations; include direct collaboration and communication with

researchers outside of a PI's social network; include frequent communication with participant communities; and leave a positive footprint when a trial ends. We classify trials that use these approaches as 'highly efficient clinical trials' (HECT). Research investments do not always have these features. We are applying time and funding to educate our program officers to selectively fund studies with these features. BMGF's ideal will be to fund clinical trials with each of these attributes present.

What approaches contribute to highly efficient clinical trials?

We have identified six approaches to global health clinical research that have shown success in increasing efficiency. Some of these approaches tie back to simple fundamentals of doing proper clinical research. Others used more advanced trial designs or data science. All of them are informed by years of trial experience, multiple exemplars across a variety of geographies and pathologies, obvious examples of failures when absent, and common sense. HECT approaches include:

1. Know disease burden and epidemiology to optimally answer the question– placing trials where burden of disease is adequate to answer the research question makes for more meaningful results
2. Use clinical trial simulations – computer simulations will optimize study design and decrease risks of failure
3. Be open to adaptations – use of platform designs or adaptive designs will enable trials to get to an answer sooner and are more compassionate
4. Employ interim evaluations – use of stopping points for blinded data review enables pre-planned adjustments during the study
5. Collaborate and communicate – before, during and after, communication with peers, especially those outside of a PI's research network, decreases risks of repeating mistakes
6. Leave a footprint – implement in a way that sustains resources, is less disruptive to employment of local trial workers, and includes engagement with the community before, during and after a trial.

Know disease burden and epidemiology to optimally answer the question

Clinical research should occur in the geographies where disease burden is sufficient to allow adequate power for a clinical study. Careful selection of multiple clinical trial sites per trial, with the appropriate patient population, will allow for more generalizable results. Traditional feasibility collected prior to study initiation relies on investigators' and their study teams' best estimates and often do not have a robust verification mechanism. The quality of assessment and planning for disease burden in a

trial impacts the efficiency of the trial. Generally, a higher prevalence of the disease or condition will mean participants will be added more quickly.

Placing a trial in a site with less need may result in findings, treatments or interventions that do not 'scale' or implement well to the geographies with the greatest need. If the research includes sub-populations, then new investigation into finding the highest concentration of those subjects may be indicated. Understanding that existing data on burden of disease is not recent, frequent, granular or even direct (it may be calculated algorithmically) means investigators may need to develop and use novel approaches or original investigation to identify where to put a trial. Additionally, understanding the epidemiological choices of prior or current trials may encourage using common endpoints and definitions.

Examples are included here, showing high, moderate and low or no fidelity to the concept of knowing the disease burden of trial sites and epidemiology in question:

> *High*: Trial sites are situated in the countries or counties/regions/districts with the a high of the disease. Burden of disease estimates are based on up-to-date prevalence data sourced independently and locally. Sites are evaluated against risks their unique characteristics may not provide adequate burden to complete the study. The trial investigators collaborate with local implementers to match past history of pathology, national or regional epidemiological trends and other factors. The planned trial makes use of a short run-in stage to confirm the burden of disease. Endpoints and other definitions deliberately match standards or commonly accepted endpoints, enabling post-study research comparisons.
>
> *Medium*: The trial investigators make some effort to find up-to-date burden of disease data at the country or site level, even if it requires additional surveys or effort. The trial takes place in a country with adequate burden of disease, without making use of local or non-published knowledge. Epidemiological design choices are independent.
>
> *Low/ None*: Trial is run only where the principal investigator's (PI's) trusted implementer exists, regardless of disease burden of that location vs. other more optimal sites. Use publicly-available, non-recent estimates of disease burden and independent epidemiology choices.

Use clinical trial simulations

Clinical trial simulations have been in use for many years. This modelling uses data science (traditionally Monte Carlo simulation) in the design phase of a trial. Clinical trial simulations aim to study what the outcome data is likely to look like from the planned clinical trial. They are used to establish otherwise unknown probabilities of success or failure across a range of 'worse' and 'better' case scenarios. Multiple candidate designs are tested under these conditions with simulations, thereby aiding

investigators to pick the design that is most likely to succeed. Many decisions are made in the design phase for trial operating characteristics. Use of simulation better informs these decisions. Simulations particularly reveal the likely impact of getting the assumptions wrong in the planning stage as well as the impact of employing pre-planned adaptations during the trial. For the latter, if interim evaluations are planned, simulations will show which actions informed by interim results are most likely to optimize the trial's likelihood of success.

Examples are included here, showing high, moderate and low or no fidelity to the use of clinical trial simulations before finalizing the study design:

> *High*: Experienced simulation team works with trial planners to cover a sufficiently wide range of plausible 'worse' and 'better' case scenarios and testing several candidate designs. Multiple rounds of simulations are run based on discussion and feedback from multiple sources. The simulations include both cost and time to support a grant request. Grantee includes cost estimates for simulation work to be called out in grant requests. Includes funding for PI's own staff to shadow, train and learn from the process to build internal capability for the future. Investigators are aware of commercial off-the-shelf and open source simulation tools.
>
> *Medium*: A moderate number of simulations are conducted, exploring only a few scenarios and candidate designs. Simulations have gone through no or few iterations and are mostly performed as a supplement to the original sample size calculations. Uses internal staff for simulations, but uses external experts or firms to evaluate your work.
>
> *Low/ None*: No use of simulations in the trial design.

Be open to adaptations

Some of the most sophisticated designs in use today are found in platform trials and adaptive trials. Today, these are more prevalent among pharmaceutical industry-funded trials. Now, adaptive designs and platform trials are being seen in more and more global health studies. We believe platform and adaptive designs clearly meet the criteria of leading edge designs meeting mature standards. Adjusting certain aspects of a trial based on accumulating evidence mid-trial leads to stronger inferences early on and decreased likelihood of the need to perform additional trials. If many questions can be answered in a single trial, there is a higher return for a fixed research spend. Increasingly, global funders desire the chance to fund researchers who can deliver more answers per investment.

Examples are included here, showing high, moderate and low or no fidelity to the use of study designs that allow for adaptations during the trial:

High: Uses a platform trial design or adaptive trial design. Uses a design that stops the trial for superiority or futility. Uses interim adjustments to adjust the participant groups toward ones with more risk or responsiveness. Add more participants to treatment arms that show superiority. Uses experienced partners, contract research organizations or consultants to ensure good clinical practice and design sophistication. Considers adaptations that originate from alternate sites in a multi-country or multi-site study.

Medium: Uses a platform or adaptive trial design that includes only one type of adjustment, e.g. a dosage, sample size, more targeted patient cohort, or early stopping.

Low/ None: Does not include adaptations in the study design.

Employ interim evaluations

One of the most effective ways to make research faster is to use pre-planned interim evaluations and make use of a Data Monitoring Committees (DMC). Interim analyses accompanied by a qualified DMC enable controlled changes mid-trial. Trials with interim evaluations do not need to be 'adaptive' trials. While the pharmaceutical industry is often interested in using interim evaluation to stop trials for futility or superiority, BMGF is interested funding research that can continue the study by adding more treatment arms over time. Trials with interim evaluations are well-known to many regulators, have regulatory guidance available, and have been used in virtually all pathologies—there is little novelty left in trial with interim evaluation.

Examples are included here, showing high, moderate and low or no fidelity to the use of interim evaluations:

High: Uses interim evaluations. Employs design options for adjusting the sample size and allowing for new treatments to take the place of inferior ones will ensure that answers are reached as quickly as possible and that valuable resources are not wasted. The interim analysis plan should include a plan for how to invest grant funds to provide the institutional review board(s) supporting materials that help them understand the rationale for the proposed interim evaluation strategy, as well as consulting help to design the trial and provide support during implementation. Uses an engaged DMC.

Medium: Employs proper approaches for interim evaluations. Uses stopping points for only a single adaptation. May use a DMC that is not able to meet frequently. Works iteratively with institutional review boards to gain a consensus understanding.

Low/ None: Uses a data monitoring committee for needs of the study unrelated to interim evaluations. Does not use interim evaluations or stopping points.

Collaborate and communicate

Collaboration is used here in a specific way. Collaboration in the context of highly efficient research means partnering in ways that may be uncomfortable. The traditional approach for a PI may include working with close colleagues or a few personally trusted institutions in one country or site, using the lowest cost partners, working in the easiest geographies with an abundance of resources, and starting a trial without a lot of communication in the region about previous studies. Ideally, collaboration would include spreading research funding across multiple countries, with significant communication in advance of the trial starting, to understand previous unpublished results, best practices and prior fails. While collaboration and communication is key during and at the end of a trial, it is most impactful in designing a trial that is more well-informed via on-the-ground knowledge and spreads the research.

Examples are included here, showing high, moderate and low or no fidelity to practicing collaboration before, during, and after the study:

> *High*: Communicates with 'previously-unknown-to-you' local or regional experts in the disease or condition, especially for endpoints and outcome measures. Looks for unpublished findings from previous studies that failed, in order to identify unforeseen mistakes the trial might include. Researches and ensures the trial does not unknowingly replicating another trial without a clear reason. Assesses if the clinical question is overlapping with those of previous research in a way that will make go-forward investments less likely or make generalizability less possible. Continues to communicate through the trial to increase on-the-ground adaptations. Communicates after the trial, to share the unique challenges, and with ideas of how they might be addressed better by others in the future. Makes trial data findable, accessible, interoperable and reusable (FAIR).
>
> *Medium*: Make some effort to consult already known investigators or implementers for some insights, prior to designing the trial. Addresses prior published studies and how the trial's question and outcome measures are overlapping or contiguous with other trials. Does not communicate during or after the trial any more than a funder or publication requires.
>
> *Low/ None*: Implements the trial without interviewing and absorbing feedback from local experts. Does not reach out to previous academic or PI teams, either in-country or otherwise. Uses only peer-reviewed literature to understand previous research. Keeps results, data, and insights internal when the trial ends.

Leave a footprint

Infusion of study-specific funding to a specific site may result in short-term positives for employment, commerce, education, and health. A trial that ends with

supplies perishing, lab equipment and single-study data collection software abandoned, and now-trained trial implementers left jobless can be a sour experience for a community. In fact, there needs to be a new model where we can sustain these investments across multiple studies and trials. An ideal trial might be a platform trial that continues, with multiple interventions being dropped and added over time. An ideal trial might offer follow-up study opportunities in the same site, or the opportunity for trial implementers to be intervention implementers regardless of results. All the skills used, from trial simulation to data monitoring committee staffing to surveying for true disease burden, are needed in future research—what is missing are the active approaches to sustain them.

Examples are included here, showing high, moderate and low or no fidelity to leaving a footprint:

> *High*: Engages the community early and often. Discusses disruption, including in employment as workers transition from other jobs to work for the study. Has a plan for attempting to extend the trial with additional arms. Uses a platform design to add new treatments and drop others, continue of the trial on specific sub-populations of interest, or proceed with long-term surveillance of patients to extend the trial. If a separate group investigators are planning a trial and can make use of the existing infrastructure, hands assets over. Addresses how investigation staff or implementation staff will invest in driving future post-trial research funding to that site. Includes in the implementation the use of data capture software, physical infrastructure, study-specific lab equipment, and other assets that are sustainable with a specific plan for ongoing costs-to-sustain. Offers additional follow-on trial concepts that could leverage the same local staff and site. Uses FAIR data principles. Leaves a ‘data’ footprint in enabling local data scientists by sharing the raw data appropriately in the region.
>
> *Medium*: Specifies plans and concepts addressing what will happen when the trial ends. Uses trial designs that may help ongoing footprints, short of using a platform design. Uses FAIR data principles. Rather than envisioning and driving toward continued research, collaborates with government, local, clinical site or other stakeholders for their adoption of trial infrastructure.
>
> *Low/ None*: Considers the research investment a one-time infusion in the site and that it is the responsibility of the clinical staff to find future opportunities to be part of other research studies. Only shares aggregated data, and to the minimum degree required. Does not engage the community.

Will Gates’ approach change in the future?

Our definition of highly efficient clinical trials will change in time. New trial designs will emerge. More findings will show success or failure of novel approaches. Artificial intelligence, real-world evidence, and model-based designs may offer ways to help trials

be faster. While we will look for grantees with the most fidelity to today's highly efficient trial approaches, we will also be trying to understand what to add to our current thinking.

Conclusion

We plan to collaborate with, learn from, and evangelize to other funders the benefits of highly efficient clinical trials. In the future, trial attributes such as those documented here will be required to receive funding. Existing principle investigators with experience from the pharmaceutical industry or from working with contract research organizations may be using these techniques already. BMGF continues to become more deliberate with evaluating research it funds, and making follow-on funding decisions based on design characteristics. Research that makes use of highly efficient clinical trial attributes will prove to be the bellwether and acclaimed studies that change health for the better, and make for the most well-funded research centers.

The Bill & Melinda Gates Foundation, 2019