Evaluation Framework for the NIH Single IRB Policy

September 13, 2019

This report contains findings and proposed next steps from the Single IRB (sIRB) Evaluation project team assembled by the Clinical Trials Transformation Initiative to support an NIH workgroup developing a comprehensive plan for the evaluation of the NIH sIRB policy. Its contents, and views expressed, are solely the responsibility of the authors and do not necessarily represent their employers, the official views of US Department of Health and Human Services, the Clinical Trials Transformation Initiative, the National Institutes of Health, or EnDyna Inc.

The work was supported with funds provided by the National Institutes of Health under Contract No. HHSN2762016000124U. Funding for a project meeting was provided by the US Food and Drug Administration (grant number R18FD005292) awarded to Duke University as the host of CTTI.

TABLE OF CONTENTS

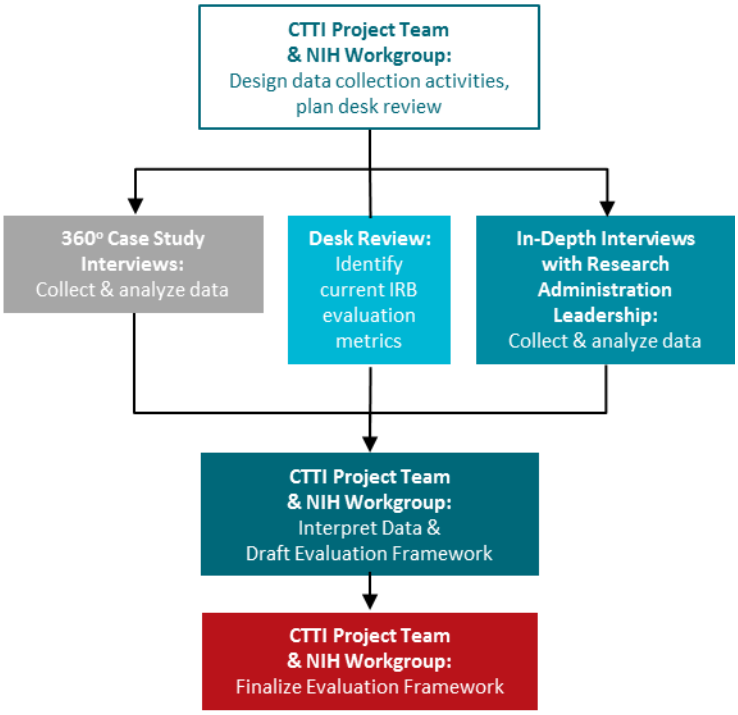
Introduction

Activating high-quality clinical trials is critical to advancing science and improving and saving lives. The Institutional Review Board (IRB) review process has been criticized for delaying clinical trial activation.[1-4] In 2014, when the National Institutes of Health (NIH) issued its draft single IRB (sIRB) policy, its stated intent was "to enhance and streamline the process of IRB review and reduce inefficiencies so that research can proceed efficiently without compromising ethical principles and protections."[5] The Final NIH Policy on the Use of a Single Institutional Review Board for Multi-Site Research was released in June 2016.[6] As of January 25, 2018, with limited exceptions, U.S. sites participating in multi-site, nonexempt, human subjects research that receive funding from the NIH are required to use a sIRB of record for ethical review required for the protection of human subjects.[7]

Additional information about the goals of the NIH policy was provided in an October 2017 presentation from the NIH Office of Extramural Programs.[8] These goals are:

1) enhance and streamline IRB review for multi-site research,
2) maintain high standards for human subjects protections,
3) allow research to proceed effectively and expeditiously,
4) eliminate unnecessary duplicative IRB review,
5) reduce administrative burden, and
6) prevent system inefficiencies.

Figure 1: Project Overview



On April 17, 2018, EnDyna, Inc., as part of an Office of Extramural Programs support services contract, released a request for proposals to develop a comprehensive evaluation plan for the NIH sIRB policy in collaboration with a policy evaluation workgroup.[9] The NIH workgroup is led by the Office of Extramural Programs and includes representatives from several NIH Institutes and Centers.

The Clinical Trials Transformation Initiative (CTTI) project team consisted of national experts in human subject protections, evaluation, and clinical research (see Appendix 1 for team member listing). The CTTI project team prepared a proposal in response to the RFP. The team's collective experience includes evaluating government policies, serving as principal investigator (PI) on NIH-funded research, participating as IRB members, administering IRBs and research programs, and holding leadership positions on the Secretary's Advisory Committee on Human Research Protections (SACHRP) and the Public Responsibility in Medicine and Research (PRIM&R) Board of Directors.

To ensure that the experts on the CTTI Project Team had a comprehensive assessment of available resources to support development of the evaluation framework (Figure 1), the project team reviewed

existing data sources and designed supplemental data collection activities. See the Data Collection Summary and Appendices 1 and 2 for full details. The three data collection activities were:

1) a desk review of existing sIRB/central IRB evaluation methods, with a literature review prepared by NIH Library services staff and the NIH workgroup;
2) 360° case study interviews at two universities that have implemented the sIRB policy as both the reviewing sIRB and the relying institution; and
3) in-depth interviews with research administration leadership.

In addition to the frequent CTTI project team teleconference discussions and check-ins with the NIH workgroup, an in-person meeting on July 31, 2019, brought the CTTI and NIH groups together to discuss and interpret the data collected in the context of creating the evaluation framework. Through additional CTTI project team teleconference meetings after the in-person meeting, this report was finalized.

Project Team Considerations and Conclusions

After an assessment of the available literature and of the sIRB processes at a sample of institutions that were following the NIH sIRB policy, the CTTI project team concluded that an evaluation of the direct impact and effectiveness of the policy would require clear definition of key data points and a case-control approach before the evaluation's implementation. Attributing outcomes directly to the policy is challenging in a clinical research environment where the sIRB model was already being implemented in response to the sIRB policy and other sponsor preferences; in preparation for the sIRB policy and Common Rule requirements; and due to pre-policy requirements of some NIH networks and National Cancer Institute studies.[7,10-15]

Although an evaluation to measure whether the NIH sIRB policy alone has enhanced and streamlined IRB review is not recommended, the CTTI project team observed and recommends the following specific actions for NIH to assist the IRB community with widespread sIRB implementation and ongoing evaluation.

1. **Define critical time points and factors in the sIRB review and approval process that all NIH grantee institutions serving as an sIRB should regularly measure**. The sIRB policy is in a relatively early stage of implementation, and the in-depth interviews confirmed the CTTI project team's experience that there is wide variation in how the policy is implemented by institutions (see Table in section 4.3 of qualitative data report). Retrospective data are available for long-performing sIRBs such as the National Cancer Institute–funded and NIH-funded trial networks and initiatives,[10,14,16-18] but these implementation examples use discordant definitions and time point collection methods. Although the time and effort to establish reliance agreements have been included in some prior comparisons of IRBs and sIRBs, measurement of other administrative demands, such as the need for lead study team communication between sites and time needed to enter information into the sIRB system on behalf of relying sites, is less common.[19-21] The CTTI project team strongly believes that establishing standard definitions and approaches to time point collection and process expectations would be valuable, not only to assess the effectiveness of using sIRB, but also to establish much needed consistency within and between institutions.

2. **Routinely collect and share established metrics with NIH Institutes and grantee institutions to promote a continuous learning environment and best practices.** Most IRBs have

addressed their internal "quality" through quality improvement (QI) program self-assessments, accreditation, and/or staff and investigator training.[22,23] These QI efforts allow for assessment of human research protection program (HRPP)/IRB qualifications and procedures and potential compliance with regulations. However, they do not provide the ability to assess efficiency or effectiveness specifically for sIRB reviews, and they offer limited comparisons across programs. Creation of well-defined quantitative and qualitative IRB- and sIRB-specific metrics, and routine collection and reporting, are needed to assess the efficiency of sIRB review. Analysis of the metrics is important to understanding how sIRB is being implemented, to identify best practices and standards, to identify unnecessary administrative burden, and to help grantees improve human subject protections while managing changes in business practice.

3. **Engage a diverse group of NIH grantee institution representatives** to address actions 1 and 2 above. The stakeholders should include a mix of large and small NIH grantee organizations, independent IRBs, multi-site study investigators and study staff, policy organizations, and other relevant parties to develop consensus practices that are feasible for all organizations.

The assumption, supported by previous comparison studies of local IRBs and sIRBs,[1,19-21] is that implementation of the sIRB model—particularly for studies subject to full board review—will reduce overall IRB member effort and time, since fewer full boards will review the same protocol across multiple sites. However, it is important that IRB review time not be the only factor considered. Details on how implementation is occurring should also be included, along with time and effort for other parts of the process, such as establishing reliance agreements; communications between lead study teams, sites, and the sIRB; and entering information into the sIRB system on behalf of relying sites. The CTTI project team recommends metrics focused on benchmarking and process improvement, as there is little agreement on metrics to measure the quality of IRB review.[24-26]

The CTTI project team recommends that the best use of resources moving forward is to develop a learning system to measure and improve the sIRB process and realize the goals of the sIRB policy. This report describes the suggested next steps and provides a framework for their implementation. These steps include engaging stakeholders; developing a foundational database to identify the population of organizations implementing the sIRB policy; developing, testing, and deploying an instrument to evaluate sIRB functions and establish and measure metrics across NIH grantee institutions and sIRBs; and using the results to continually improve the sIRB process.

Suggested Next Steps for the NIH

Engage Stakeholders

It is imperative that those implementing the sIRB model for multisite studies be involved as key participants in communicating their needs and suggesting areas for improvements in implementation of, or enhancements to, the sIRB policy. A group of stakeholders from a mix of large and small NIH grantee organizations, independent IRBs, multi-site study investigators and study staff, policy organizations, and other relevant parties should be created to advise on every step of the process described below. A series of meetings could be used, but other methods of ongoing community dialogue should be considered in order to include organizations that may lack the resources to attend in-person meetings.

Develop a Foundational Database of Organizations Implementing the sIRB Policy

It is unclear what organizations currently serve as sIRBs in accordance with the sIRB policy. A database of planned or active sIRBs is needed. There is no comprehensive list of reviewing sIRBs and relying institutions, and the absence of such a list makes it difficult to identify the population to query in any evaluation. The list could be developed from existing or new data fields on the R&R Other Project Information Form already used for all grant applications, or collected during the Just-in-Time period when a prime awardee indicates whether they are serving as the sIRB of record, relying on the IRB of a sub-awardee institution or NIH program central IRB, or contracting with an independent IRB. Only collecting the minimum necessary additional information, and doing so within an existing process, will minimize the burden of creating this database.

Proposed Evaluation of sIRB Functions Across NIH Grantee Institutions and IRBs

Purpose of Evaluation

The CTTI project team recommends that a survey instrument be developed—with broad input from NIH grantee institutions and stakeholders—and used to globally evaluate sIRB functions in order to create a learning system that continuously improves the sIRB process and furthers the goals of the policy. Working group expertise, validated by the qualitative data collection, was used to identify key questions and potential metrics (see Tables 1-7) to guide the creation of the proposed survey instrument. The survey should be pilot tested to assess the practicality and adjust the questions and definitions of metrics and milestones as needed. The final survey would then be deployed, and assessment of survey responses would be used to understand the way the sIRB model is being implemented, to identify best practices and standards, to identify unnecessary administrative burden, and to help grantees improve human subject protections while managing changes in business practices. Standard metrics and measures established through the survey should then be collected annually by the NIH, NIH Institutes, implementing organizations, and/or other relevant groups to be used for continuous improvement of the sIRB process.

Target Sample

The sIRB database should be used to select a representative sample of NIH grantee institutions implementing the sIRB policy to take part in the pilot and final surveys. Surveyed organizations at both

stages should include institutions of varying sizes and levels of NIH funding, conducting different types of research, and with different levels of experience using the sIRB model.

Type of Evaluation

The proposed evaluation framework borrows its underlying evaluation approach from the "expertise-oriented" approaches.[20] Because there is no baseline for evaluating IRB or sIRB efficiency or effectiveness for multi-site clinical trials, this evaluation should be used to compile and share nationwide sIRB processes and metrics, establish standards where appropriate, and identify areas where changes and improvements are feasible. Organizations in the process of implementing the sIRB policy will also benefit from access to the survey results to use in their own QI programs.

Evaluation Methods

The survey should consist of a set of guided assessments that ask grantee organizations to share their practices and processes. Selected organizations should be asked to collect information from their own experiences as sIRBs, the experiences of relying institutions, and the experiences of lead study teams under their grants. The information that the self-assessments provides to the NIH and participating organizations about the implementation of the sIRB policy is expected to assist in systemic improvement. The survey should be piloted, modified, and released to selected grantee organizations.

Consideration should be paid to, and consensus of stakeholders should be obtained on, the way the questions in the survey are asked. One of the challenges identified during this work is the wide variability in how different organizations define terms and tasks and how they are implementing the sIRB policy. Comparing data across institutions will, therefore, require deliberate up-front work developing the instructions for the survey. At a minimum, (1) key terms should be clearly defined; (2) quantitative data to support certain metrics (ie, time to approval) may require the development of detailed help text; and (3) an effort should be made to limit free text where possible, with look-up fields or drop-down menus provided where appropriate to ease reporting and analysis. Adjustments to annual surveys should be made to add or remove elements as sIRB measures and processes become standardized or questions about implementation are no longer relevant.

Given the approaching compliance date for the sIRB requirement in the Common Rule, the NIH should leverage its leadership and experience by working with other Common Rule agencies and the Office for Human Research Protections (OHRP) to implement the survey, promoting consistency and best practices in the use of sIRBs. Standard definitions and milestones established through the survey should be established as national standards for IRBs.

Key Evaluation Questions and Methods Crosswalk

The goals of the NIH sIRB policy were used as the domains to guide the development of the in-depth interview guides (see Appendix A of Appendix 1). As mentioned above, CTTI project team expertise, validated by qualitative data, was used to create the key questions organized by policy goal domains in Tables 1-6, and to collect organization information from the survey population in Table 7. The questions in Tables 1-7 are not intended as final survey questions, but to provide a framework of the key information to be gathered.

Timeline

If information about organizations that are implementing the sIRB model is already available at the NIH from sIRB plans or other sources, creation of the foundational database could begin immediately and be completed in 2 to 6 months. If a collection mechanism is needed, this process may take up to a year.

While the database of involved institutions and organizations is being developed, the preliminary survey should also be built. Questions should capture the practices of organizations involved in the sIRB mandate, organized by the domains described below. It is likely that several different and overlapping questions will need to be tested in the preliminary survey in each domain to find those that will be most consistently understood and reliable. Professional organizations such as the Association of American Medical Colleges (AAMC), PRIM&R, the Association for the Accreditation of Human Research Protection Programs (AAHRPP), established regional networks, the National Center for Advancing Translational Sciences (NCATS), and others could be approached for partnership.

Once this evaluation instrument is created, it should be pilot tested with a small group of grantee organizations drawn from the database of sIRB implementing organizations. The results of this test should inform selection of questions and framing for the final survey, which should then be deployed on an annual basis to assess progress in reaching the goals of the sIRB policy.

The specific metrics that will be included in the final survey should be publicly released in order for grantee organizations to begin planning for data collection ahead of the required collection period. Early public release of stakeholder-informed metrics and definitions would allow for the best evaluation results, as institutions would be better prepared for the collection. For example, if the survey is to be deployed in January 2021, it could be released in June or July of 2020. This approach is similar to how changes in Public Health Service regulations on conflicts of interest in federally funded research were evaluated in a national study.[27] One year before the effective date of the rule, the metrics that were going to be collected were announced, allowing institutions to prepare and provide the best possible data.

The stakeholder community should continue to be involved in reviewing the survey results, defining standard definitions for efficiency and effectiveness measures, determining where development of standards and tools would be useful for the research community, and suggesting improvements or changes to the policy.

Conclusion

While the project team agreed that it would be infeasible to conduct a definitive evaluation of the direct impact and effectiveness of the NIH sIRB policy due to multiple factors, NIH could and should lead the way, in partnership with other Common Rule agencies, in the ongoing evaluation of the implementation and process improvement of the sIRB model. This effort includes, but is not limited to, the development of standards and best practices based on the evaluation. NIH leadership has previously been effective in developing required training programs for human subjects research and standards for review of potential conflicts of interest, even for organizations outside the Public Health Service funding environment. We hope the NIH will take the same leadership role in the implementation and continuous improvement of the sIRB model.

Tables: Key Evaluation Questions and Methods Crosswalk

Table 1: Goal 1 Crosswalk - Enhance and Streamline IRB Review for Multi-site Research

Key Evaluation Questions	Data	Source		
		Reviewing sIRB	Relying Institution	PI
What activities does the institution consider to be included in sIRB review?	Initial ethical review of the protocol, review of consent form, ancillary reviews, continuing review, local considerations.	X	X	
What are the roles and responsibilities of the Human Research Protection Program (HRPP)/IRB staff when serving as the **reviewing** IRB?	• Number of full-time employees (FTEs) required to serve as sIRB • Titles and roles of employees in the sIRB process • Amount of time (hours) spent by these employees on sIRB activities	X		
Describe how, if at all, resource allocation has changed for the HRPP/IRB when the institution is serving as the **reviewing** IRB?	• Change in number of HRPP/IRB staff handling sIRB process (FTEs in 2017, 2018, 2019, and 2020) • Changes in roles of employees	X		
What activities does the institution consider to be part of local institutional review (reviews occurring at the **relying** institution)?	• Departmental review, ancillary reviews, HIPAA, other	X	X	
What are the roles and responsibilities of the HRPP/IRB staff when **relying** on an outside IRB?	• Number of FTEs required for relying site institutional review activities • Amount of time (hours) spent on sIRB activities		X	
Describe how, if at all, resource allocation has changed for the HRPP/IRB when the institution is **relying** on an outside IRB?	• Change in number of HRPP/IRB staff handling sIRB process (FTEs in 2017, 2018, 2019, 2020) • Changes in roles of employees	X	X	
What are the roles and responsibilities of lead study team when: submitting initial protocol to sIRB, communicating with other sites about sIRB submissions, other activities specifically related to the sIRB process?	• Number of FTEs required to complete IRB submissions and communicate with sites about sIRB submission • Amount of time (hours) spent on sIRB activities. • Change in site staff due to need to conduct sIRB activities in 2017, 2018, 2019			X
How, if at all, is the process for serving as the **reviewing** sIRB standardized? Process for serving as relying institution?	Which standardized processes or systems are being used?	X	X	
In what ways, if any, could the sIRB process be enhanced and/or streamlined?		X	X	X
How, if at all, is the process different depending on type of study (large multi-site clinical trials vs socio-behavioral/minimal risk research)?	Multi-site interventional trials vs socio-behavioral/minimal risk research?	X	X	

Key Evaluation Questions	Data	Source		
		Reviewing sIRB	Relying Institution	PI
In what ways, if any, has variability in research process and conduct changed with implementation of the sIRB mandate?	Ask for each process below: • Reliance • Submission process/Initial Review • Addition of sites • Institutional review/ancillary reviews • Informed consent forms • Events reporting	X	X	X

Table 2: Goal 2 Crosswalk - Maintain High Standards for Human Subjects Protection

Key Evaluation Questions	Data	Source		
		Reviewing sIRB	Relying Institution	PI
How does the **reviewing sIRB** obtain local considerations/context from relying institutions relevant to the study, including information related to vulnerable populations?	Written policies or procedures, reliance agreement specifications, other process	X		
In the past 12 months, how many selected study sites (relying institutions) have dropped out of a research study before sIRB review?	Number and reason for drop-out: unresolved issues around local considerations, inability to agree and execute reliance agreement, refusal to rely on sIRB	X		
How is the **reviewing** sIRB selected for a multi-site study?	• IRB characteristics, availability of expert scientific reviewer(s). • Is level of vetting dependent on risk level of study?		X	X
Is participant/patient/non-researcher viewpoint represented with the use of a sIRB?	How is viewpoint incorporated? Has the amount of input changed with sIRB review compared to multiple local IRB reviews? Note: Participating grantee organization should request viewpoint of non-research IRB member with research participant experience.	X	X	
How are unanticipated problems involving risks to subjects or others handled by the **relying** institution? By the **reviewing** sIRB?	Does institution have policies in place for reporting events specifically in sIRB model? How are differences in reporting requirements tracked? Has amount of work required changed for PI?	X	X	X
How are allegations of serious or continuing noncompliance handled by the **relying** institution? By the **reviewing** sIRB?	Who writes the corrective and preventative action plan? Who is responsible for reporting to regulatory agencies? Who is responsible for determining whether an activity constitutes serious or continuing noncompliance? Is there an appeal process? Who is responsible for reporting possible noncompliance to the sIRB?	X	X	
What suggestions, if any, would help institutions maintain high standards for human subjects protection in sIRB review?		X	X	X

Table 3: Goal 3 Crosswalk – Allow Research to Proceed Effectively and Expeditiously

Key Evaluation Questions	Data	Source		
		Reviewing sIRB	Relying Institution	PI
How does the sIRB interact with relying institutions, the study lead PI, and local investigators?	• Who is responsible for collecting site reports for submission to the reviewing sIRB? • If an eIRB system is used, who is responsible for entering information for research sites? • Who is responsible for reporting unanticipated problems involving risks to subjects or others and possible noncompliance to the reviewing sIRB?	X	X	X
What suggestions, if any, would allow research reviewed by an sIRB to proceed more effectively and expeditiously?		X	X	X
What kinds of training programs for implementation of the sIRB mandate does the institution have?	• Who receives training? • Who provides training? • What additional training programs would be helpful?	X	X	X
Describe the process for ensuring necessary institutional reviews are occurring.	Who gives the final approval for research to start at site?		X	
Suggested review time metrics included below. Consider if it is feasible and worthwhile to collect and report separately for studies undergoing expedited review and full board review. Proportion of studies of each type are collected in Organization Profile (Table 7). Specific definitions should be established in the next step of the process.				
For NEW submissions over the past 12 months, describe the median time required for approval of non-exempt human subject research at your institution when your institutional IRB is reviewing research NOT subject to sIRB requirements	Provide the median time for the following: • Time from submission to the office responsible for processing human subject research applications to final approval to conduct research at your organization • Time from submission to the office responsible for processing human subject research applications to IRB review • Time from IRB review to final approval	X		

Key Evaluation Questions	Data	Source		
		Reviewing sIRB	Relying Institution	PI
For NEW submissions over the past 12 months, describe the median time required for approval of non-exempt human subject research at your institution when your institutional IRB is serving as the sIRB on a multisite study	Provide the median time for the following: For your site • Time from submission to the office responsible for processing human subject research applications to final approval to conduct research at your organization • Time from submission to the office responsible for processing human subject research applications to IRB review • Time from IRB review to final approval For relying sites • Time to complete IRB Reliance Agreement • Time from submission to the office responsible for processing human subject research applications to approval for the relying organization • Time from submission to the office responsible for processing human subject research applications to IRB review • Time from IRB review to final approval	X		

Key Evaluation Questions	Data	Source		
		Reviewing sIRB	Relying Institution	PI
For NEW submissions over the past 12 months, describe the median time required for approval of non-exempt human subject research at your institution when you are the prime and you have chosen to sub-contract the sIRB on a multisite study to a commercial IRB or are utilizing a NIH network IRB.	Provide the median time for the following: For your site • Time from submission to the office responsible for processing human subject research applications to final approval to conduct research at your organization • Time from submission to the office responsible for processing human subject research applications to IRB review • Time from IRB review to final approval For relying sites • Time to complete IRB Reliance Agreement • Time from submission to the office responsible for processing human subject research applications to approval for the relying organization • Time from submission to the office responsible for processing human subject research applications to IRB review • Time from IRB review to final approval	X		
For NEW submissions over the past 12 months, describe the median time required for approval of non-exempt human subject research at your institution when you are a **relying site**	Provide the median time for the following: For your site • Time from submission to the office responsible for processing human subject research applications at your institution to final approval to conduct research at your organization • Time to complete IRB Reliance Agreement • Time from submission to your institution to the relying IRB • Time from submission to the relying site to IRB review • Time from IRB review to final approval	X		

Table 4: Goal 4 Crosswalk - Eliminate Unnecessary Duplicative IRB Review

Key Evaluation Questions	Data	Source		
		Reviewing sIRB	Relying Institution	PI
To what extent, if any, has the sIRB process eliminated duplicative IRB review?	IRB review of the protocol	X	X	X
What, if anything, could the sIRB process do to eliminate duplicative review?	Communication processes about which parties are completing which reviews	X	X	X
What IRB reviews are occurring at relying institutions (in purview of IRB, not other ancillary reviews)?	Who is conducting reviews? Is informed consent reviewed? If reviewed, before or after approved by sIRB?		X	
What documents are collected and stored at relying institutions?	• Informed consent, protocol, approval document from sIRB, investigator training and qualifications, other? • Purpose of collection: reference/documentation at relying institutions, used for ancillary reviews, other purpose.		X	

Table 5: Goal 5 Crosswalk - Reduce Administrative Burden

Key Evaluation Questions	Data	Source		
		Reviewing sIRB	Relying Institution	PI
What, if any, additional burdens does the sIRB process create at relying institutions? At institutions serving as the reviewing IRB?	How many different authorization agreements are being used? How many different eIRB systems are being used?	X	X	X
How, if at all, might the administrative burden be reduced? At the relying institution? At the reviewing sIRB?	Document sharing systems, access for external personnel to sIRB electronic system, communication tracking systems	X	X	X

Table 6: Goal 6 Crosswalk - Prevent Systemic Inefficiencies

Key Evaluation Questions	Data	Source		
		Reviewing sIRB	Relying Institution	PI
What, if any, systemic inefficiencies are created by the sIRB process?		X	X	X
How, if at all, is the sIRB process standardized across reviewing institutions?		X	X	X
How, if at all, might current inefficiencies be reduced or eliminated?		X	X	X
How, if at all, have IRB/HRPP policies, practices, and/or eIRB systems been updated due to the sIRB model? At the relying institution? At the reviewing sIRB?	• Practices removed, practices added/building? • Software changes? • What have the costs been to implement these changes?	X	X	

Table 7: Organization Profile Crosswalk

Key Evaluation Questions	Data	Source		
		Reviewing sIRB	Relying Institution	PI
Type of organization? • Reviewing IRB organization type? • Relying institution organization type? • Principle investigator/Lead study team organization type?	Academic institution; hospital; independent IRB; dedicated research facility; VA facility; governmental organization; contract research facility, or sponsor	X	X	X
Number of NIH funded multi-site studies for which the organization is serving as the reviewing sIRB? For NEW submissions over the past 12 months	Total number of pending/open NIH funded studies where organization is serving as the sIRB? • expedited review, full board review	X	X	
Number of NIH funded multi-site studies for which the organization is relying on an external sIRB?	Total number of pending/open NIH funded studies where organization is relying on an external IRB? • expedited review, full board review	X	X	
Total number of reliance agreements for NIH multi-site studies for which the organization is serving as the sIRB? Relying institution?	Total number of reliance agreements for pending or open studies	X	X	
Total number of electronic IRB systems used by PI/study teams?	Number of different eIRB systems used (PI)			X

Data Collection Summary

Desk Review Summary

Metrics currently collected by institutions, IRBs, and other groups to evaluate the performance and effectiveness of local IRBs and sIRBs were collected through a literature review; presentations at the 2018 Advancing Ethical Research Conference; and correspondence with established NIH sIRBs/central IRBs, academic IRBs, and multicenter study coordinating centers. The list of metrics collected and sources are included in Appendix 2. Metrics are grouped into five categories: volume, review time, staffing, costs, and quality.

Review Time: Time from IRB submission to IRB approval is often collected internally at HRPPs/IRBs and is compiled and reported by AAHRPP[32] and the Clinical and Translational Science Award (CTSA) Program Common Metrics.[16] There are also studies that compare review times between sIRBs and local IRBs.[1,14,19-21] To capture the full picture of the sIRB process, institutions have started collecting total HRPP/IRB review time (see Desk Review Metrics: Appendix 2), though total review time definitions vary. For example, some start from the request to rely and others start with HRPP/IRB submission. Similarly, end date definitions range from IRB approval of the main protocol to IRB approval of all relying sites. At this time, available review time benchmarks are limited to those collected by AAHRPP and the CTSA program. They are not sIRB-specific and do not include time for reliance agreements or local institutional reviews.[16,32] The IRB Reliance Exchange (IREx) reports full time-to-approval metrics for lead and relying member institutions starting with when sites are contacted to begin the reliance process through the time they are reviewed by the sIRB. Local (ie, relying) institution review dates, time with the IRB, and time with the study team are also measured.[17]

Volume: Volume of HRPP/IRB submissions is commonly collected and separated by level or review and type of submission (eg, full board review, expedited review; and initial, continuing review, and other). Volume metrics specific to sIRBs include the total number of studies relying on an external IRB, the number of requests for an IRB to serve as the sIRB or to rely on an outside IRB, and the total number of reliance agreements (which may be fewer than the number of relying studies if a single reliance agreement covers multiple studies).

Staffing and Costs: The total number of full-time equivalent staff is a standard metric collected by IRBs. More recently, some IRBs that are transitioning to the sIRB model have implemented time tracking programs for their employees to record staffing costs related to sIRB review.[29-31] Although costs are mainly calculated using staff time, other costs, such as upgrading or changing information systems needed for sIRB review, are also measured.

Quality and Effectiveness of IRB Review: Little information is collected on the effectiveness or quality of IRB and sIRB review.[3,24-26] Assessment of qualifications, procedures, and compliance with HRPP regulations are completed through accreditation or certification by third parties or OHRP QI program self-assessments. They do not provide the ability to assess the effectiveness of IRB or sIRB reviews or compare the quality of review across IRB programs. Post-review surveys of researchers are used by IRBs and HRPPs to identify areas for QI.[26] The absence of significant findings on external inspections or audits have been used as criteria for assessing quality when selecting an sIRB.[28] Groups have suggested conducting studies to determine the impact of common effectiveness surrogate measures—such as IRB composition, staffing, decision making, review times, regulatory compliance, and auditing—on the protection of human subjects.[24,25] However, standardized outcome measures were not established before the effective date of the NIH sIRB policy.

Most of the available literature on sIRB use and evaluation focus on quantitative methods and are primarily collected only at the institution or network level, with limited aggregate data reporting or specific measures for

sIRB review. Existing sources of compiled data on IRB operations and review time, the AAHRPP and the CTSA Common Metrics program, are limited to member organizations (n=254, n=58 respectively in 2018) and are not specific to the use of sIRBs.[16,22] Measurement of components specific to the sIRB process are being collected by individual institutions and initiatives such as IREx.[17,29-31] However, the definitions and time points used to define IRB and sIRB review times vary across organizations. Creation of well-defined IRB- and sIRB-specific metrics, and routine collection and reporting, are needed to assess the efficiency of sIRB review. Standard outcome measures are not available for assessment of the effectiveness of IRB and sIRB review or to compare quality across programs.[24-26] Development and pilot testing of reliable measures of IRB and sIRB effectiveness are needed before an assessment of the effect of sIRB on enhancing IRB review and maintaining high standards for human subjects protection can be completed.

Qualitative Research Summary

A qualitative descriptive study was conducted using in-depth interviews with (a) individuals at two universities that have implemented the sIRB process as both a reviewing sIRB and a relying institution (referred to as 360° case study interviews); and (b) research administration leadership who represent academic, independent, and health center–based IRBs and institutions (n=34). The objectives of the interviews were to describe key stakeholder experiences in implementing the NIH sIRB policy, describe steps involved in operationalizing the sIRB process at IRBs and institutions, and identify potential metrics to evaluate the implementation of the NIH sIRB policy. **The qualitative findings summarized here and fully described in the report found in Appendix 1 informed the development of the NIH sIRB evaluation framework.** The final deliverable is the evaluation framework.

In brief, the main findings that informed the development of the evaluation framework are:

1. Generally, most participants believed that the sIRB model improves, or has the potential to improve, inefficiencies associated with the local IRB model (ie, IRB review at each site) by creating consistency in the review process, standardizing documents produced for a study, reducing workload for staff at relying sites, and reducing overall duplication in ethics reviews. Most participants described that implementing the NIH sIRB policy has not streamlined ethics review when their institution has served as the sIRB; however, it has streamlined the amount of involvement of their IRBs when they are a relying institution. In addition, reviews are still required by the relying institution. These include privacy reviews and determinations, ancillary reviews, and activities related to compliance and oversight.

2. Most participants believed that the sIRB process typically becomes more efficient, or has the potential to become more efficient, once systems are created, systematic processes are followed (eg, use of common reliance agreements), and institutions gain experience and IRBs establish working relationships.

3. The sIRB model also creates new inefficiencies due to unclear roles and responsibilities for staff and institutions; a lack of systems and processes for implementing the sIRB process (eg, retooling IRB workflows, incompatibility of IRB software, and inability of relying sites to directly access the reviewing IRB's electronic systems); and added workload, particularly for investigators who must now submit the same documents to both reviewing and relying IRBs.

4. There was variation in the order and specific manner in which the sIRB steps are implemented across and within institutions (see section 4 “Process Mapping” of the qualitative data report). The steps included:
    a) The PI identifies a need for a sIRB plan.
    b) The PI and/or site investigators submit the study protocol to their own institution; the protocol is submitted to the reviewing institution as the sIRB (if their own institution is not the reviewing IRB).
    c) The relying and reviewing institutions negotiate reliance agreements.
    d) The relying institution completes ancillary reviews.
    e) The relying institution provides information on local context.
    f) The reviewing institution conducts the ethics review.
    g) The reviewing institution approves the study protocol, and the relying institution provides institutional approval.
    h) The institution(s) (the relying institution or the reviewing IRB) notifies the study teams of the protocol and institutional approvals.
    i) The institution(s) (the relying institution or the reviewing IRB) conducts post approval oversight, monitoring, and auditing.

5. Concerns were raised about the need for extensive monitoring and reporting to ensure that the high standards for human subjects protections are maintained when using a sIRB process.

6. “Shadow reviews”—in which relying IRBs still provide an ethics reviews—are being conducted by some institutions.

7. The development and use of resources and tools, such as the NCATS Streamlined, Multisite, Accelerated Resources for Trials (SMART) IRB, are helpful and assist in standardizing the process.

8. Additional processes and systems are needed and will improve the efficiency of the sIRB process (eg, establishing a well-defined definition of local context and having a central repository for institutional information).

9. Study participants’ experiences with research do not appear to have changed with the use of sIRBs.

Numerous current and new metrics were suggested for evaluating the sIRB process. Similar to the findings of the desk review, current metrics measure time in each step of the review process. Some participants reported measuring time spent pre-reviewing documents before IRB submission and time for PI training on the sIRB process. Measurement of the volume of IRB submissions and communications between IRBs and investigators were also reported. A few quality metrics were noted, including the number of modifications requested, the percentage of initial study applications approved by the reviewing IRB, and the number of errors in approved documents found by relying sites. It was noted that quality metrics will be important in evaluating the sIRB process and should continue to be developed. Suggested metrics include number of staff and time spent on sIRB activities; costs of required infrastructure changes; determining what activities are being conducted by the reviewing IRBs and relying institutions; number of communications between parties involved; and satisfaction surveys. Participants noted that the ability to collect standard metrics could be improved by the use of standardized processes and increasing the ability of relying sites to access the sIRB software system or portal. (see Section 5.0 “Metrics” of the qualitative data report for all proposed metrics.)

Definitions

Ancillary review[33] – Review conducted in coordination with IRB review to ensure that risks associated with the research are minimized and compliance requirements are met. Areas of ancillary review include radiation safety, institutional biosafety (recombinant DNA/gene transfer studies), embryonic stem cell oversight, scientific review committees, conflict of interest, IT security, clinical trials office, genomic data sharing institutional certification, environmental health and safety, nursing, and research pharmacy/controlled substances. Ancillary reviews can be deferred to the reviewing IRB with some exceptions. The responsible party should be specified in a reliance agreement or study-specific addendum.

Lead study team – Group responsible for communications, coordination, and document management associated with the use of a sIRB across all sites in a multi-site study. The overall PI should identify who will take on the role of the lead study team. This may be the PI's own study team, a coordinating center, both, or a contract research organization.

Local considerations – Any applicable state or local laws, regulations, institutional policies, standards, or other local factors, including local ancillary reviews, relevant to an instance of research.

Reviewing sIRB – The IRB of record, which provides the ethical review for all sites participating in a particular multi-site study, for the duration of the study. Also known as the sIRB.

Relying institution – The participating institution that will rely on (ie, cede IRB review to) an IRB from another institution to conduct the ethics review of a study that will be conducted at the relying institution. The NIH sIRB policy refers to these institutions as "participating sites."

Research administration leadership – Individuals in leadership positions (eg, IRB chairs, regulatory administrators) who have implemented the sIRB process (as a reviewing IRB, a relying institution, or both) either at an academic institution or with an independent IRB.

sIRB plan – A written description of how the multi-site study will comply with the NIH sIRB policy. The plan is required to be submitted as an attachment in the grant submission. Required components are available at https://grants.nih.gov/grants/how-to-apply-application-guide/forms-e/general/g.500-phs-human-subjects-and-clinical-trials-information.htm.

References

1. Wagner TH, Murray C, Goldberg J, Adler JM, Abrams J. Costs and benefits of the national cancer institute central institutional review board. *J Clin Oncol.* 2010;28(4):662-666.
2. Office of Human Research Protections (OHRP), the National Institutes of Health (NIH), the Association of American Medical Colleges (AAMC), the American Society of Clinical Oncology (ASCO), and the Department of Veterans Affairs (VA). *National conference on alternative IRB models: Optimizing human subjects protection*. Washington, DC, November 20-21, 2006.
3. Abbott L, Grady C. A systematic review of the empirical literature evaluating IRBs: what we know and what we still need to learn. *J Empir Res Hum Res Ethics.* 2011;6(1):3-19.
4. CenterWatch. 2019 Financial & Operating Benchmarks. 2019; https://store.centerwatch.com/p-592-2019-financial-operating-benchmarks.aspx.
5. National Institutes of Heath. Request for Comments on the Draft NIH Policy on the Use of a Single Institutional Review Board for Multi-Site Research. 2014; https://grants.nih.gov/grants/guide/notice-files/NOT-OD-15-026.html.
6. National Institutes of Heath. Final NIH Policy on the Use of a Single Institutional Review Board for Multi-Site Research. 2016; https://grants.nih.gov/grants/guide/notice-files/NOT-OD-16-094.html.
7. National Institutes of Heath. Single IRB Policy for Multi-site Research. 2016; https://grants.nih.gov/policy/clinical-trials/single-irb-policy-multi-site-research.htm. Accessed August 16, 2019.
8. Duque D. Single IRB and the Expectations Request Process for the Research Community. 2017; https://grants.nih.gov/sites/default/files/Single%20IRB%20&%20Exceptions%20Process%20Webinar%20October%2018%202017.pdf. Accessed August 16, 2019.
9. EnDyna I. Support NIH Workgroup to Develop a Methodology to Evaluate NIH's sIRB Policy: EnDyna, Inc; 2018.
10. Kaufmann P, O'Rourke PP. Central institutional review board review for an academic trial network. *Acad Med.* 2015;90(3):321-323.
11. Broderick JP, Palesch YY, Janis LS, National Institutes of Health StrokeNet I. The National Institutes of Health StrokeNet: A User's Guide. *Stroke.* 2016;47(2):301-303.
12. Massett HA, Hampp SL, Goldberg JL, et al. Meeting the Challenge: The National Cancer Institute's Central Institutional Review Board for Multi-Site Research. *J Clin Oncol.* 2018;36(8):819-824.
13. Bingley PJ, Wherrett DK, Shultz A, Rafkin LE, Atkinson MA, Greenbaum CJ. Type 1 Diabetes TrialNet: A Multifaceted Approach to Bringing Disease-Modifying Therapy to Clinical Use in Type 1 Diabetes. *Diabetes Care.* 2018;41(4):653-661.
14. Diamond MP, Eisenberg E, Huang H, et al. The efficiency of single institutional review board review in National Institute of Child Health and Human Development Cooperative Reproductive Medicine Network-initiated clinical trials. *Clin Trials.* 2019;16(1):3-10.
15. Federal Policy for the Protection of Human Subjects. Fed Regist. 2017;82(12):7149-7274. https://www.federalregister.gov/documents/2017/01/19/2017-01058/federal-policy-for-the-protection-of-human-subjects. Accessed July 26, 2019.
16. Clinical and Translational Science Awards (CTSA) Program. CTSA Program Common Metrics Operational Guideline: IRB Duration 2017; https://clic-ctsa.org/sites/default/files/common_metrics/Common-Metrics-Operational-Guidelines-IRB-Duration_2.2.17.pdf. Accessed August 12, 2019.
17. IRB Reliance Exchange. IREx Overview. 2018; https://www.irbexchange.org/p/wp-content/uploads/2018/11/IREx_Overview-1pager_20181114.pdf. Accessed August 12, 2018.
18. SMART IRB. The Streamlined, Multisite, Accelerated Resources for Trials IRB Reliance platform (SMART IRB) website. Available at: https://smartirb.org/about-us/. Accessed July 8, 2019.
19. Abbott D, Califf R, Morrison BW, Pierre C, Bolte J, Chakraborty S. Cycle Time Metrics for Multisite Clinical Trials in the United States. *Ther Innov Regul Sci.* 2013;47(2):152-160.

20. Vardeny O, Hernandez AF, Cohen LW, et al. Transitioning to the National Institutes of Health single institutional review board model: Piloting the use of the Streamlined, Multi-site, Accelerated Resources for Trials IRB Reliance. *Clin Trials.* 2019;16(3):290-296.
21. Neuman MD, Gaskins LJ, Ziolek T, Investigators R. Time to institutional review board approval with local versus central review in a multicenter pragmatic trial. *Clin Trials.* 2018;15(1):107-111.
22. Association for the Accreditation of Human Research Protection Programs (AAHRPP). Overview of Accreditation. http://aahrpp.org/apply/process-overview/overview. Accessed June 29, 2019.
23. Office for Human Research Protections (OHRP). OHRP QA Self Assessment Tool. 2017; https://www.hhs.gov/ohrp/education-and-outreach/human-research-protection-program-fundamentals/ohrp-self-assessment-tool/index.html. Accessed September 6, 2019.
24. Taylor HA, Ervin AM. A Measure of Effectiveness Is Key to the Success of sIRB Policy. *Am J Bioeth.* 2017;17(7):41-43.
25. Resnik DB. Some reflections on evaluating institutional review board effectiveness. *Contemp Clin Trials.* 2015;45(Pt B):261-264.
26. Hall DE, Hanusa BH, Ling BS, et al. Using the IRB Researcher Assessment Tool to Guide Quality Improvement. *J Empir Res Hum Res Ethics.* 2015;10(5):460-469.
27. Pierce HH DA, Grayer D. Implementing the Regulations on Financial Conflicts of Interest: Results from the AAMC Conflict of Interest Metrics Project. *AAMC Analysis in Brief.* 2015;15(4):1-2.
28. Coleman LM. Exploring Single IRB Models to Support Clinical and Translational Research. 2018; https://osp.od.nih.gov/wp-content/uploads/Session_II_1_Yale_University_CTSA.pdf.
29. Bailey J, Success with a single reviewing IRB serving a federally funded consortium. *2018 Advancing Ethical Research Conference*. San Diego, CA.
30. Linke M, Pivovarova E, Meschia J, Luke S, Roll S, Broderick J. Single IRB Review Improves Approval Times for a NIH-Funded Multi-Site Study. *2018 Advancing Ethical Research Conference*2018.
31. Hines S, Identifying Costs in the Evolving IRB World: Data-Tracking for Effective Costing in the Implementation of Single IRB Review. *2018 Advancing Ethical Research Conference*. San Diego, CA2018.
32. Association for the Accreditation of Human Research Protection Programs (AAHRPP). Metrics on Human Research Protection Program Performance. 2018; http://www.aahrpp.org/apply/resources/metrics-on-hrpp-performance. Accessed August 14, 2019.
33. sIRB Alliance at NYU School of Medicine. *Ancillary Review Committees*. Available from: https://static1.squarespace.com/static/5ca20d1801232c1d41308f4d/t/5d167661dc7ff100017a1997/1561753185817/Ancillary+Review+Committees.pdf. Accessed August 14, 2019.

Appendices

1. Findings from Qualitative Research to Inform the Framework
2. Desk Summary Metrics Listing