ADMINISTRATIVE STEPS FOR LAUNCHING A RANDOMIZED EVALUATION IN THE UNITED STATES

This checklist provides guidance on the logistical and administrative steps that are necessary to launch a randomized evaluation that adheres to legal regulations, follows transparency guidelines required by many academic journals, and complies with security procedures required by regulatory or ethical standards. Many of these steps require advanced planning at the beginning of the research process. The order of completion may vary by project; this list is not necessarily chronological, and many steps are interdependent as illustrated in Figure 1.

1. Identify Data Sources & Requirements
2. Obtain Ethical Approval From an Institutional Review Board (IRB) or Privacy Board
3. Design Informed Consent Process
4. Plan for Obtaining HIPAA Individual Authorization (if applicable)
5. Establish Data Use Agreements
6. Create a Data Security Plan
7. Establish Data Sharing Permissions & Protocol
8. Register the Trial

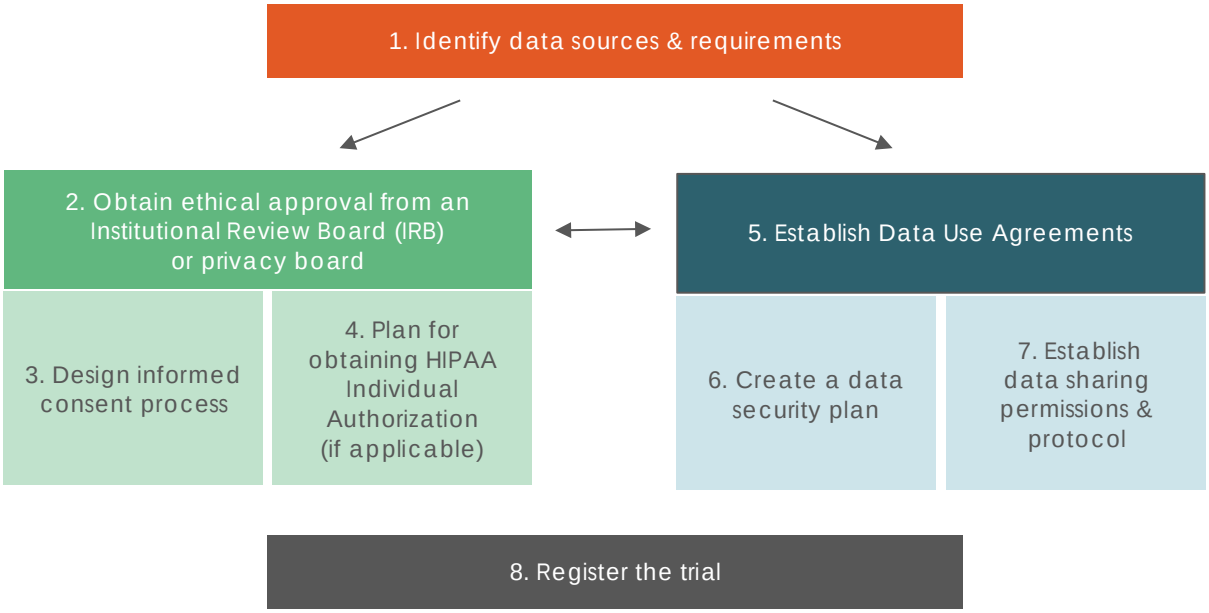
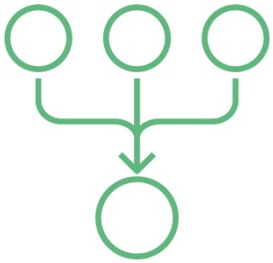


FIGURE 1. MAIN RELATIONSHIPS BETWEEN THE ADMINISTRATIVE STEPS

1. IDENTIFY DATA SOURCES & REQUIREMENTS



If the data necessary for an evaluation are collected in administrative data, using such data may lower data acquisition costs and lower the risk of certain biases relative to primary data. However, the researcher will likely need to apply for permission to use the data (see step 5 for more details) to identify a viable strategy for matching the data to her study sample, and to understand the quality, limitations, and regulations governing the use of those data. The process of getting permission to use and match the data can be time consuming, so researchers should adjust the timeline of the evaluation launch accordingly. The subject or source of the administrative data typically determines which regulations apply; education data may be subject to the Family Educational Rights and Privacy Act (FERPA), health data may be subject to the Health Insurance Portability and Accountability Act (HIPAA), and data from certain U.S. states are subject to state law.

Planning for primary data collection involves developing instruments for collecting the data (e.g., surveys, biomarkers, direct observation, spatial geography applications, etc.), validating the instruments through field-testing, training data collectors, and providing ongoing oversight and monitoring to ensure data are being collected consistently and reliably.

The level of identification of secondary data, as well as the sensitivity and types of information obtained through primary data collection, have implications for the requirements for ethics approvals, data use agreements, informed consent, and data security. Understanding how sensitive the required data are will enable the researcher to better approximate the length and rigor of the approval processes.

Resources:

- J-PAL North America's Catalog of Administrative Data Sets catalogs a number of key U.S. data sets and documents procedures on how to access data based on information provided by the originating agencies.
- This guide developed by J-PAL North America provides general, practical guidance on how to identify administrative data sources, assess their quality and contents, understand relevant requirements, and obtain and use nonpublic administrative data for a randomized evaluation.
- Resources on primary data collection can be found on the J-PAL website.
- Information about integrated data systems can be found on the Actionable Intelligence for Social Policy website.
- The U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, and National Center for Health Statistics have a toolkit on how to collect, use, protect, and share health data responsibly.
- Researchers at Harvard University are working to develop tools to help researchers understand what regulations apply to data based on an interactive questionnaire. Institutional Review Boards, data librarians, and compliance offices may also be useful to determine which regulations apply.

2. OBTAIN ETHICAL APPROVAL FROM AN INSTITUTIONAL REVIEW BOARD (IRB) OR PRIVACY BOARD

Research involving humans, or individual-level data about living humans, is likely[1] subject to review by an Institutional Review Board (IRB), even if no direct interaction between subjects and researchers is involved. Their review may determine that a research project is 1) not human subjects research and thus not subject to any further review, 2) exempt from ongoing review, 3) eligible for expedited review by an IRB administrator, or 4) subject to review by a full IRB. These determinations are made by the IRB, not by the research team, based on the study's potential impact on the health and well-being of study participants.

If multiple institutions with IRBs are involved in the research (e.g., when co-investigators are affiliated with different universities, or an implementing partner has a separate IRB), you will either need to get approval from each IRB individually, or apply for an IRB Authorization Agreement (IAA), which allows one institution to rely on another for IRB review, approval, and continuing oversight. This mechanism is recommended for research projects that involve multiple intuitions with IRBs, as it greatly expedites the IRB review process when one IRB is designated as the primary reviewer.

The IRB will likely require all investigators and study staff who will have direct interaction with study participants (e.g., to obtain consent or conduct study enrollment) or access to identifiable information to complete a human research training course such as the Collaborative Institutional Training Initiative (CITI) Program or the National Institutes of Health's (NIH) Human Subjects Protection Training.

Applying for IRB approval requires describing the research protocol and intervention in a way that enables the IRB to conduct a substantive review. In addition, any materials used to collect data (including any questionnaires or surveys used, or data use agreements used to acquire secondary data) or recruit participants (including advertising materials) must be reviewed and approved.

After approval, protocols that have not been exempt are subject to ongoing review. This includes reporting of any new staff, changes to study procedures, changes to surveys used, and reporting of adverse events. Contact the IRB of record with any questions regarding what should be reported or included in annual reviews.

For more information, see the Compliance section and the Consent and Authorization appendix of "Using Administrative Data for Randomized Evaluations."

Resources:

Review and application requirements differ by institution.

- IRBs review protocols based on the requirements of the Common Rule (45 CFR 46.116). Referencing the rules and writing protocols and applications with the rules in mind may help improve communication between the research team and the IRB and increase the chances of a successful application.

[1] There are very limited exceptions to this rule, and the exceptions may vary by institution. Most universities apply the Federal Policy for the Protection of Human Subjects (i.e., the "Common Rule") to all "human subjects" research (which includes most research involving human subjects or individual-level data about living humans) regardless of any federal requirements. Check with your IRB for details.

- The Massachusetts Institute of Technology's (MIT's) Committee on the Use of Humans as Experimental Subjects (COUHES) provides examples of what various types of application forms might look like and includes instructions.
- COUHES also has an investigator quick guide, covering when review is needed, and guidelines for various procedures in the IRB review process.
- Harvard University's Committee on the Use of Human Subjects (CUHS) provides further examples of application forms.
- The University of Colorado has several guidance documents relating to all elements of the IRB process.
- The U.S. Department of Health & Human Services presents decision charts for understanding IRB requirements.

Timeline:

Most IRBs require at least one month to complete their review, and many advise allowing for a longer review period. Harvard University, for example, suggests submitting applications at least two months before the anticipated start date of research.

Additional Considerations:

IRBs often require researchers to furnish signed data use agreements (DUAs) before fully approving a study protocol, and data providers often require IRB approval before executing a DUA. Both processes can involve lengthy review periods, and changes made by one entity must be reviewed and approved by the other. See Step 5, Establish Data Use Agreements, for more information on DUAs.

3. DESIGN INFORMED CONSENT PROCESS

Informed consent is a process in which research subjects are informed of the research procedures, goals, risks, and benefits and consent to participate in research activities – or share their personal data – voluntarily. Consent may be required even in cases where the involvement of the individual is limited to the use of his/her data. The Common Rule and institutional policy of both the researcher's institution (providing IRB review) and the data holder's agency dictate the required elements of informed consent. Educational data subject to FERPA and health data subject to HIPAA may have additional consent requirements.

Certain informed consent requirements, including the requirement to obtain written documentation of consent, may be waived or altered at the discretion of the IRB with adequate justification from the researcher. For example, IRBs may waive the written consent requirement if risk to study participants is minimal and the burden of collecting written consent is high.

In addition to IRB requirements, data providers may have requirements regarding what data from their organization will be shared and with whom, when the data will be released, and how it will be protected. The organization's review may result in changes to the study's informed consent form, which must be approved by the IRB. Data providers often require that the consent form include a description of the type of data they can provide. If such a description was not included, researchers who request administrative data for a study population from whom consent

has already been obtained may be required to re-obtain consent for each participant before a data provider will agree to release his or her administrative records (Lee, Warren, & Gill, 2015).

For more information, see the Compliance section and the Consent and Authorization appendix of "Using Administrative Data for Randomized Evaluations."

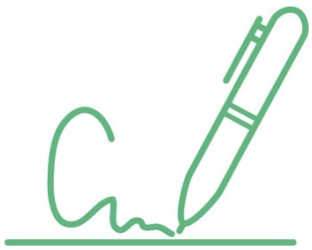
Resources:

- 45 CFR 46.116 – Common Rule requirements for informed consent (original text).
    - Paragraph (a) describes the basic, required elements of informed consent;
    - (b) describes additional elements that may be included or required depending on the study;
    - (c) and (d) describe the conditions under which IRBs may waive or approve an alteration of informed consent.
- The U.S. Department of Health & Human Services maintains a tips sheet on informed consent.
- MIT's Committee on the Use of Humans as Experimental Subjects (COUHES) provides template forms and instructions for obtaining informed consent and authorization.
- J-PAL researchers can find resources under the Human Subjects section on the J-PAL website.
- The University of Colorado's IRB provides examples and discussion of waivers of informed consent.
- The University of Chicago's IRB provides guidance on the implications of FERPA to informed consent.

Additional Considerations:

While certain elements of informed consent are requirements, slight tweaks in the written language or delivery style of the staff member guiding potential subjects through the informed consent document may impact take-up rates and compliance. Consider careful training of any individuals who will be guiding potential subjects through the informed consent process, and consider piloting the acceptability of different consent forms and methods of explanation.

4. PLAN FOR OBTAINING HIPAA INDIVIDUAL AUTHORIZATION (IF APPLICABLE)



A signed record of an individual's authorization may be required in order to obtain data from certain healthcare or health-related entities. This includes, but is not limited to, data containing personally identifiable information that are considered Protected Health Information (PHI) under HIPAA. HIPAA regulations impose criminal and/or civil penalties on individuals who use or share data inappropriately. These penalties apply to data providers and may also apply to researchers who obtain such data; therefore, data providers take many precautions before releasing data to researchers.

Certain requirements of authorization may be waived or altered at the discretion of the IRB with adequate justification from the researcher. In many cases, the authorization and informed consent process may be combined.

For more information, see the Compliance section and the Consent and Authorization appendix of "Using Administrative Data for Randomized Evaluations."

Resources:

- 45 CFR 164.502 – Uses and disclosures of protected health information (original text).
- 45 CFR 164.508 – Uses and disclosures for which an authorization is required (original text). HIPAA regulations pertaining to authorizations for the release of health information, and requirements of the authorization.
- NIH guidance on complying with HIPAA, including Authorization for research and waivers of Authorizations.
- The U.S. Department of Health & Human Services' guide to understanding the HIPAA Privacy Rule's relationship to research includes descriptions of the specific requirements of an authorization for research.

Additional Considerations:

While certain elements of authorization are requirements, slight tweaks in the written language or delivery style of the staff member guiding potential research subjects through the authorization process may impact take-up rates and compliance. Consider careful training of any individuals who will be guiding potential subjects through the authorization process, and consider piloting different methods of explaining the authorization.

5. ESTABLISH DATA USE AGREEMENTS

A data use agreement (DUA), documenting the terms under which a provider shares data and a researcher uses data, is often required in order to access data from another institution. DUAs, which typically must be approved by legal counsel at the researcher's home institution, often contain restrictions on data use, security requirements, and publication requirements that can significantly impact the underlying research or academic freedom.

For more information, including elements of particular importance in reviewing DUAs, and tips for negotiating these agreements, see the Data Use Agreements section of "Using Administrative Data for Randomized Evaluations."

Resources:

Many universities have a standard template that includes terms and conditions that are acceptable to the university and were created with researchers' needs in mind. Using a pre-vetted template may simplify the review process at the institution that created the template.

- MIT has sample data use and nondisclosure agreements here.

Timeline:

In a 2015 analysis of data acquisition efforts with 42 data agencies, MDRC found that it typically takes 7 to 18 months from initial contact with a data provider to the completion of a legal agreement.[2] Much of this time is spent in a tandem process of obtaining both IRB and legal approval for a data request. IRBs often require researchers to furnish signed DUAs before approving a study protocol, and data providers often require IRB approval before signing a

[2] See (Lee, Warren, & Gill, 2015), especially Figure ES.1 Data Acquisition Process: Typical Length of Time to Complete Each Step in The Mother and Infant Home Visiting Program Evaluation-Strong Start report.

DUA. Both processes can involve lengthy review periods, and changes made by one entity must be reviewed and approved by the other.

6. CREATE A DATA SECURITY PLAN

Any data set containing personally identifiable information – especially if it also contains potentially sensitive information – should be safeguarded. IRBs, data providers (e.g., the Centers for Medicare & Medicaid Services), and some funders (e.g., the National Institutes of Health) require descriptions of data security procedures. This includes a description of user access controls, data sharing policies, encryption or password policies, and potential for re-identification.

Ensuring proper data storage is an important component of data security. Plan to prevent loss of essential data, including raw data, treatment assignment lists, and crosswalks between de-identified study IDs and personally identifiable information. These data should be backed up regularly in at least two separate, secure locations.

Some projects may require plans for new data storage locations, such as an institutionally hosted server, an offline-only computer stored in a secure location, or simply a folder on a cloud-based storage provider (e.g., Box, Dropbox, Google Drive). Choice of storage location depends on data security requirements and data use agreement provisions.

For more information, see the Data Security section and Data Security Appendix of "Using Administrative Data for Randomized Evaluations."

Resources:

Many research universities provide support and guidance for data security through their IT departments and through dedicated IT staff in their academic departments. Researchers should consult with their home institution's IT staff when setting up data security measures, as the IT department may have recommendations and support for specific security software.

Cloud-based storage providers offer a range of options for data backups and may offer additional packages to back up data for longer periods of time to protect against the unintentional erasure of data. Institutional servers may have data backup plans, with device-level backup plans also available. Backing up data to an external hard drive (stored in a *separate location* from daily computers) is an option for low-connectivity environments.

- For researchers at MIT, Dr. Micah Altman, Director of Research at MIT Libraries, regularly presents talks on Managing Confidential Data.
- MIT's Information Systems & Technology Department provides resources on:
    - Protecting data
    - Sensitivity of data
    - Data risks
    - Secure Shell File Transfer Protocol: SecureFX
    - Encryption (including software recommendations) and whole-disk encryption
    - Removing sensitive data

- Harvard University's Research Data Security Policy (HRDSP) is an excellent resource for security level classification and security requirement examples.

7. ESTABLISH DATA SHARING PERMISSIONS & PROTOCOL

J-PAL and a number of grant-making institutions, including the National Science Foundation and the National Institutes of Health, have adopted data-sharing policies for research that they fund. Many top academic journals require data and replication code sharing as a condition of publication (e.g., American Economic Review, Econometrica, and Science.)

Developing and storing data sets and code with publication in mind will decrease the burden of preparing data later in the research process.

Permission from data providers – both individuals and data providers such as agencies or companies – may be necessary to publish even de-identified data. Data use agreements may need to include provisions allowing the publication of data, and may influence the level of dis-aggregation allowed. Review boards may require that informed consent and/or HIPAA authorization processes include any plans for the publication of data, as well.

Resources:

- J-PAL hosts resources and links to additional information about transparency and reproducibility.
- The Berkeley Initiative for Transparency in the Social Sciences (BITSS) is an excellent source of information on reproducibility and data publication. They frequently host in-person educational sessions. Their Manual of Best Practices in Transparency in Social Science Research offers suggestions on how to write replicable analysis code from the beginning.
- Innovations for Poverty Action (IPA) hosts resources on research transparency. For IPA-affiliated researchers, they offer additional support for data curation and code checking. Their manual of Best Practices for Data and Code Management covers the principles of organizing and documenting materials at all steps of the project lifecycle with the goal of making research reproducible.

8. REGISTER THE TRIAL

Registration on clinicaltrials.gov is **required** by the FDA for many medical, clinical, or health related trials.

Registration on other sites such as the American Economics Association's RCT Registry (which is supported by J-PAL) or OSF may be required by project funders, including J-PAL. Registering trials is a method of ensuring transparency. Trial registries provide a source of results for meta-analysis and may serve as a resource for locating available survey instruments and data.

Trial registration must take place prior to launch of the intervention being studied.