NEST
Coordinating Center
An initiative of MDIC

National Evaluation System for health Technology Coordinating Center (NESTcc) Methods Framework

A Report of the Methods Subcommittee of the NEST Coordinating Center – An initiative of MDIC

February 2020

National Evaluation System for health Technology Coordinating Center (NESTcc) Methods Framework

Subcommittee Members:

- *Sharon-Lise Normand, PhD (Chair); Harvard Medical School/Harvard TH Chan School of Public Health*
- *Jesse A. Berlin, ScD; Johnson & Johnson*
- *Mitchell Krucoff, MD; Duke University Medical Center/Duke Clinical Research Institute (DCRI)*
- *John Laschinger, MD; W.L. Gore and Associates*
- *Heng Li, PhD, U.S. Food and Drug Administration (FDA)/Center for Devices and Radiological Health (CDRH)/Office of Product Evaluation and Quality (OPEQ)/Office of Clinical Evidence and Analysis (OCEA)/Division of Clinical Evidence & Analysis (DCEA) 2**
- *Nilsa Loyo-Berrios, PhD, MSc, FDA/CDRH/OPEQ/OCEA/DCEA 1**
- *Joao Monteiro, PhD; Medtronic, Inc.*
- *Didier Morel, PhD; Becton Dickinson*
- *Nilay Shah, PhD; Mayo Clinic*
- *Scott Snyder, PhD; Cook Research Incorporated*

Additional Contributors:

- *Haley Abing; Harvard Medical School*
- *Tiffany Abushaikha, MS; NESTcc*
- *Rachael Fleurence, PhD, MA, MSc; NESTcc*
- *Jess Gasvoda, MPH; NESTcc*
- *Louis Jacques, MD; ADVI*
- *Dure Kim, PharmD; NESTcc*
- *Robbert Zusterzeel, MD, PhD, MPH; NESTcc*

Notes:

* This publication reflects the views of the author and should not be construed to represent FDA's views or policies.

Conflict of Interest Disclosure Information for all Subcommittee members and additional contributors can be found here.

The Methods Subcommittee appreciates the thoughtful review and public comments regarding the Methods Framework. The Subcommittees' response to these comments can be found here.

Table of Contents

Preface

The National Evaluation System for health Technology Coordinating Center (NESTcc) seeks to support the sustainable generation and use of timely, reliable, and cost-effective Real-World Evidence (RWE) throughout the medical device total product lifecycle (TPLC), using high-quality Real-World Data (RWD) that is analyzed using robust methodological standards.

Stakeholders across the medical device ecosystem, including health systems, patient groups, industry, clinicians, payers, and regulators, stand to benefit from improved use of RWE and RWD generated in the course of clinical care and everyday life. Opportunities include increased patient awareness of device safety issues, efficient and low-cost evidence generation for regulatory review and reimbursement purposes, and improved patient and provider ability to make care decisions based on robust data.

NESTcc is growing its relationship with Data Network Collaborators to advance the use of RWE generation and foster collaboration with stakeholders across the medical device field. NESTcc has surveyed its Data Network to determine current capabilities, gaps, and priority areas for improving patient outcomes using high-quality RWD generated during the routine course of care. Our Data Network currently consists of 12 Network Collaborators. Together, they represent more than 195 hospitals and 3,942 outpatient clinics and have access to over 494 million patient records. Available data sources include electronic health records (EHRs), pharmacies, public and private claims, registries, and patient-generated data (PGD).

In conjunction with the development of the Data Network, NESTcc has established Data Quality and Methods Subcommittees to support its efforts to conduct RWE studies for medical devices. Each subcommittee has developed a Framework, the content of which follows this preface. These Frameworks build upon existing bodies of work and leverage subcommittee members' knowledge and experience from similar initiatives, including PCORnet, Sentinel, and MDEpiNet, and are intended to serve as guides for medical device ecosystem stakeholders wishing to collaborate with NESTcc to ensure the quality of data and research methodology.

The Subcommittees, established in 2018, are composed of representatives from health systems, including NESTcc Network Collaborators, medical device manufacturers, and the U.S. Food and Drug Administration (FDA). The 12-member Data Quality Subcommittee and 9-member Methods Subcommittee held monthly meetings to develop their respective Framework documents from June 2018 to November 2019.

Draft versions of NESTcc's Data Quality and Methods Frameworks were circulated to Network Collaborators for review and comment, followed by a public comment period. The public comment period took place over two months from May 2019 to July 2019, during which time the Frameworks received comments from seven organizations across the medical device ecosystem. The comments were then incorporated into these initial versions for publication through the continued efforts of subcommittee members and NESTcc leadership.

The Frameworks will be updated in the future based on key findings and lessons learned from NESTcc's RWE Test-Case projects, which address two primary objectives. First, they will explore the feasibility for medical device ecosystem stakeholders to work with RWD sources and NESTcc's initial set of Network Collaborators. Second, the Test-Cases will help identify areas where NESTcc could play a role in reducing transaction costs (e.g., contracting, IRB, data sharing agreements, publication policies etc.). Test-Case concepts were solicited from stakeholders across the ecosystem, including health systems, government organizations, non-profit patient organizations, and medical device manufacturers.

The Data Quality Framework is, in its current state, based mostly around EHR data in the clinical care setting while the Methods Frameworks is applicable to many different data sources. In future iterations, the Frameworks will be moving to a more complete version, incorporating other data sources for Data Quality assessment and further real-world evidence examples and best practices for methodology to provide a more complete resource for medical device ecosystem stakeholders.

Robbert Zusterzeel, MD, PhD, MPH
Data Network Director, NESTcc

Lesley Curtis, PhD
Chair and Professor, Department of Population Health Sciences, Duke University School of Medicine
Interim Executive Director, Duke Clinical Research Institute

Sharon-Lise Normand, PhD
S. James Adelstein Professor of Health Care Policy, Department of Health Care Policy, Harvard Medical School
Professor, Department of Biostatistics, Harvard T.H. Chan School of Public Health

Introduction

The NESTcc Methods Subcommittee, consisting of a diverse range of stakeholders who each contributed their unique academic, regulatory, and industry methodological expertise, advised the NESTcc Governing Committee and staff on constructs of study design and statistical methods. The Subcommittee helps ensure that NESTcc's projects can be interpreted based on the most efficient, appropriate, and rigorous methods of analysis. Specifically, the Methods Subcommittee was tasked with developing a pragmatic methodological framework or "living playbook" that can be used by all stakeholders across the NESTcc medical device ecosystem in designing, executing, and evaluating research studies based on RWD. The Methods Framework is also intended to highlight device-specific considerations in benefit/risk studies based on both observational and experimental designs. While the Framework is closely linked to regulatory science, the principles described are applicable to any study intending to quantify cause and effect, and to descriptive studies.

The Subcommittee adopted two principles during their deliberations: pre-specification and justification for control of confounders. As a first step in developing the Methods Framework, the Subcommittee created a protocol template, which builds upon existing bodies of work and leverages the Subcommittee members' knowledge and experience from similar initiatives, including the Medical Device Epidemiology Network (MDEpiNet), PCORnet, and Sentinel. The template is intended to promote pre-specification of as much detail as possible prior to data analysis to be transparent regarding what was and was not pre-specified when presenting findings. The Subcommittee noted that the data supporting medical device evaluations could be retrospectively or prospectively collected; the data may be from electronic health records, clinical registries, insurance claims data, patients, or a combination of these sources. A critical strategy in bolstering the validity of RWE, however, is pre-specification.

The second principle adopted by the Subcommittee related to justification for and clarification as to how confounders, variables related to both medical device use and outcomes, will be controlled. Randomization that can control for both measured and unmeasured confounders is one approach. In the absence of randomization, regression, matching, or other statistical tools attempt to provide statistical control of the measured confounders. For this reason, the template developed applies to both randomized and non-randomized designs. Pre-specification of study design features and of analytical strategies will help minimize selective reporting of study results.

These efforts were conducted in parallel and are mutually complementary to the NESTcc Data Quality Subcommittee tasked with developing a Data Quality Framework. Consequently, this report does not focus on data quality but assumes that the data proposed in the protocol have been evaluated for reliability and validity for use in medical device evaluation.

Study Protocol

The planning of a study, whether a randomized trial or an observational study, involves the construction of a detailed document prospectively indicating how the study will be conducted. This document, denoted the study protocol, describes fundamental features of study design that are precisely defined at an early stage, prior to study subject enrollment. Key aspects of a study protocol, many of which are found in a PICOTS (population, intervention, comparator, outcome, timeframe, setting) framework, are described in **Table 1**. The study protocol and corresponding statistical analysis plans should be signed and dated prior to commencement of the enrollment of the first participant and data analyses, respectively. The Subcommittee developed this template with a focus on describing, at a high level, the key content relevant to each component of the protocol. While other protocol templates exist, generally for clinical trials, the point of departure made in this document is a single protocol template for retrospectively acquired or prospectively acquired data. It represents the minimum components required for *any* study. The Subcommittee's intention was to provide guidelines on what is required to conduct a scientifically valid medical device study.

Table 1. Key Components of a Study Protocol

1	Background including an understanding of the disease, available therapies, and device risk
2	Description of the device
3	Study-specific objectives
4	Target population and patient selection
5	Outcomes: primary, secondary, procedural, and device
6	Device exposure and outcome schedules
7	Study design including comparison treatments/devices, blinding, and control of confounders
8	Study procedures*
9	Required sample size
10	Study registration
11	Monitoring plans
12	Statistical analysis plan

*Note: Research involving human subjects (whether randomized or observational) should conform to standard principles. This Framework provides some of the informing or consenting considerations but emphasizes such ethical issues should be described in the protocol.

Different evidentiary requirements are needed based on the **stage of device development** (e.g., early feasibility/first-in-human, new device for new indication vs. existing approved device for indication expansion, new proprietary device for currently approved indication, iteration of approved devices and surveillance of approved devices) and the perceived risk of the device. Such diverse device assessments may require different study designs and endpoints and have different uses. For instance, for a medical device to be coverable by Medicare, evidence is needed demonstrating that the device is **reasonable and necessary** for the diagnosis or treatment of illness or injury or to improve the functioning of a malformed body member in subjects representative of the affected Medicare beneficiary population. This entails evidence generation supporting improvements in clinically meaningful patient health outcomes. Alternatively, manufacturers may be interested in understanding long-term performance of a device or researchers may focus on understanding the benefits of a marketed device compared to other devices. Study features for device evaluation at a specific stage of the device's lifecycle and for the consumers of the information are likely to differ. While this document does not discuss in detail design features specific to device stage, some examples are provided.

Medical devices are classified by FDA based on perceived risk and the regulatory controls necessary to provide a reasonable assurance of safety and effectiveness. Class I devices generally pose the lowest risk to the patient and/or user and Class III devices pose the highest risk. For regulatory purposes, most Class II devices require submission of a 510(k) (premarket notification) to demonstrate that the device is at least as safe and effective as (substantially equivalent to) a legally marketed device (predicate device). Some Class II devices require clinical data to support the substantial equivalence decision. Most Class III devices require submission of a premarket approval application (PMA) to demonstrate reasonable

assurance of safety and effectiveness. Some Class III devices are humanitarian use devices (HUDs) and are required to demonstrate reasonable assurance of safety and probable benefit through the submission of a humanitarian device exemptions (HDE) application. Guidance for specific features discussed in this Framework can be found in FDA Guidance Documents: https://www.fda.gov/RegulatoryInformation/Guidances/default.htm.

A high-level summary of guiding principles for medical device evaluation corresponding to each component listed in the protocol follows.

0.1 **References or Supporting Literature**

1. CMS. Medicare Program; Revised Process for Making Medicare National Coverage Determinations. September 26, 2003;187(68). https://www.cms.gov/Medicare/Coverage/DeterminationProcess/Downloads/FR09262003.pdf.
2. European Medicines Agency. Guidance for Industry, E6 Good Clinical Practice: Consolidated Guidance, Revision 1 (R1). July 1996. https://docplayer.net/14803071-Ich-topic-e-6-r1-guideline-for-good-clinical-practice-step-5-note-for-guidance-on-good-clinical-practice-cpmp-ich-135-95.html.
3. Sedrakyan A, Campbell B, Merino JG, et al. IDEAL-D: a rational framework for evaluating and regulating the use of medical devices. *The BMJ*. June 9, 2016;10.1136**(**353**)**:i2372. https://doi.org/10.1136/bmj.i2372**.**

1. Background: Disease, Available Therapies, and Device Risk

Introductory material presented in the protocol should include a thorough discussion of the underlying disease or physical condition enough to allow an understanding of the disease, as well as a discussion of available standard of care therapy and outcomes. This discussion should take into account the patient impact, including: disease burden, the safety and effectiveness of currently available therapies, gaps or insufficiencies in the pathophysiologic understanding of the disease or its available (standard of care) treatment options, and how the new or improved device might result in improvements in outcomes or quality of life (i.e., unmet medical needs). A description of the device (including any relevant predicate devices) and associated procedures, the device effects based on the underlying anatomy, disease pathology, and physiology, and the proposed benefits and risks of the device relative to those posed by the underlying disease as well as to those posed by currently available therapy.

This information (quantitative or qualitative) provides the backdrop necessary for understanding the proposed device's intended use (what the device label will say the device is to be used for) and indication for use (reasons for using the device), the study objective, the rationale for the proposed study design, and the adequacy of the planned clinical and statistical evaluations of evidence provided by the data. Such background information will usually be derived from various real-world data sources, including health insurance claims data, electronic health record data, and registry data. Procedural and long-term risks associated with devices that require insertion or implantation should also be discussed.

Overall, the goal of the background information is to demonstrate that based on the information presented, there is a justified rationale for conducting the study, that the study objective is reasonable and achievable, and that both ethical equipoise and sufficient safety oversight exist in order to proceed with an appropriately designed study.

1.1 General Principles to Follow

A. A description of the **disease target**, its natural history, and patient impact
B. A summary of the **currently available therapy** or therapies including:
    I. The known risks and benefits in specified patient populations
    II. The critical assessment of evidence
    III. The known outcomes
    IV. The rationale for selection of comparator therapy for the investigational protocol
C. An assessment of the **underlying/unmet need** for the therapy proposed – why the device is needed and where the device fits in:
    I. The pathophysiologic rationale for development of the device including identification of gaps or insufficiencies with current therapy
    II. The experience with existing cleared (e.g., predicate) or approved devices, drugs, biologics, or combination products, or other standard of care treatments
    III. The anatomic rationale for development of the device
    IV. A discussion of known and new risks that might result from use of the device
    V. A discussion of known and new clinical benefits that might result from use of the device
D. Inclusion of evidence predictive for finding reasonable **assurance of safety and effectiveness**, and likelihood of benefit relative to the likelihood of risk
    I. Expected safety profile for the procedure and device (expected adverse events)
    II. Expected main clinical benefit and likelihood of demonstrating the benefit is clinically meaningful
E. A summary of the reports of prior investigations, including but not limited to a summary of the **literature, clinical experience, or investigations**, relevant to the clinical study
    I. Include a discussion of why the clinical study is needed based on the absence or limitations of existing pre-clinical or clinical data
F. A discussion of a **clear mechanistic integration of how device performance results in clinical benefit** to patients specific to the device and to the clinical syndrome being studied; for instance, how a coronary stent, opening an infarcted artery, conveys benefit to a patient suffering a heart attack

1.2 **References or Supporting Literature**

1. Global Harmonization Task Force Study Group 5. Clinical Evidence – Key Definitions and Concepts. April 26, 2006. http://www.imdrf.org/docs/ghtf/archived/sg5/technical-docs/ghtf-sg5-n1r7-guidance-definitions-060426.pdf.
2. International Organization for Standardization. Clinical investigation of Medical Devices for Human Subjects – Good Clinical Practice. *ISO*. February 2011;14155(2). https://www.iso.org/standard/45557.html.
3. Kramer DB, Tan YT, Sato C, et al. Postmarket surveillance of Medical Devices: A Comparison of Strategies in the US, EU, Japan, and China. *PLoS Medicine*. September 24, 2013;10(9). https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1001519.

2. Device Description

A detailed description of the device(s) being evaluated should be included in the protocol. Relevant information for each important component, ingredient, or material that will be in contact with tissues or body fluids of the study subject is required. If the device is marketed already, specify the **brand and model number** of the device; if more than one generation of the device is used, specify all models. If Unique Device Identifiers (UDIs) are available in the data source, those should be used.

2.1 **General Principles to Follow**

A. A description of the device sufficient for understanding should include:
   I. The device and its **components** (e.g., programmer), **accessories** (e.g., delivery system), and **unique device identifier** (UDI);
   II. The device **mode** of action and intended **use**;
   III. Unique features of the device designed to **mitigate risks** or **enhance performance** or clinical benefits;
   IV. Results of pre-clinical testing for relevant bench tests, animal studies, computational modeling, biocompatibility, potential hypersensitivity, toxicity, sterilization, and manufacturing;
   V. **Sizing requirements** and **technical training** for clinical insertion or implantation of devices
   VI. Characterization of the **expected** device performance over time;
   VII. For **each component**, list its status (e.g., investigational, market released)

2.2 References or Supporting Literature

1. US Food and Drug Administration. Medical Device Accessories - Describing Accessories and Classification Pathways. Guidance for Industry and Food and Drug Administration Staff. December 20, 2017. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/medical-device-accessories-describing-accessories-and-classification-pathways. No. FDA-2015-D-0025.

3. Study-Specific Objectives

The protocol of a medical device study should contain unambiguous statements of its objectives aligned with its overall purpose, such as assessing the feasibility of the device, supporting a future premarket approval, expanding the indication of an approved device, or conducting postmarket studies for its intended stakeholders. Stakeholders include patients, patient organizations, clinicians, and payors. The objectives must be relevant, specific, based on measurable quantities, and attainable within a reasonable timeframe (**Box 1** provides an example of how study-specific objectives are defined, based on clinical justification for risks and benefits and translated into outcomes with corresponding measurement types). The objectives are typically organized by order of decreasing importance. A study objective may be operationalized by inclusion of statistical hypotheses, although this is not obligatory. A description of the key parameters of interest and basis for making conclusions, however, should be included. The choice of the primary objective(s) is important and should be made explicit; secondary objectives should be identified as such.

3.1 General Principles to Follow

A. State the overall purpose of the study and correspondingly specific objectives following the SMART principle (Specific, Measurable, Attainable, Relevant, and Time-Framed) organized to:
   I. Show how the primary objective(s) was chosen on the most straightforward, distinct clinical basis to formulate hypotheses
      a. If there are multiple primary objectives, justify each
   II. Include rationale for secondary objectives and describe how they are not directly linked to primary objective(s)
B. Specify, for devices consisting of multiple components (a "system"), if the system is the device being assessed or if a specific component is being assessed for each objective; for each study objective, precisely define the outcome measure(s) from which clinically meaningful effects in terms of risks relative to benefits can be derived, and clearly specify the type of measurement (e.g., binary, time to event)
C. For each outcome measure, precisely define the measure of association or effect on which statistical inference is to be made (e.g., absolute difference, hazard ratio, relative risk)

D. For each measure of association of interest, provide a precise description of statistical inference for device effectiveness and device safety
    I. For hypothesis testing, provide the mathematical expression for each hypothesis to be tested and the corresponding verbal statement
    II. For parameter estimation, state how resulting estimates will be used to make causal inference and contribute to evidence-based decisions

3.2 **References or Supporting Literature**

1. Friedman LM, Furberg CD, DeMets SL. Fundamentals of Clinical Trials. *Springer*. 1998. https://link.springer.com/book/10.1007/978-1-4419-1586-3.

BOX 1: Defining Study Specific Objectives: Comparative Effectiveness Multicenter Trial for Adhesion Characteristics of Ventral Hernia Repair Mesh (ClinicalTrials.gov Identifier: NCT01355939 / 2011-02112 1KM1CA156708-01 (U.S. NIH Grant/Contract). This **observational study** compares the benefits, harms, and comparative effectiveness of intraperitoneal barrier-coated and non-barrier-coated ventral hernia repair (VHR) mesh in reducing adhesions, adhesion-related complications, and adhesiolysis sequelae in actual patient subpopulations and clinical circumstances. *Specific Aim 1*: To evaluate and compare the adhesion characteristics of intraperitoneal barrier-coated versus non-barrier-coated mesh during abdominal re-exploration after prior ventral hernia repair. *Specific Aim 2*: To evaluate and compare the adhesion-related complications and adhesiolysis-related complications of intraperitoneal barrier-coated versus non-barrier-coated mesh during abdominal re-exploration after prior ventral hernia repair. These aims are "translated" into one single primary outcome (Mesh adhesiolysis time: Mesh surface area [Time Frame:

---

4. Target Population, Patient Selection, and Source for Patient Recruitment

A description of the population to which the results of the study will apply should be provided. In principle, the research "participants" (whether they are actively enrolled in a study or contained in an existing data source) should closely reflect the **population of intended use** (i.e., the target population) or sub-group of interest in certain postmarket studies. Detailed inclusion and exclusion criteria for patient selection should be established **(Box 2)**. If the criteria limit the enrolled population relative to the intended target population, those differences should be highlighted, and the exclusions should be justified. Additionally, the source of patient recruitment should be described, and if appropriate, the experience of the physicians or device operators. Other factors that should be considered are the setting in which the device is intended for use in routine practice, and the previous and current treatments of the patients being enrolled. For example, if the target population are adults with peripheral vascular disease (PVD) and a national disease registry exists, patients could be recruited and randomized.

Alternatively, a local registry comprised of health record data for adults with PVD collected during routine clinical care could serve as the basis for an observational retrospective or prospective study.

4.1 General Principles to Follow

A. Factors to consider and specify in describing the population of intended use (**target population**) should include:

    I. Disease state under study (e.g., previously untreated, measurable disease)
        a. Descriptors might include stage and severity of the condition, duration of the condition, existence (or exclusion of) specific comorbidities (e.g., diabetes), age of the population (e.g., adult vs. paediatric, adults restricted to certain age ranges), or geographic region, etc.
    II. Use of **objective criteria** for defining inclusion or exclusion features
    III. The study device (class versus specific device). In some situations, the target population will be defined by having had (or about to have) a particular procedure (e.g., implantation of a total knee replacement), regardless of the specific device implanted. Sometimes, the specific device will define the population (e.g., women who have a specific brand and type of breast implant).

B. **Specify Source of Patient Recruitment**
    I. Describe clinical centers that will be enrolling participants (for prospective, primary data collection) or treating patients (retrospective or prospective observational data)
    II. Describe readers, operators, or surgeons in centers participating in the study; the ability to gather this type of data will depend on the data sources. For example, in existing administrative data, examined retrospectively, institutions and surgeons are likely to be de-identified, but it may still be possible to provide descriptive information on procedure volume, even without identifiable operators.
    III. If the study is to be limited to certain sites (e.g., high-volume centers with highly experienced operators who are specialized and trained), note in the protocol that this population of operators may not reflect the operators in broader practice (who would be using the device once it is marketed). Indicate what plans, if any, are in place for subsequent data collection in a broader set of centers with operators who may be less highly trained.
    IV. Provide a high-level description of steps taken to assess data quality as described in the NESTcc Data Quality Framework. The description of data quality will apply to several aspects of the protocol, however, in the context of describing the target population, this assessment would be in terms of the ability to identify the target population in a valid and reliable manner.

4.2 References or Supporting Literature

1. Little RJ, Rubin DB. Statistical Analysis with Missing Data: Second Edition. August 26, 2002.
   https://onlinelibrary.wiley.com/doi/book/10.1002/9781119013563.
2. Premier Applied Sciences. Premier Healthcare Database: Data that Informs and Performs.
   November 4, 2019.
   https://www.premierinc.com/downloads/PremierHealthcareDatabaseWhitepaper.pdf.

BOX 2: EXAMPLE OF TARGET POPULATION: This study will use hospital billing records contained in the Premier Hospital Database (PHD). The Premier Health Care Alliance was formed for hospitals to share knowledge, improve patient safety, and reduce risks. Participation in the Premier Health Care Alliance is voluntary. Although the database excludes federally funded hospitals, the hospitals included are nationally representative based on bed size, geographic region, location (urban/rural) and teaching hospital status. The PHD contains complete clinical coding, hospital cost, and patient billing data from more than 600 hospitals throughout the United States. Premier collects data from participating hospitals in its health care alliance. The database contains a date-stamped log of all billed items by cost-accounting department including medications; laboratory, diagnostic, and therapeutic services; and primary and secondary diagnoses for each patient's hospitalization. Identifier-linked enrollment files provide demographic and payer information. Detailed service level information for each hospital day is recorded; this includes details on medication and devices received.

Population: *The study setting will be hospital admissions for VATS lobectomy or laparoscopic right colectomy identified within the Premier database. The study population will comprise patients undergoing VATS lobectomy or laparoscopic right colectomy during a hospital admission occurring between January 1, 2012, and September 30, 2016, for whom the endoscopic surgical stapler used in the procedure can be identified with respect to being powered vs. non-powered and with respect to manufacturer (Ethicon vs. Medtronic).*

Subject Selection: Inclusion Criteria:
- Underwent VATS lobectomy or laparoscopic right colectomy (elective or nonelective) during a hospital admission occurring between January 1, 2011, and September 30, 2016
- The first observed hospital admission, beginning on January 1, 2012, or later, meeting these criteria during this period will be designated the *index hospital admission*
- Aged ≥ 18 years or older at time of index hospital admission
- Endoscopic surgical stapler used during the index hospital admission can be identified with respect to being powered vs. non-powered and with respect to manufacturer (Ethicon vs. Medtronic)

Subject Selection: Exclusion Criteria:
- Both powered and non-powered staplers used during the index hospital admission
- da Vinci EndoWrist surgical staplers used during index admission
- *Provisional exclusion criterion*: Non-specific (i.e., not identifiable with respect to powered vs. non-powered status or brand) staplers used during index admission
- *Provisional exclusion criterion*: Evidence of robotics (laparoscopic right colectomies only; for laparoscopic non-robotics it is assumed that regardless of powered vs. non-powered stapler, the majority of anastomoses are done extracorporeally with a certain percentage intra-corporeally; with robot almost all are done intra-corporeally; there is evidence that intra-corporeal anastomoses are associated with better outcomes)
- Point of origin or admission from another institution
- *Provisional exclusion criterion*: Medicare Severity-Diagnosis Related Group that is not predominant in overall sample, not accounting for comorbidities and complications

5. Outcomes: Primary, Secondary, Procedural, and Device

The primary outcome is directly linked to the primary study objective; sometimes, more than one primary outcome may be of interest. For instance, for joint replacement, the primary outcome may be both time to revision and one-year pain assessed by a questionnaire. Secondary outcomes provide additional information that is intended to support the primary hypotheses. If the primary outcome is overall survival, the secondary outcome may be progression-free survival.

Procedural data are information generated as part of the procedure that is associated with the device use. The need for and use of procedural data will be dependent on the question of interest and data sources that may be available.

In terms of device performance, device outcomes depend upon the risk of the device. In high-risk devices, linking device performance mechanistically to outcomes in conjunction with determinations of effectiveness, safety, and benefit/risk in the context of well-defined clinical syndromes is required. Device performance measures may be multidimensional in that performance may relate to biomaterials, design features, manufacturing tolerances, operator proficiency, patient selection criteria, anatomic variations, lesion variations, or adjunctive therapies. When appropriate, involving patients in identifying outcome measures that are directly relevant to their experience of the condition should be considered.

Justification for the types of measurement for the outcomes (e.g., binary, continuous, time-to-event) should be supplied. If the expected event rates are small for the main question of interest (e.g., ipsilateral stroke), investigators may opt to create a composite measure such as the occurrence of a stroke or death or MI or hospitalization, to decrease the duration and study size. Justification that the components of the composite generally share a similar pathophysiologic mechanism should be supplied.

In observational studies, the inclusion of a control outcome, defined as an outcome unaffected by exposure to the "study" device, can strengthen the study design. While such outcomes cannot unequivocally prove the absence of bias between treatment arms, it can test a putative mechanism of bias (**Box 3**). Justification for the choice of control outcomes should be supplied. The effect of the device on the control outcome should adopt the same analytical procedure used to assess the effect of the device on the study primary outcome.

Finally, the schedule of outcomes assessments (patient or device) should be directly linked to the study objectives.

5.1 General Principles to Follow

A. **Primary and Secondary Outcomes:**
   I. Provide clear definitions of primary and secondary endpoints (outcomes) and method of outcomes assessment
   II. Primary and secondary outcome(s):

a. Provide criteria for objective classification of the outcome
b. If endpoint adjudication is required, describe rules as well as number of independence and qualifications of adjudicators
c. Characterize the misclassification rate associated with the outcome
d. Describe measures adopted to minimize data collection biases (e.g., standardized structured data capture, with harmonized definitions) including missing data
e. Prespecify the minimal clinically important difference and justify it based on existing literature
f. In non-inferiority designs, justify the choice of clinical difference and address other vulnerabilities related to adequacy of assessment instruments and negation of blinding

III. Specify the scales of each outcome (e.g., binary, failure time, categorical)
IV. Describe the rationale for using composite outcomes or surrogate outcomes, and considerations for interpretation of results
V. Specify and justify timepoints of data collection
VI. Describe what outcomes, if any, were discussed or prioritized with input from patients

B. **Procedural Outcomes:**
I. List **specific procedural** outcomes; these may include procedure time, physiological and biological data captured as part of the procedure, and procedure-specific data
    a. Capture procedural details (approach, length, etc.), success (was intended device successfully implanted), and complications (related to access, approach or acute device malfunction)
II. Describe if the data are **standardized** (e.g., are the data routinely available in a similar format across systems)
III. Characterize the **expected completeness** of data capture

C. **Device Outcomes:**
I. For permanently implantable devices, aspects of device performance may change over time; thus, clearly identify **which features** of the device will be measured
    a. Initial ability of the device to perform as intended may be eroded over time, through wear and tear, materials failures, battery depletion, infection, or temporal changes in the implant site
    b. Indicate if and how both short- and long-term device outcomes are collected
II. Report on device performance from information obtained in **pre-clinical testing**, including computer simulation, bench testing, and animal studies
    a. Include adequate assessment/re-assessment of device performance features in conjunction with adverse clinical endpoint reporting
III. Indicate and document justification why independent adjudication of whether adverse outcomes are "device related" is **not warranted**

D. **Control Outcomes** in observational studies (falsification):
    I.   Describe why the outcome is highly **unlikely** to be causally related to the device or comparator
    II. Demonstrate that the suspected confounders of the effect of the device on the control endpoint are the same as those of the effect of the device on the primary study endpoint

E. **Outcome Schedule:**
    I.   Specify timing of patient evaluation and justify the schedule, including:
         a. Baseline measurements related to patient characteristics, clinical history, and prognostic factors
         b. Measure baseline primary outcome if goal is to measure change
         c. If using patient reported outcomes, it is important to collect one or more baseline outcomes
         d. Specify that any baseline data must be measured or have occurred prior to treatment exposure
    II. Provide rationale for both short-term (e.g., 30 days) outcomes such as length of stay, intensive care unit duration, acute complications related to access or device, and late outcomes (months or years)
         a. The scheduled assessments should be based on expectations of safety events or expected benefits – is the device performing safely and having the desired effect
    III. If assessing change, then describe the schedule of assessments and justify the need to repeatedly measure
    IV. Pre-specify a list of potential adverse effects and justify the frequency of assessment

> **BOX 3: CONTROL OUTCOME**: To assess the effectiveness of arterial closure devices (ACD) for preventing complications with percutaneous coronary intervention (PCI), Wimmer, et al. (2016) undertook a retrospective analysis using the CathPCI Registry from 2009-2013 at 1,470 sites across the United States. The primary outcome was defined as vascular access site complications in patients undergoing transfemoral PCI. The control outcome was non-access site bleeding. It was found that the use of ACDs was associated with a modest absolute risk reduction in vascular access site complications. Absolute differences in non-access site bleeding were negligible, suggesting acceptable statistical control of confounding in the comparison with regard to the study primary endpoint.
> Reference: Wimmer NJ, Secemsky EA, Mauri L, et al. Effectiveness of Arterial Closure Devices for Preventing Complications with Percutaneous Coronary Intervention: An Instrumental Variable Analysis. Circulation Cardiovascular Interventions. April 9, 2016. https://www.ahajournals.org/doi/10.1161/CIRCINTERVENTIONS.115.003464.

For additional information on data capture for outcomes, please refer to the Data Quality Framework.

5.2 References or Supporting Literature

1. Lipsitch M, Tchetgen E, Cohen T. Negative Controls: A Tool for Detecting Confounding and Bias in Observational Studies. *Epidemiology.* 2010;3(21):383-388. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3053408/.
2. Prasad V, Jena AB. Prespecified Falsification Endpoints: Can they Validate True Observational Associations? *JAMA*. 2013;309(3)241-242. https://jamanetwork.com/journals/jama/article-abstract/1557739.
3. Rosenbaum PR. Design of Observational Studies. *Springer Series in Statistics, Springer*. 2010; Chapter 5. http://www.stewartschultz.com/statistics/books/Design%20of%20observational%20studies.pdf.
4. US Food and Drug Administration. Investigational Device Exemptions (IDEs) for Early Feasibility Medical Device Clinical Studies, Including Certain First in Human (FIH) Studies. Guidance for Industry and Food and Drug Administration Staff. October 1, 2013; https://www.fda.gov/ucm/groups/fdagov-public/@fdagov-meddev-gen/documents/document/ucm279103.pdf.
5. Velentgas P, Dreyer NA, Nourjah P, et al. Developing a Protocol for Observational Comparative Effectiveness Research: A User's Guide. *AHRQ*. 2013;12(13)Chapter 6-EHC099. *Agency for Healthcare Research and Quality (Rockville, MD)*. https://www.ncbi.nlm.nih.gov/books/NBK126186/.
6. Wimmer NJ, Resnic FS, Mauri L, et al. Comparison of Transradial Versus Transfemoral Percutaneous Coronary Intervention in Routine Practice. *Journal of the American College of Cardiology*. 2013;22(62):2147-2150. http://www.onlinejacc.org/content/62/22/2147.
7. Wimmer NJ, Secemsky EA, Mauri L, et al. Effectiveness of Arterial Closure Devices for Preventing Complications with Percutaneous Coronary Intervention: An Instrumental Variable Analysis. *Circ*

Cardiovasc Interv. 2017;9(4):e003464.
https://www.ahajournals.org/doi/10.1161/CIRCINTERVENTIONS.115.003464.

6. Patient Exposure to the Device

The main goals of the underlying study should be used to define exposure and outcomes. Exposure may vary based on types of devices that are being studied. For example, a device that is implanted may have a different exposure measurement compared to a device that is used to perform a procedure. The latter involves time-limited exposure while with the former, exposure could be lifelong. Exposure definitions should be as specific and detailed as possible. For studies in which detailed device information is collected de novo, the device or procedure to which patients are exposed should be known exactly. Additionally, assessment of when exposure might change for the specific device and plans to capture when and how exposure changed are critical. For example, an implanted device may be removed and knowing when this occurred and why it occurred are essential in device evaluation. The schedule of exposure assessments (patient or device) should be directly linked to the study objectives.

6.1 **General Principles to Follow**

A. Define any induction (time from device use and expected time of primary outcome) or latent (time from outcome initiation to outcome detection such as malignant tumor initiation to detection) periods. For example, an induction (run-in) period of two months was planned in which insulin treatment was intensified with a standardized titration protocol, designed to achieve optimum injection treatment (Reznik et al. [2014])

B. Describe the **units** for exposure measurement
   I. Indicate if exposure is "any" (randomized to new implant or received new implant) versus duration of exposure (e.g., number of days since breast implant date)
   II. Describe whether multiple exposures are inherent to the clinical situation. For instance, if multiple stents are implanted in a single procedure in a single patient, describe if the measurements to be made are for each patient-stent or for the first stent only

C. Describe the **precision** with which exposure will be measured; this includes the data source, misclassification error, and measurement error
   I. Specify how the device or "system," for devices consisting of multiple components, will be identified within the RWE data source (e.g., model number, UDI) and the specificity of information regarding the device use (e.g., anatomic location) (also see NESTcc Data Quality Framework)

D. Describe the approach to **confirming exposure** to the investigational device

E. Identify specific clinical or surgical aspects that **may narrow or broaden** the definition of the exposure (e.g., anterior approach for hip replacement)

F. As noted in the section on Target Population, provide information on the training and **experience of device operator/surgical team.** For instance, do surgeons require 25 hours of training or 15 cases to be proficient for the device?

G. Include **dose** of exposure (where relevant), **changes** in exposure status, and exposure to **other devices** (if multiple devices are used for the same procedure) that may impact the performance of the device being evaluated. For instance, using an intra-arterial line during a procedure likely would not affect the performance of a coronary stent.

6.2 **References or Supporting Literature**

1. Reznik Y, Cohen O, Aronson R, et al. Insulin Pump Treatment Compared with Multiple Daily Injections for Treatment of Type 2 Diabetes (OpT2mise): A Randomized Open-label Controlled Trial. *Lancet*. July 2, 2014;384(9950):1265-1272. https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(14)61037-0/fulltext.
2. Velentgas P, Dreyer NA, Nourjah P, et al. Developing a Protocol for Observational Comparative Effectiveness Research: A User's Guide. *AHRQ.* 2013;12(13)Chapter 4-EHC099. *Agency for Healthcare Research and Quality (Rockville, MD)*. https://www.ncbi.nlm.nih.gov/books/NBK126190/.

7. Study Design

A study protocol for a controlled trial or an observational study should include a detailed description of the design features used to evaluate the medical device. Fundamental features required include the number and type of comparison groups, blinding, outcomes (primary, secondary, procedural, device etc.), if a controlled trial, the experimental unit of randomization, and how randomization will occur. Additional aspects associated with device evaluations related to the effects of the device operator, the device procedure, and the complexity of the device should also be considered. The choice of study design will depend upon the ability to minimize bias, ethical issues, the practicality of executing the design, data quality, data availability, and the objectives of the study.

Specific Design

This includes a characterization of the specific study design, the number and type of treatment arms, and whether blinding is used to mask treatment (for controlled trials) or mask outcomes (for retrospective studies).

7.1 **General Principles to Follow**
A. Describe and justify the **choice of design** as precisely as possible, using standard descriptors (e.g., "a 2-group parallel sham-controlled fully blinded randomized trial," "a prospective observational cohort study," "a case-control study")
   I. If a prospective study, provide rationale for using randomization (controlled) or for not using randomization

B. Define the **primary study objective** (e.g., superiority, non-inferiority, equivalence comparison with OPC/PG, descriptive study)
C. If a randomized trial, describe and **justify treatment allocation**
    I. If unequal allocation, provide evidence that statistical efficiency is not too compromised and how such an allocation may impact the detection of adverse events in the various treatment arms. For example, if the experimental arm is twice as large as the control arm, observing a specific adverse event will be twice as likely, even in the absence of an actual device effect.
D. If an observational study and utilizing matching, describe number of matches, size of match sets, closeness of the matches, and algorithm to find matches. If weighting, specify how weights will be obtained and how extreme weights will be handled.
E. If adopting a machine learning approach to adjust for differences between participants in different treatment groups, details on the creation of training, validation, and test sets should be provided and justified.

7.2 **References or Supporting Literature**

1. Rosenbaum PR. Modern Algorithms for Matching in Observational Studies. *Annual Review of Statistics and Its Application.* March 2020;7. https://doi.org/10.1146/annurev-statistics-031219-041058.
2. Setoguchi S, Gerhard T. Comparator Selection. In: Velentgas P, Dreyer NA, Nourjah P, et al. Developing a Protocol for Observational Comparative Effectiveness Research: A User's Guide. *Agency for Healthcare Research and Quality (Rockville, MD)*. January 2013; Chapter 5. https://www.ncbi.nlm.nih.gov/books/NBK126184/.
3. US Food and Drug Administration. Design Considerations for Pivotal Clinical Investigations for Medical Devices. Guidance for Industry, Clinical Investigators, Institutional Review Boards and Food and Drug Administration Staff. November 7, 2013. https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm373766.pdf.
4. US Food and Drug Administration. ICH Harmonized Tripartite Guideline: Choice of Control Group and Related Issues in Clinical Trials E10, Step 4 Version. ICH Expert Working Group. July 20, 2000. https://database.ich.org/sites/default/files/E10_Guideline.pdf.
5. US Food and Drug Administration. Guidance on Legislation: Clinical investigators of medical devices -- statistical considerations. MHRA, Medicines and Healthcare Products Regulatory Agency. November 2013. https://www.fdanews.com/ext/resources/files/11/11-18-13-StatConsiderations.pdf.

Blinding (Masking)

The treatment that a study subject receives may be masked to all or some individuals involved in the study, including subjects, investigators, outcome assessors, and data analysts. To the extent possible, whether a randomized or observational study, proper blinding/masking is encouraged.

7.3 General Principles to Follow

A. Describe who is blinded, when they are blinded, procedures used to blind, and when the blind will be broken
   I. Rationale for lack of blinding of investigators, participants, outcome evaluators, or statisticians should be provided; other strategies to conceal treatment allocation, outcome data, and covariates should be described
   II. In observational studies, investigators should remain blinded to all endpoints until the estimation of the treatment assignment mechanism is adequate (good balance on observable characteristics between treatment arms and sufficient overlap of treatment arms)
B. Procedures used to **maintain the blind** should be included in the protocol

7.4 References or Supporting Literature

1. Karanicolas P, Farrokhyar F, Bhandari M. Blinding: Who, what, when, why, how? *Canadian Journal of Surgery*. 2010;53(5):345-348. http://canjsurg.ca/wp-content/uploads/2013/12/53-5-345.pdf.
2. Schulz KF, Grimes DA. Blinding in randomised trials: hiding who got what. *Lancet*. 2002;359:696-700. https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(02)07816-9/fulltext.
3. US Food and Drug Administration. Design Considerations for Pivotal Clinical Investigations for Medical Devices. Guidance for Industry, Clinical Investigators, Institutional Review Boards and Food and Drug Administration Staff. November 7, 2013. https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm373766.pdf.

Units of Randomization, Observation, and Analysis

Units of randomization and observation are the unit that is randomized and the unit of outcome measurement, respectively. Often, the unit of randomization is the individual subject. However, for logistical reasons, the unit of randomization could be larger, such as randomly assigning families rather than individuals to receive treated versus untreated nasal tissues. Conversely, the unit of randomization could be "smaller" than the participant, such as randomizing the right limb to receive a device and the left limb to the comparison treatment. In the limb example, the unit of observation is the "person-limb"

given outcomes are measured on each limb within a participant, a distinction that must be specified throughout study procedures as well as statistical analyses.

7.5 General Principles to Follow

A. Provide a precise definition of the randomization unit, including the rationale for the specific choice of unit
B. Include a clear definition of the unit of observation and analysis and the rationale for the choice

7.6 References or Supporting Literature

1. Rosenberger WF, Lachin JM. Randomization in Clinical Trials: Theory and Practice, Edition 1. *Wiley*. 2002. https://onlinelibrary.wiley.com/doi/book/10.1002/0471722103.

Mechanism of Treatment Assignment

This is the manner by which a treatment (device A versus B) is assigned (randomized study) or administered (observational study) to a unit when there is more than one treatment option. In randomized trials, the treatment assignment mechanism is described as known because the investigators have control of the process. In observational studies, the treatment assignment mechanism is characterized as unknown and must be estimated.

7.7 General Principles to Follow

A. Characterize and justify the **treatment assignment mechanism** when the assignment is randomized, including:
   I. Whether a fixed or adaptive randomization
   II. Whether randomization is centralized
   III. Describe stratification variable(s) such as center, operator, etc.
   IV. Describe choice of a fixed or random block size and justify choice
   V. Indicate how and by whom assignment will be communicated (in-person, phone, web, etc.)
   VI. Indicate who will know the allocation and when it will be known
   VII. Describe the time between randomization and treatment initiation and justify the length
   VIII. Provide an accounting of the number of participants: approached, eligible, provided consent, and randomized as depicted in a CONSORT diagram
B. Characterize the treatment assignment mechanism when the assignment is non-randomized (observational study) and indicate **how confounding will be controlled**:
   I. Describe variables that will be used to estimate the treatment assignment mechanism (e.g., the propensity score). If adopting machine learning, describe the process.

II. Describe procedures used to determine comparability of units in the treatment arms (e.g., standardized mean differences)

III. Specify and justify thresholds used to include subjects (e.g., any constraints on the closeness of the matches, what size weights will be truncated, variables used to match exactly, size of overlap deemed acceptable)

IV. Provide an accounting of the number of participants: approached or identified, eligible, provided consent (if required), and included in study as depicted in a CONSORT diagram

C. Describe how the treatment assignment mechanism will work when competing products enter the market while assessing a medical device

7.8 References or Supporting Literature

1. CONSORT Transparent Reporting of Trials. http://www.consort-statement.org/consort-statement/flow-diagram.

2. Götberg, M, Christiansen EH, Gudmundsdottir IJ, et al. Instantaneous Wave-free Ratio versus Fractional Flow Reserve to Guide PCI. *NEJM*. 2017;19(376):1813-23. https://www.nejm.org/doi/full/10.1056/NEJMoa1616540.

3. Rao SV, Hess CN, Barham B, et al. A Registry-Based Randomized Trial Comparing Radial and Femoral Approaches in Women Undergoing Percutaneous Coronary Intervention. *JACC Cardiovasc Interv*. 2014;8(7):857-67. https://www.sciencedirect.com/science/article/pii/S193687981400853X.

4. Rosenbaum PR, Rubin DB. The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*. April 1983;1(70):41-55. https://www.stat.cmu.edu/~ryantibs/journalclub/rosenbaum_1983.pdf.

5. Wittes, J. Randomized Treatment Assignment. *Encyclopedia of Biostatistics*. https://onlinelibrary.wiley.com/doi/abs/10.1002/0470011815.b2a01050.

Other Covariates

The collection and use of other covariates may be of interest in some designs. The identification of confounders and control outcomes are particularly important in observational studies.

7.9 General Principles to Follow

A. The following aspects should be pre-specified in the protocol:

I. Subgroups: Define (continuous vs categorical) and justify covariates describing groups of participants for which the device effect may vary

II. Confounding: Define (continuous vs categorical) and justify covariates that may impact treatment selection and outcomes in observational designs

III. If covariates are not pre-specified, justification of the approach to select variables (e.g., empirical variable selection). If adopting machine learning approaches, pre-specification of the procedure to implement the algorithm should be detailed.

IV. If categorizing covariates, provide the rationale for the choice of categories **and** ensure that the category definitions are not based on how the definition influences the estimated treatment effect

V. Characterize the completeness, quality, validity, and replicability of the covariates
   a. For additional information on completeness, quality, validity, and replicability of data, please refer to the Data Quality Framework

7.10 **References or Supporting Literature**

1. Committee for Medicinal Products for Human Use (CHMP). Guideline for adjustment of baseline covariates. *European Medicines Agency*. April 26, 2013. https://www.ema.europa.eu/documents/scientific-guideline/draft-guideline-adjustment-baseline-covariates_en.pdf.
2. Rosenbaum P. Observation and Experiment: An Introduction to Causal Inference. *Harvard University Press*. 2017. https://www.hup.harvard.edu/catalog.php?isbn=9780674975576.
3. US Food and Drug Administration. Evaluation and Reporting of Age-, Race-, and Ethnicity-Specific Data in Medical Device Clinical Studies. Guidance for Industry, Food and Drug Administration Staff. September 12, 2017. https://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM507278.pdf.
4. Velentgas P, Dreyer NA, Nourjah P, et al. Developing a Protocol for Observational Comparative Effectiveness Research: A User's Guide. *AHRQ*. 2013;12(13)-EHC099. *Agency for Healthcare Research and Quality (Rockville, MD)*. https://www.ncbi.nlm.nih.gov/books/NBK126190/.

8. Study Procedures

A clear description of how the study will be conducted ("study procedures") should be included in the protocol. Information regarding how patients are approached and consented (if required), how randomization will be conducted, how data will be collected, definitions of protocol deviations and how these will be treated, what constitutes subject withdrawal or discontinuation, and what stopping rules will be utilized, if applicable, should also be included.

Informed Consent
Consent involves informing the patient or study participant what the study involves, why it is important, what is required of the participant, and who to contact in the event of a question, among other items. It is a critical feature of clinical trials and a growing area in observational studies. Use of secondary data may also require participant consent or an IRB waiver. The Department of Health & Human Services has placed informed consent policies on the Office for Human Research Protections' website.

8.1 General Principles to Follow

A.   If no consent is required, provide rationale and supporting documents from the relevant Institutional Review Board or Research Ethics Committee
B.   Consent should be obtained prior to subject enrollment
C.   The consent process in special circumstances (e.g., subject unable to read or write, emergency treatments) should be described
D.   Include a statement indicating if vulnerable populations are included and the process for obtaining consent
E.   Provide explanation of the research (e.g., risks, benefits, study completion, study discontinuation) using language that is non-technical and understandable to the subject in a separate informed consent form (ICF), if required
F.   Provide ample time for the subject to read and understand the informed consent and to ask questions, receive answers, and consider participation
G.   Obtain dated signature acknowledging that his/her participation is completely voluntary

8.2 References or Supporting Literature

1.   University of California, Irvine Office of Research, The Regents of the University of California. How to Consent. 2019. https://research.uci.edu/compliance/human-research-protections/researchers/how-to-consent.html.
2.   US Department of Health and Human Services. Informed Consent. https://www.hhs.gov/ohrp/regulations-and-policy/guidance/informed-consent/index.html.
3.   US Department of Health and Human Services. The HIPAA Privacy Rule. https://www.hhs.gov/hipaa/for-professionals/privacy/index.html.
4.   US Food and Drug Administration. A Guide to Informed Consent. January 1998. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/guide-informed-consent.
5.   US Food and Drug Administration. IRB Waiver or Alteration of Informed Consent for Clinical Investigations Involving No More Than Minimal Risk to Human Subjects: Guidance for Sponsors, Investigators, and Institutional Review Boards. July 2017. https://www.fda.gov/media/106587/download.

Protocol Deviation Handling

Describes what types of deviations are anticipated, strategies to avoid them, and how the deviations will be handled in the study/analysis.

8.3 **General Principles to Follow**

A. Describe procedures in place to **minimize the inclusion of ineligible** participants as well as whether ineligible patients are included in the analyses
B. Describe strategies to **reduce non-compliance** (or treatment crossovers) or participant withdrawal
    I. For observational studies, clearly specify the treated populations (e.g., received at least one drug-coated stent, received exactly one drug-coated stent), describe the protocol for direct patient contact for follow-up if permitted, and specify any additional data sources that may be used to supplement follow-up information (e.g., state registry of vital statistics).
C. Because **study withdrawal and non-compliance are separate mechanisms**, distinct approaches to minimizing both should be included
    I. For observational studies, a minimum duration of follow-up should be specified (if any). Non-compliance may be determined in some studies (prospective observational studies) but for many, intended treatments may be unknown.
D. Provide procedures to **minimize the number of assessments** made outside of a follow-up interval

8.4 **References or Supporting Literature**

1. Mohan S, Mehra M, Petrizzo M, et al. A Toolkit for the Management of Protocol Deviations. *Therapeutic Innovation & Regulatory Science.* November 2016;50(6):791–800. https://link.springer.com/article/10.1177%2F2168479016647987.

9. Required Sample Size

The determination of sample size is a critical component of the design of a clinical study, whether randomized or observational (**Box 4**). If the sample size is too small, firm conclusions are unlikely to be inferred or results might be obtained by chance. On the other hand, an excessively large sample size would be wasteful and unethical and could lead to a statistically significant finding for an effect that is not clinically meaningful. A clinically meaningful effect size should be used as the basis for the sample size calculations. In practice, the study sample size is determined based on several design parameters and follows a set of statistical principles. Not all study designs require that sample size be fixed before the beginning of the study. In a group sequential design or an adaptive design, the eventual sample size depends on the trajectory of outcome data. In these designs, a stopping rule is used rather than a unique sample size, which will then fluctuate based on the results of interim analyses. Nonetheless, the same basic statistical principles apply.

9.1 General Principles to Follow

A. Indicate the type of **study design**:
   I. Fixed sample size
   II. Group sequential (see interim analysis and stopping rule topic, section 11 Monitoring Plan)
   III. Adaptive (see interim analysis and stopping rule topic, section 11 Monitoring Plan)

B. Indicate approach to **evaluation**:
   I. If an estimation approach is adopted, provide and justify assumptions regarding widths of confidence intervals and estimated effect size
   II. If a hypothesis testing approach is adopted, specify null and alternative hypotheses (basis for margin for a non-inferiority test), method of testing, test statistic, anticipated effect size (justify), power, and type I error rate/significance level
   III. Justify the selection of one-sided versus two-sided confidence intervals (or one-sided vs two-sided hypothesis test)

C. Indicate and justify **additional features** of the study that impact sample size:
   I. Adjustment for multiplicity (e.g., hierarchical testing or simultaneous confidence intervals)
   II. Adjustment for clustering (e.g., center effects)
   III. Approach to controlling for confounding variables
   IV. Prevalence/incidence rates (reference and control cohort)
   V. Accounting for missing data
   VI. Correction for loss to follow-up, treatment discontinuation, or other forms of censoring

9.2 References or Supporting Literature

1. European Medicines Agency. ICH Topic 3: Structure and Content of Clinical Study Reports. July 1996. https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e-3-structure-content-clinical-study-reports-step-5_en.pdf.
2. Goodman SN, Berlin JA. The use of predicted confidence intervals when planning experiments and the misuse of power when interpreting results. *Ann Intern Med*. 1994;121:200-6. https://www.ncbi.nlm.nih.gov/pubmed/8017747.
3. US Dept of Health and Human Services, US Food and Drug Administration. Guidance for Industry: E9 Statistical Principles for Clinical Trials. ICH Expert Working Group. February 1998. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/e9-statistical-principles-clinical-trials.
4. US Food and Drug Administration. Use of Real-World Evidence to Support Regulatory Decision-Making for Medical Devices. Guidance for Industry, Food and Drug Administration Staff. July 27, 2017.

https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm513027.pdf.

5. Velentgas P, Dreyer NA, Nourjah P, et al. Developing a Protocol for Observational Comparative Effectiveness Research: A User's Guide. *AHRQ*. 2013;12(13)Chapter 9-EHC099. *Agency for Healthcare Research and Quality (Rockville, MD)*. https://www.ncbi.nlm.nih.gov/books/NBK126183/.

BOX 4: SAMPLE SIZE JUSTIFICATION. *Insulin Pen Needles: Effects of Extra-Thin Wall Needle Technology on Preference, Confidence, and Other Patient Ratings* (ClinicalTrials.gov Identifier: NCT01852136 / DBC-11-NEXXT01). A sample size of 180 patients (all patients pooled) was determined to give 95% power to detect an average relative difference of 10 mm on the VAS (assuming an SD of 37 mm for relative VAS scores, based on results from a previous study and a t-test procedure). In addition, a sample size of 180 patients was sufficient to provide 90% power to detect a significant preference for investigated PNs (based on a Monte-Carlo simulation). A sample size of 60 patients for each pen brand with the same SD gives 90% power to detect an average relative difference of 16 mm on the VAS. To obtain at least 180 evaluable patients, target enrollment was 210 patients. The enrollment of 30 patients over the target was considered to be sufficient because the attrition rate was anticipated to be low due to the short study duration, without any changes to patients' usual insulin therapy. Reference: Aronson R, Gibney MA, Oza K, et al. Insulin Pen Needles: Effects of Extra-Thin Wall Needle Technology on Preference, Confidence, and Other Patient Ratings. *Clinical Therapeutics*. 2013;7(35):923-933.e4. https://www.ncbi.nlm.nih.gov/pubmed/23790553.

---

10. Study Registration

Registration of randomized trials is standard practice and is required by publication policies at major journals and by governmental regulations. Trial registration helps prevent selective analysis and reporting of endpoints. As an example, when trial results for the primary endpoint are not favorable, and secondary endpoints are favorable, registration allows the reader to make an informed judgment about the appropriateness of the reporting and the validity of the emphasis on secondary endpoints, if those endpoints become the focus of a publication.

The value of registering observational study protocols is increasing with the goal of enhancing reproducibility and credibility. Because observational studies are not the focus of www.clinicaltrials.gov, registration can be cumbersome. However, other venues are available. The Center for Open Science, for instance, provides pre-registration (among other services) for observational studies. Pre-specification and publication for all studies is strongly encouraged, will make the best evidence available, will assure a high degree of transparency, and will reduce ethical questions of conflict of interest.

10.1 General Principles to Follow

A. Trials should be **registered** on the www.clinicaltrials.gov website prior to enrolling the first patient, with no exceptions; observational studies could also be registered at this site or at the Center for Open Science: https://cos.io/.

10.2 References or Supporting Literature

1. Berger ML, Mamdani M, Atkins D, et al. Good Research Practices for Comparative Effectiveness Research: Defining, Reporting and Interpreting Nonrandomized Studies of Treatment Effects Using Secondary Data Sources: the ISPOR Good Research Practices for Retrospective Database Analysis Task Force Report--Part I. *Value in Health*. November-December 2009;12(8):1044-1052. https://www.valueinhealthjournal.com/article/S1098-3015(10)60308-7/abstract.
2. Chavers S, Fife D, Wacholtz M, et al. Registration of Observational Studies: Perspectives from an Industry-Based Epidemiology Group. *Pharmacoepidemiology and Drug Safety*. October 2011;20(10):1009-1013. https://onlinelibrary.wiley.com/doi/abs/10.1002/pds.2221.
3. ClinicalTrials.gov. https://clinicaltrials.gov/.
4. Davis F, Caban-Martinez A. Sense of the College: Should Observational Studies be Registered? *Ann Epidemopl*. May 2011;5(21):382-6. https://www.sciencedirect.com/science/article/abs/pii/S1047279710004588?via%3Dihub.
5. Editors. The Registration of Observational Studies--When Metaphors Go Bad. *Epidemiology*. September 2010;5(21):607-609. https://journals.lww.com/epidem/FullText/2010/09000/The_Registration_of_Observational_Studies_When.9.aspx.
6. Feinstein AR. Scientific Standards In Epidemiologic Studies of the Menace of Daily Life. *Science*. December 2, 1988;4883(242):1257-1263. https://science.sciencemag.org/content/242/4883/1257.
7. Gliklich RE, Dreyer N. Registries for Evaluating Patient Outcomes: A User's Guide. *AHRQ*. April 2007. http://bok.ahima.org/PdfView?oid=79053.
8. ICMJE. http://www.icmje.org/.
9. ISPE. Guidelines for Good Pharmacoepidemiology Practices (GPP). *Pharmacoepidemiology and Drug Safety*. 2008;17(2):200-208. https://www.pharmacoepi.org/pub/?id=1c2a23af-2354-d714-516a-7175549e3a88.
10. ISPOR. Good Practices for Outcomes Research. https://www.ispor.org/heor-resources/good-practices-for-outcomes-research.
11. Lash T. Preregistration of Study Protocols Is Unlikely to Improve The Yield From Our Science, but Other Strategies Might. *Epidemiology*. September 2010;5(21):612–613. https://journals.lww.com/epidem/Fulltext/2010/09000/Preregistration_of_Study_Protocols_Is_Unlikely_to.11.aspx.

12. Motheral B, Brooks J, Clark MA, et al. A Checklist for Retrospective Database Studies--Report of The ISPOR Task Force on Retrospective Databases. *Value in Health*. March-April 2003;2(6):90-97. https://www.valueinhealthjournal.com/article/S1098-3015(10)60137-4/pdf?_returnURL=https%3A%2F%2Flinkinghub.elsevier.com%2Fretrieve%2Fpii%2FS1098301510601374%3Fshowall%3Dtrue.
13. Poole C. A Vision of Accessible Epidemiology. *Epidemiology*. September 2010;5(21):616-618. https://journals.lww.com/epidem/Fulltext/2010/09000/A_Vision_of_Accessible_Epidemiology.13.aspx.
14. Samet JM. To Register or Not to Register. *Epidemiology*. September 2010;5(21):610-611. https://journals.lww.com/epidem/Fulltext/2010/09000/To_Register_or_Not_To_Register.10.aspx.
15. Takkouche B, Norman G. Meta-Analysis Protocol Registration: Sed Quis Custodiet Ipsos Custodes? [But Who Will Guard the Guardians?]. *Epidemiology*. September 2010;5(21):614-615. https://journals.lww.com/epidem/Fulltext/2010/09000/Meta_analysis_Protocol_Registration__Sed_quis.12.aspx.
16. Vandenbroucke J. Preregistration of Epidemiologic Studies: An Ill-Founded Mix of Ideas. *Epidemiology*. September 2010;5(21):619-620. https://journals.lww.com/epidem/Fulltext/2010/09000/Preregistration_of_Epidemiologic_Studies__An.14.aspx.

11. Monitoring Plan

Monitoring clinical investigations **(Box 5)** is essential not only for the protection of human subjects, but also for the conduct of high-quality studies. Appropriate monitoring plans help ensure protection of the rights, welfare and safety of the human subjects, and the quality of the study data pursuant to Good Clinical Practice (GCP) standards. Reasons for study monitoring include protocol compliance, adverse events, treatment comparisons to stop a trial (early if needed), data management to identify data errors or missingness, and study futility. Use of an independent Data Safety Monitoring Board/Committee may not only ensure human subject safety but also reduce bias in study management. In certain RWE applications, such as retrospective analysis of existing data, monitoring functions may be more limited to assessment of data quality (e.g., missingness, out of range variables) and potential pathways to address such issues (also see the NESTcc Data Quality Framework).

11.1 **General Principles to Follow**

A. **Data Safety Monitoring Boards/Committees**
   I. Describe the **charge** of the data safety monitoring committee, members and their expertise, frequency of meetings, and procedures in the DSMC Charter
   II. Describe the **processes for providing unblinded** data tables to independent committees without undermining central study integrity (indicate who is blinded to what information and when blinding is revealed)

III. Provide a description for periodicity of data review and formal approach to stopping rule(s)

B. **Site-Based and Central Data Monitoring**
I. Describe the process for site-based and central data quality monitoring including members and how data issues will be resolved
II. Describe data query, resolution and final documentation processes including audit trail technology consistent with the electronic records, electronic signatures – scope and application portion of FDA Part 11 compliance

C. **Interim Analyses**
I. Define operational procedures for the committee interpreting interim analyses (Steering Committee, Data Safety Committee, etc.)
II. Define the purpose of any interim analyses (e.g., early stopping for futility, for efficacy, for safety, for adaptive designs, or potential mid-course corrections)
III. Describe and justify the number and frequency of analyses
  a. If stopping rules are part of a specific dynamic study design, describe rules for stopping for futility, efficacy, or continuing and how sample size is impacted
  b. Pre-specify rule for stopping for safety
  c. Provide clinical and statistical justification for stopping rules
IV. Describe and justify sample size, type I error, and alpha spending functions, and how the interim analyses impact the sample size needed for the primary outcome

11.2 **References or Supporting Literature**

1. Bruno A, Kent TA, Coull BM, et al. Treatment of Hyperglycemia in Ischemic Stroke (THIS): A Randomized Pilot Trial. *Stroke*. 2008;39(2):384-389. https://doi-org.su.idm.oclc.org/10.1161/STROKEAHA.107.493544.
2. Chang, M. Introductory Adaptive Trial Designs: A Practical Guide with R. *Chapman & Hall/CRC Biostatistics Series.* 2015. ISBN-10: 1498717462. https://www.crcpress.com/Introductory-Adaptive-Trial-Designs-A-Practical-Guide-with-R/Chang/p/book/9780367377601.
3. Chow SC, Chang M. Adaptive Design Methods in Clinical Trials. *CRC Press*. December 2011. ISBN 9781439839874. https://www.crcpress.com/Adaptive-Design-Methods-in-Clinical-Trials/Chow-Chang/p/book/9781439839874.
4. Johnston KC, Hall CE, Kissela BM, et al. Glucose Regulation in Acute Stroke Patients (GRASP) Trial: A Randomized Pilot Trial. *Stroke*. October 2009;40:3804-3809. https://doi.org/10.1161/STROKEAHA.109.561498.
5. Mehta CR, Pocock SJ. Adaptive Increase in Sample Size when Interim Results are Promising: A Practical Guide with Examples. *Statistics in Medicine*. December 10, 2011;30(28):3267-3284. https://doi.org/10.1002/sim.4102.

6. National Institute of Neurological Disorders and Stroke rt-PA Stroke Study Group. Tissue Plasminogen Activator for Acute Ischemic Stroke. *New England Journal of Medicine*. December 14, 1995;333(24):1581-1587. https://www.nejm.org/doi/full/10.1056/NEJM199512143332401.

7. O'Brien PC, Fleming TR. A Multiple Testing Procedure for Clinical Trials. *Biometrics*. September 1979;35(3):549-556. *JSTOR*. www.jstor.org/stable/2530245.

8. Palesch YY, Hill MD, Ryckborst KJ, et al. The ALIAS Pilot Trial: A Dose Escalation and Safety Study of Albumin Therapy for Acute Ischemic Stroke--II: Neurologic Outcome and Efficacy Analysis. *Stroke*. August 2006;37(8):2107-2114. https://www.ahajournals.org/doi/10.1161/01.str.0000231389.34701.b5.

9. Pocock SJ. Group sequential methods in the design and analysis of clinical trials. *Biometrika*. 1977;64(2):191-199. *JSTOR*. https://www.jstor.org/stable/i315486.

10. US Food and Drug Administration. Adaptive Designs for Clinical Trials of Drugs and Biologics: Guidance for Industry. November 2019. https://www.fda.gov/downloads/drugs/guidances/ucm201790.pdf.

11. US Food and Drug Administration. Adaptive Designs for Medical Device Clinical Studies. Guidance for Industry and Food and Drug Administration Staff. July 2016. https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm446729.pdf.

12. US Food and Drug Administration. Balancing Premarket and Postmarket Data Collection for Devices Subject to Premarket Approval: Guidance for Industry and Food and Drug Administration Staff. April 2015. https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm393994.pdf.

13. US Food and Drug Administration. Design Considerations for Pivotal Clinical Investigations for Medical Devices: Guidance for Industry, Clinical Investigators, Institutional Review Boards and Food and Drug Administration Staff. November 2013. https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm373766.pdf.

14. US Food and Drug Administration. Guidance for Industry: Oversight of Clinical Investigations – A Risk-Based Approach to Monitoring. August 2013. https://www.fda.gov/downloads/Drugs/Guidances/UCM269919.pdf.

BOX 5: EXAMPLE FOR STOPPING RULES IN AN ADAPTIVE DESIGN USING O'BRIEN AND FLEMING GUIDELINES.

The Stroke Hyperglycemia Insulin Network Effort (SHINE) trial protocol: a randomized, blinded, efficacy trial of standard vs. intensive hyperglycemia management in acute stroke (ClinicalTrials.gov Identifier: NCT01369069). The sample size estimate was based on data from the two NIH funded pilot trials, as well as other relevant acute stroke trials (see references 11-14 above). These data supported an estimate of 25% favorable outcome rate in the control group. The minimal clinically relevant absolute difference in favorable outcome between the two treatment groups was estimated to be 7% (control group = 25%; intervention group = 32%). The study is therefore powered to detect an absolute 7% difference in favorable outcome between the groups. The study design includes four interim analyses for both efficacy and futility of the primary outcome (after 500, 700, 900, and 1,100 patients complete the study) and a final analysis for a total of five planned analyses of the primary outcome. Including a 3% non-adherence rate and the four interim analyses, approximately 1,400 randomized patients are needed to provide 80% power with a two-sided type I error rate of 0.05. Reference: Bruno A, Durkalski VL, Hall CE, et al. The Stroke Hyperglycemia Insulin Network Effort (SHINE) Trial Protocol: A Randomized, Blinded, Efficacy Trial of Standard vs. Intensive Hyperglycemia Management in Acute Stroke. *International Journal of Stroke*. 2014;9(2):246-251. https://doi.org/10.1111/ijs.12045.

---

12. Statistical Analysis Plan (SAP)

The statistical analysis plan provides the detailed description of all statistical analyses to be conducted once the data are available. The **contents of the SAP in the protocol** are often less detailed than the final SAP, which might be a separate document. The SAP appearing in the protocol must be approved prior to enrollment of first subject. **If the SAP is a separate document it must include a description of the study objective, design, procedure, endpoints, and analysis population to provide enough context for the correct evaluation of the SAP.**

12.1 General Principles to Follow

    A. Definition and justification of target population and study samples
        I. Intention-to-Treat sample for effectiveness: all randomized subjects
        II. Safety Sample: all subjects receiving the study treatments (e.g., implanted with a pacemaker)
    B. Indicate the **treatment of missing data,** associated assumptions, and how the associated assumptions will be validated

C. Provide information about specific datasets, where they are stored, and what analyses are planned
D. Specify the intended statistical software
E. Define/describe computation of derived variables
F. Define study success criteria
G. Provide description of statistical models, statistical hypotheses, tests, and estimation for:
    I.   Analyses of primary, secondary, procedural, device, and safety outcomes
    II. Interim analyses
    III. Subgroup analyses
H. Provide a plan for adjustment **for multiplicity** of all endpoints, with the possible exception of safety endpoints
I. Describe sensitivity analyses including the feature addressed and assumptions made
J. Provide examples of tables and graphs
K. Describe and justify the interim analysis plan and its impact on statistical design (type 1 error spending function, similar to previous section)
L. Pre-specify how learning curve effects will be handled
M. If a noninferiority design, justify the acceptable or tolerable clinical difference
N. For observational studies using machine learning or variable selection procedures, specify the process that will be adopted. For instance, if using Bayesian Additive Regression Trees, describe the number of trees, the size of the cross-validation samples, and the prior distributions for the number of variables. For regression selection approaches, indicate if stepwise selection will be used, the requirements for variables to enter or to exit, etc.

12.2 **References or Supporting Literature**

1. Gamble C, Krishan A, Stocken D, et al. Guidelines for the Content of Statistical Analysis Plans in Clinical Trials. *JAMA*. 2017;318(23):2337-2343. https://jamanetwork.com/journals/jama/fullarticle/2666509.
2. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH Harmonised Tripartite Guideline: Statistical Principles for Clinical Trials E9. *European Medicines Agency (London, England)*. 1998. https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e-9-statistical-principles-clinical-trials-step-5_en.pdf.
3. Rose S, Normand S-L. Double Robust Estimation for Multiple Unordered Treatments and Clustered Observations: Evaluating Drug-Eluting Coronary Stents. *Biometrics*. 2018. https://doi.org/10.1111/biom.12927.
4. Schuler MS, Rose S. Targeted Maximum Likelihood Estimation for Causal Inference in Observation Studies. *American Journal of Epidemiology*. 2017;185(1):65-73. https://doi.org/10.1093/aje/kww165.

13. Future Work

The methodological considerations presented describe the basic components of a protocol for device evaluation. Future work will focus on study designs that leverage both randomization and observational data, approaches for balancing treatment arms including machine learning, and strategies for assessing the robustness of conclusions drawn from a specific design and data. The Subcommittee developed the Framework in the context of implantable or limited duration use devices. Additional considerations will be required when assessing imaging modalities and diagnostics.

CONTACT INFORMATION

For more information,
please contact
NESTcc at nestcc@mdic.org

NEST
Coordinating Center
An initiative of MDIC