Methodological Considerations in Medical Device Evaluations

A Report from the National Evaluation System for health Technology Coordinating Center's Methods Subcommittee

Preamble

The National Evaluation System of health Technology Coordinating Center (NESTcc) launched the Data Quality and Methods Subcommittees on August 24, 2018 to support the conducting of efficient, timely, and high-quality real-world evidence (RWE) studies for evaluating medical devices. The NESTcc Methods Subcommittee, consisting of a diverse range of stakeholders who each lend their unique methodological and industry expertise, advised the NESTcc Governing Committee and staff on constructs of study design and statistical methods. The role of 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" for NESTcc. This playbook was intended to highlight device-specific considerations in benefit/risk studies (both observational and randomized) as well as for safety signal detection. While this framework is closely linked to regulatory science, the principles described are applicable to any study intending to quantify cause and effect.

Subcommittee Members: Sharon-Lise Normand, PhD (Chair); Jesse Berlin, ScD; Mitchell Krucoff, MD; John Laschinger, MD; Heng Li, PhD; Nilsa Loyo-Berrios, PhD; MSc, Joao Monteiro, PhD; Didier Morel, PhD; Nilay Shah, PhD; Scott Snyder, PhD

Additional Contributors: Haley Abing; Tiffany Abushaikha, MS; Rachael Fleurence, PhD, MA, MSc; Jess Gasvoda, MPH; Robbert Zusterzeel, MD, PhD, MPH

To achieve this task, the NESTcc Methods Subcommittee developed a Protocol Framework, which builds upon existing bodies of work and leverages the subcommittee members' knowledge and experience from similar initiatives, including PCORnet, Sentinel, and the Medical Device Epidemiology Network (MDEpiNet). The document is intended to promote prospective study design – that is, pre-specification of as much detail as possible prior to data analysis to make clear what was and was not pre-specified. The Principle Investigator and the team members should work together to complete the study Protocol. Once undertaking analysis, any deviations from the Protocol should be reported and justified.

The information requested in the Protocol was developed over the course of several months **(Table 1)** with input from all subcommittee members who incorporated feedback from multiple rounds of comments and revisions.

Our efforts were in parallel and 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.

Table 1. Summary of meetings of the Methods Subcommittee	
Date	**Activity**
August 15, 2018	Data Quality and Methods Planning Committee Call
August 24, 2018	In-Person Meeting
Sept 17, 2018	Virtual Meeting
October 12, 2018	In-Person Meeting: Discuss Round 1 Protocol Assignments
December 10, 2018	Virtual Meeting
December 14, 2018	Round 1 Revisions DUE
December 19, 2018	Compiled/Revised Protocol with new Assignments
December 23, 2018	Round 2 Revisions DUE
January 9, 2019	Virtual Meeting
January 28, 2019	Virtual Meeting
February 13, 2019	Virtual Meeting
February 27, 2019	In-Person Meeting
April 22, 2019	Disseminated for Comments
May 3, 2019	Comments Received
May 7, 2019	Virtual Meeting
May 28, 2019	Posted for Public Comment

Introduction

A key task in planning a study, whether a randomized clinical 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, namely prior to statistical analyses. 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 2. Study design features specific medical devices are highlighted. The subcommittee developed this protocol template with focus on describing, at a high level, the key content relevant to each component of the protocol. The subcommittee's intention was to provide guidance on what is required to conduct a scientifically valid medical device study. The study protocol and corresponding statistical analysis plans should be completed (signed and dated) prior to commencement of data analyses.

Table 2: 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 treatment assignment
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 report provides some of the informing or consenting considerations but emphasize such ethical issues should be described in the protocol.	

The subcommittee recognized that different evidentiary requirements are needed based on the **stage of device development** (e.g., new device for new indication vs. existing approved device for indication expansion vs. surveillance of approved devices) and whether the device itself is new, iterative, or a second-generation device. Such diverse device assessments may require different study designs and endpoints. Study features specific to device evaluation at a particular stage of the device's lifecycle are thus also highlighted.

Medical devices are classified based on risk: Class I (minimal risk), Class II (moderate risk), and Class III (high risk or life-sustaining). Most Class I devices are exempt from Premarket Notification 510(k). Most Class II devices require submission of a 510(k) to demonstrate that the device **is at least as safe and effective** as **(substantially equivalent to)** a legally marketed device (predicate device) and hence not subject to a Premarket Approval Application. Finally, most Class III devices require submission of a Premarket Approval Application to demonstrate **reasonable assurance of safety and effectiveness** to meet statutory requirements. Some Class III devices are Humanitarian Device Exception (HDE) devices and are required to demonstrate reasonable assurance of safety and **probable** benefit. https://www.fda.gov/medicaldevices/deviceregulationandguidance/howtomarketyourdevice/premarketsubmissions/humanitariandeviceexemption/ucm563286.htm#HUD).

The subcommittee also recognized that virtually all of the components in the protocol apply to both randomized and non-randomized designs. For example, treatment assignment is made via randomization in experimental studies, whereas treatment is observed and not randomly assigned in non-randomized settings. In both instances, a description of the randomization process (randomized studies) and the estimated treatment assignment mechanism (non-randomized studies) should be pre-specified. Thus, this report does not provide separate principles for randomized trials and observational studies. Rather,

the subcommittee advocated pre-specification of study design features and of analytical strategies to minimize selective reporting of study results.

Guidance for specific features discussed in this report can be found at 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.

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

Introductory material presented in the protocol should include a thorough discussion of the underlying disease and available therapies sufficient to allow an understanding of the disease, the patient impact, and unmet medical needs, the device (including any 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 and indication for use, 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 from scheduled assessments and proposed endpoint definitions. 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 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 benefits and outcomes of each
    II. The strength of evidence supporting each
    III. The known risks of each
    IV. The rationale for selection of comparator therapy for the investigational protocol
    V. The therapeutic gaps or insufficiencies evident with current therapy – identification of an unmet clinical need
C. An assessment of the **underlying need** for the therapy proposed– why is the device needed and where does the device fit in:
    I. The physiologic rationale for development of the device
    II. The experience with existing cleared (e.g., predicate) or approved devices
    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 **literature, clinical experience or investigations**, relevant to the clinical study

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 (e.g., how a coronary stent, opening an infarcted artery, conveys benefit to a patient suffering acute myocardial infarction)

1.2 Specific Principles by Stage of Device

A. New/Early Stage Device
   I. Describe unmet need
   II. Justify initiation of a clinical trial due to the absence of prior clinical information or to limitations of existing pre-clinical data regarding the device's performance, safety, and benefits
   III. Describe early feasibility study results

B. Iterative/Late Stage Device & Indication Expansion
   I. Provide clinical discussion to justify use of performance goals or other historical controls
   II. Describe current device utilization including indications (on and off-label) and demographics if relevant
   III. Include clinical outcomes from other or prior devices with similar physiologic, anatomic, or mechanistic modes of action
   IV. Be clear about what the new device adds in terms of meeting previously unmet needs
      a. For example, for a new indication, describe if off-label use of the existing device has been observed
   V. Describe safety profile of device observed post-approval
   VI. Define and justify relevant surrogate endpoints

C. Surveillance
   I. Describe aspects of a device's safety and effectiveness that require investigation or monitoring after market introduction if a concern or doubt remains after approval of related studies
   II. Supply rational of what triggered the need for a surveillance study

1.3 References or Supporting Literature

1. Global Harmonization Task Force Study Group 5. "Clinical Evidence – Key Definitions and Concepts." 26 April, 2006, http://www.imdrf.org/docs/ghtf/archived/sg5/technical-docs/ghtf-sg5-n1r7-guidance-definitions-060426.pdf
2. Kramer DB, Tan YT, Sato C, Kesselheim AS. Postmarket surveillance of medical devices: a comparison of strategies in the US, EU, Japan, and China. PLoS Med. 2013;10(9):e1001519.

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.

2.1 General Principles to Follow

A. A description of the new 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, toxicity, sterilization, and manufacturing
   V. Sizing requirements 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 Specific Principles by Stage of Device

A. New Device/Early Stage:
   I. Describe the specific technical, structural, or procedural shortcomings of existing devices that are addressed by the new device
   II. Describe the context and what makes the device new
   III. Describe the need for operator training: 1) of didactic nature or 2) hands-on operator training/proctoring
   IV. Identify potential role of underlying patient or device factors impacting device performance

B. Iterative/Late Stage Device & Indication Expansion:
   I. Provide the specific technical, structural, or procedural shortcomings of existing devices that are addressed by the (new) iterative device
   II. Detail novel device design features with rationale for "iteration" rather than "new device" designation
   III. Identify the role of operator training and underlying patient or device characteristics in device safety

C. Surveillance:
   I. Indicate which features of the device(s) will be followed
   II. Provide biological plausibility of the life of the device
   III. Identify the primary device characteristic and rationale for studying it
   IV. Specify approach to capturing "unknown" unknowns

2.3 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." 20 December, 2017, https://www.fda.gov/ucm/groups/fdagov-public/@fdagov-meddev-gen/documents/document/ucm429672.pdf

3. Study-Specific Objectives

The protocol of a medical device study should contain unambiguous statements of its objectives aligned with its overall purpose (e.g., assessing the feasibility of the device, supporting a future premarket approval, expanding the indication of a predicate device, or post-market surveillance). The objectives must be relevant, specific, based on measurable quantities, and attainable within a reasonable time-frame **(Box 1)**. The objectives are typically organized by order of decreasing importance. A study objective may be operationalized by inclusion of the statistical hypotheses, although this is not necessary. A description of the key parameters of interest and basis for making conclusions, however, should be included. The choice of the primary objective is important and should be made explicit; secondary objectives should be identified as such.

3.1 General Principles to Follow

A. Define a general objective and derive several specific objectives (use the SMART terminology: Specific, Measurable, Attainable, Relevant, and Time-framed) that will be organized to:
    I. Show how the primary objective was chosen to provide the most straightforward, distinct clinical basis to formulate hypotheses
        a. If there are many primary objectives, justify each
    II. Include rationale for secondary objectives and describe how they are not directly linked to primary objective
    III. Specify, for devices consisting of multiple components (a "system"), if the system is the device being assessed or if a particular component is being assessed for each objective
B. Provide a precise description of the hypotheses or of the causal parameters for device effectiveness and device safety
    I. Precisely define the outcome measure(s) for each study objective, clinically meaningful effects in terms of risks relative to benefits
    II. For each outcome measure, include a precisely-defined causal parameter on which statistical inference is to be made (e.g., absolute difference, hazard ratio, etc.)
    III. If adopting a hypothesis testing approach, provide the mathematical expression for each hypothesis to be tested
    IV. If adopting an estimation approach, state how resulting estimates will be used to make causal inference and contribute to evidence-based decisions

3.2 Specific Principles by Stage of Device

A. Must clearly identify specific objectives for all device types, regardless of stage of development

3.3 References or Supporting Literature

1. Friedman LM, Furberg CD, DeMets SL. Fundamental of Clinical Trials. Springer, 1998.

BOX 1: 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 II*: 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: Intraoperatively (day 1)]) and several secondary outcome measures (Mesh contracture, mesh adhesion tenacity, via adhesiolysis time to abdominal wall, to mortality rate).

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** (e.g., the target population). Detailed inclusion and exclusion criteria 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. For example, if a national registry exists, randomization could be embedded in the national registry to achieve a prospective registry-based randomized clinical trial that may better reflect the target population. Alternatively, a local registry comprised of highly curated electronic health record data collected during routine clinical care could serve as the basis for an observational retrospective 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, etc.)
        a. Descriptors might include 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. pediatric, 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)

a. 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 particular 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 (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 source
      a. 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 protocol that this population of operators may not reflect the operators in broader practice (who would be using the device once it's marketed)
      a. 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. For observational studies utilizing a database or electronic health record (EHR), provide the name of the database and description of sampling frame; 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
   V. For example, how valid is the method for identifying patients with the condition of interest? How valid and granular is the approach to device identification? Specific steps include:
      a. Describe data sources, including linkages; for instance, if the data source is based on a Common Data Model (e.g., the Observational Medical Outcomes Partnership Common Data Model www.ohdsi.org), advantages and disadvantages of the data should be described
      b. Provide a high-level description of steps taken to assess data quality described in the NESTcc Data Quality Framework
      c. If available, include results of quantitative assessments of the reliability, sensitivity, specificity, and other features associated with the database
      d. If data linkages are performed, provide methods used (e.g., probabilistic or deterministic) and verification/validation planned
      e. If the data have been converted to a Common Data Model (e.g., the Observational Medical Outcomes Partnership Common Data Model), that should be specified
      f. Plans to account for quality issues such as sensitivity analyses

4.2 Specific Principles by Stage of Device

A. New/Early Stage Device:
    I. Describe how participating centers or entry criteria impact the composition of the study population relative to the target population
B. Iterative/Late Stage Device & Indication Expansion:
    I. If the iteration is intended to address a particular limitation of previous devices (e.g., a device is designed to accommodate larger lesions than previous devices), specify how patients in this expanded population will be identified in the data source or enrolled
C. Surveillance:
    I. Specify if and how "learning curve" effects of readers, operators, or surgeons impact the target population

4.3 References or Supporting Literature

1. "Premier Healthcare Database: Data that Informs and Performs." Premier Applied Sciences, the Research and Analytics Division of Premier Inc., July 29, 2018; https://www.premierinc.com/downloads/PremierHealthcareDatabaseWhitepaper.pdf

315

BOX 2 - EXAMPLE OF TARGET POPULATION: This study will use hospital billing records contained in the Premier Hospital Database (PHD). 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 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 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:

1. Underwent VATS lobectomy or laparoscopic right colectomy (elective or nonelective) during a hospital admission occurring between January 1, 2011 and September 30, 2016
    a. 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*
2. Aged ≥18 years or older at time of index hospital admission
3. 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:

1. Both powered and non-powered staplers used during the index hospital admission
2. da Vinci EndoWrist surgical staplers used during index admission
3. *Provisional exclusion criterion*: Non-specific (i.e., not identifiable with respect to powered vs. non-powered status or brand) staplers used during index admission
4. *Provisional exclusion criterion*: Evidence of robotics (laparoscopic right colectomies only; for laparoscopic non-robotics it is assumed that regardless of powered vs. nonpowered 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)
5. Point of origin or admission from another institution
6. *Provisional exclusion criterion*: Medicare Severity-Diagnosis Related Group which is not predominant in overall sample, not accounting for comorbidities and complications

316

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 1-year pain assessed by a questionnaire. Secondary outcomes provide additional information that are 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. For example, in low-risk devices, device performance may be sufficient to support a regulatory "tool claim," (e.g., a blood pressure cuff may accurately measure blood pressure independently of whether it is high, low or normal). 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 device performance may relate to biomaterials, design features, manufacturing tolerances, operator proficiency, patient selection criteria, anatomic variations, lesion variations or adjunctive therapies. Patient-driven outcomes may also be considered, and when appropriate, involving patients in identifying patient important outcomes.

In observational studies, pre-specification of a control outcome (e.g., an outcome unaffected by exposure, can strengthen the study design). While such outcomes cannot unequivocally prove the absence of bias in the association between exposure and study endpoint, it can test a putative mechanism of bias (Box 3).

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 outcome must be appropriate for desired instructions for use
    - a. Provide criteria for objective classification of the outcome
    - b. If endpoint adjudication is required, describe rules as well as number 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
- III. If using International Classification of Diseases (ICD codes), explain how the outcomes will be captured (algorithm), what codes will be used, describe sensitivity and specificity of

the ICD codes, whether the codes have been previously validated (e.g., are all potential cases with the outcome captured, do all identified cases have the outcome of interest)

IV. If using patient reported outcomes (PROs), describe the PRO instrument, describe how its validity will be evaluated (or has been evaluated)

V. If using patient generated data measured through devices (e.g., remote device monitoring, hemodynamic monitoring devices), describe both the internal and external validity of the data generated through such devices

VI. Specify the scales of each outcome (e.g., binary, failure time, categorical, etc.)

VII. Justify the use of surrogate outcomes and the use of composite outcomes

VIII. Specify and justify time points of data collection

IX. 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 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 why independent adjudication of whether adverse outcomes are "device related" is **not warranted**

D. **Control Outcomes** in observational studies:

I. Describe why the outcome is highly **unlikely** to be causally related to the device

II. Demonstrate that the suspected confounders of the association between the device and the potential control endpoint match those of the association between the device and the primary study endpoint

III. Analyze the association between the device and control endpoint according to the same procedure used to analyze the association between the device and the study primary outcome

IV. If more than one control outcome, describe why additional outcomes are needed

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 baseline outcome

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

5.2 Specific Principles by Stage of Device

A. New/Early Stage Device:

I. Provide rationale for selection of the primary outcome in light of risks

II. If a surrogate outcome is proposed, justification for using this type of outcome rather than a clinical outcome should be detailed

B. Iterative/Late Stage Device or Indication Expansion:

I. Provide justification if proposing new/different outcomes than those used in studies on first of a kind device

C. Surveillance:

I. Describe if safety concerns are related to specific modes of device failure (e.g., if a lead fractures in an ICD device, failure to shock the patient could result in death)

II. Describe why the need for serial monitoring over time of the device's anatomic position, electrical characteristics, or other performance attributes when assessing permanently implanted devices is unwarranted

5.3 References or Supporting Literature

1. Marc Lipsitch, Eric Tchetgen, and Ted Cohen (2010) Negative Controls: A Tool for Detecting Confounding and Bias in Observational Studies. Epidemiology Vol 21 No 3, 383-388.
2. Neil J. Wimmer et al. (2013) Comparison of Transradial Versus Transfemoral Percutaneous Coronary Intervention in Routine Practice. Journal of the American College of Cardiology Vol. 62, No. 22 2147-2150.

3. Neil J. Wimmer et al. (2016) Effectiveness of Arterial Closure Devices for Preventing Complications with Percutaneous Coronary Intervention: An Instrumental Variable Analysis. Circ Cardiovasc Interv. 9(4): e003464.
4. Prasad V, Jena AB. Prespecified Falsification Endpoints: Can they Validate True Observational Associations? JAMA 2013 ;309(3)241-242.
5. Rosenbaum PR. Design of Observational Studies. Springer Series in Statistics, Chapter 5, 2010; Springer, New York, NY
6. 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." 1 October, 2013; https://www.fda.gov/ucm/groups/fdagov-public/@fdagov-meddev-gen/documents/document/ucm279103.pdf
7. Velentgas P, Dreyer NA, Nourjah P, Smith SR, Torchia MM, eds. Developing a Protocol for Observational Comparative Effectiveness Research: A User's Guide. AHRQ Publication No. 12(13)-EHC099. Rockville, MD: Agency for Healthcare Research and Quality; January 2013. www.effectivehealthcare.ahrq.gov/Methods-OCER.cfm. (Chapt 6).

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 endpoint 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 control of confounding in the comparison with regard to the study primary endpoint.

6. Device Exposure

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 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. Specify the **brand and model number** of the device
    I. If more than one generation of the device is used, specify all models
    II. If Unique Device Identifiers are available in the data source, those should be used

B. Clearly identify the device being studied; for instance, is the focus on the main component or is it on the system?
C. 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
    I. For example, an induction (run-in) period of 2-months was planned in which insulin treatment was intensified with a standardized titration protocol, designed to achieve optimum injection treatment (Reznik et al. 2014)
D. 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
        a. For instance, if multiple stents are implanted in a single procedure in a single patient, will measurements be made for each patient-stent or for the first stent only?
E. Describe the precision with which exposure will be measured; this includes the data source, misclassification error, and measurement error
F. Describe the approach to confirming exposure to the investigational device
G. Identify specific clinical or surgical aspects that may narrow or broaden the definition of the exposure (e.g., anterior approach for hip replacement)
H. As noted in the section on Target Population, provide information on the training and experience of device operator/surgical team
I. 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 be impact the performance of the device being evaluated
    I. For instance, using an intra-arterial line during a procedure likely would not affect the performance of a coronary stent

6.2 Specific Principles by Stage of Device

A. New/Early Stage Device:
    I. Justify the duration of exposure based on the clinical objective and possible adverse events
B. Iterative/Late Stage or Indication Expansion Device:
    I. Provide rationale for exposure duration in relation to the new indication
C. Surveillance:
    I. Provide evidence that exposure duration is measured accurately

6.3 References or Supporting Literature

1. Velentgas P, Dreyer NA, Nourjah P, Smith SR, Torchia MM, eds. Developing a Protocol for Observational Comparative Effectiveness Research: A User's Guide. AHRQ Publication No. 12(13)-EHC099. Rockville, MD: Agency for Healthcare Research and Quality; January 2013. www.effectivehealthcare.ahrq.gov/Methods-OCER.cfm. (Chapt 4).

2. Yves Reznik, Ohad Cohen, Ronnie Aronson, Ignacio Conget, Sarah Runzis, Javier Castaneda, Scott W Lee - Insulin pump treatment compared with multiple daily injections for treatment of type 2 diabetes (OpT2mise): a randomized open-label controlled trial. Lancet 2014; 384: 1265–72.

7. Study Design

A study protocol should include a detailed description of the design features used to evaluate the medical device. Basic features needed include the number and type of comparison groups, blinding, outcomes (primary, secondary, procedural, device), the experimental unit of randomization, and how randomization will occur. Additional aspects associated with device evaluations relate to the effects of the device operator, the device procedure, and the complexity of the device should also be considered. Choice of the study design will depend upon the ability to minimize bias, ethical issues, practicality of executing, data quality, and data availability.

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.

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")
    I. Provide rationale for using randomization (controlled) or for not using randomization
B. Define the primary study objective (e.g., superiority, non-inferiority, equivalence)
C. 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

7.2 Specific Principles by Stage of Device

A. New/Early Stage Device:
    I. Some examples of controlled studies:
        a. Active Treatment comparator: A randomized study where the treatment arm is compared to those treated using the current standard of care; if using this design, indicate how the current standard of care arm will be described in order to interpret the effect size
        b. Subject as own control: Cross over study where the order of interventions (e.g., new device feature ON or OFF) will be randomly assigned for each patient; if using this design, justify that subjects will be unaware of which intervention they are currently receiving
B. Iterative/Late Stage or Indication Expansion Device:
    I. Some examples:

a. Objective Performance Criteria (OPC) or Performance Goal (PG): Single-arm study (subjects prospectively recruited) where the safety/effectiveness endpoints are compared to an OPC; the OPC could be a single number derived from historical data from clinical studies or registries; the use of historical versus contemporary comparison group requires justification
b. Historical Control Groups: Observational study where off-label use-cases found in a device registry (not prospectively recruited) are compared to a historical control group consisting of participants with on-label use; both arms should have been treated within the same time period, and ideally, within similar centers

C. Surveillance:
   I. Prospective observational 2-arm study where the safety/effectiveness endpoints are compared to other devices or interventions; choice of comparison devices/interventions requires justification

7.3 References or Supporting Literature

1. "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
2. Setoguchi S, Gerhard T. Comparator Selection. In: Velentgas P, Dreyer NA, Nourjah P, et al., editors. Developing a Protocol for Observational Comparative Effectiveness Research: A User's Guide. Rockville (MD): Agency for Healthcare Research and Quality (US); 2013 Jan. Chapter 5.
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." 7 November, 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://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E10/Step4/E10_Guideline.pdf

Blinding (Masking)

This may refer to the act of masking the treatment that a study subject received to participants, investigators, outcome assessors, and data analysts may all be blinded. It may also refer to the masking of outcome data to statisticians. To the extent possible, whether a randomized or observational study, blinding is encouraged.

7.4 General Principles to Follow

A. Who is blinded, when they are blinded, procedures used to blind, and when the blind will be broken should be precisely described

I. Rationale for lack of blinding of investigators, participants, outcome evaluators, or statisticians should be provided; other strategies to conceal treatment allocation should be described
II. Minimally, statisticians should remain blinded to patient outcomes with hypothesized endpoints until the Statistical Analysis Plan is completed and approved

B. Procedures used to maintain the blind should be included in the protocol

7.5 Specific Principles by Stage of Device

A. New/Early Stage Device:
I. A description of how blinding for all investigators, participants, etc. will be achieved should be included; if no blinding will be used, describe why this is not feasible for each person listed above

B. Iterative/Late Stage or Indication Expansion Device:
I. If no randomization, describe procedures in place to reduce selection biases

C. Surveillance:
I. If no randomization, describe procedures in place to reduce selection biases

7.6 References or Supporting Literature

1. Karanicolas, P., Farrokhyar, F., & Bhandari, M. (2010). Blinding: Who, what, when, why, how? Canadian Journal of Surgery, 53(5), 345–348.
2. Schulz, K.F., Grimes, D.A. (2002) Blinding in randomised trials: hiding who got what. Lancet, 359, 696-700.
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." 7 November, 2013, https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm373766.pdf

Units of Randomization and Observation

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 analysis 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.

7.7 General Principles to Follow

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

7.8 Specific Principles by Stage of Device

A. Surveillance:
    I. A clear and objective description of how the unit of observation will be determined in an observational study should be provided

7.9 References or Supporting Literature

1. Rosenberger WF, Lachin JM. Randomization in Clinical Trials: Theory and Practice, Edition 1. Wiley, 2002.

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.

7.10 General Principles to Follow

A. Characterize and justify the treatment assignment mechanism when the assignment is known (randomization) 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 & 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 & justify the length
    VIII. Provide an accounting of the number of participants: approached, eligible, provided consent, and randomized
B. Characterize the treatment assignment mechanism when the assignment is unknown (observational study) including:
    I. Describe variables that will be used to estimate the treatment assignment mechanism (e.g., the propensity score)

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., what size caliper used for matching, what size weights to 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

7.11 Specific Principles by Stage of Device

A. New/Early Stage Device:
   I. A description of the method to protect against guessing treatment assignment should be provided (e.g., permuted block randomization, adaptive randomization, etc.)
   II. If randomization is not stratified by center, a clear rationale should be included
B. Surveillance:
   I. Blinding or separation of outcome by treatment arm to all investigators is particularly important in surveillance settings where randomization does not occur; a description of how this will be achieved should be included
   II. Describe how the treatment assignment mechanism will work when competing products enter the market while assessing a medical device

7.12 References or Supporting Literature

1. Götberg, M. et al. (2017) Instantaneous Wave-free Ratio versus Fractional Flow Reserve to Guide PCI. NEJM vol 376 no 19 1813-23.
2. The Central Role of the Propensity Score in Observational Studies for Causal Effects. Paul R. Rosenbaum and Donald B. Rubin Biometrika Vol. 70, No. 1 (Apr., 1983), pp. 41-55
3. Sunil V. Rao et al. (2014) A Registry-Based Randomized Trial Comparing Radial and Femoral Approaches in Women Undergoing Percutaneous Coronary Intervention. JACC Cardiovasc Interv vol 7, no 8 857-67.
4. Wittes, J. (2005). Randomized Treatment Assignment. In Encyclopedia of Biostatistics (eds P. Armitage and T. Colton). doi:10.1002/0470011815.b2a01050

Other Covariates

These may be of interest in some designs.

7.13 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)
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. Characterization of the completeness, quality, validity, and replicability of the covariates

7.14 References or Supporting Literature

1. Committee for Medicinal Products for Human Use (CHMP). "Guideline for adjustment of baseline covariates." European Medicines Agency, 26 April 2013. https://www.ema.europa.eu/documents/scientific-guideline/draft-guideline-adjustment-baseline-covariates_en.pdf
2. US Food and Drug Administration. "Evaluation of Sex-Specific Data in Medical Device Clinical Studies. Guidance for Industry, Food and Drug Administration Staff." 22 August, 2014, https://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM283707.pdf
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." 12 September, 2017, https://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM507278.pdf
4. Velentgas P, Dreyer NA, Nourjah P, Smith SR, Torchia MM, eds. Developing a Protocol for Observational Comparative Effectiveness Research: A User's Guide. AHRQ Publication No. 12(13)-EHC099. Rockville, MD: Agency for Healthcare Research and Quality; January 2013. www.effectivehealthcare.ahrq.gov/Methods-OCER.cfm.

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, 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.

Consent

Consent involves informing the patient or study participant what the study involves, why it is important, what is required of the participant, 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.

8.1 General Principles to Follow

A. If no consent is required, provide rationale and supporting documents
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
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

A. "How To Consent." UCI Office of Research, The Regents of the University of California, 2019, https://research.uci.edu/compliance/human-research-protections/researchers/how-to-consent.html

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 cross-overs) or participant withdrawal
C. Because study withdrawal and non-compliance are separate mechanisms, distinct approaches to minimizing both should be included
D. Provide procedures to minimize the number of assessments made outside of a follow-up interval

8.4 References or Supporting Literature

1. Mohan, Sandy, et al. "A Toolkit for the Management of Protocol Deviations." Therapeutic Innovation & Regulatory Science, vol. 50, no. 6, Nov. 2016, pp. 791–800, doi:10.1177/2168479016647987.

9. Required sample size

The determination of sample size is a critical component of the design of a clinical study (Box 4). If the sample size is too small, firm conclusions are unlikely to be inferred or results might have been obtained by chance. On the other hand, an excessively large sample size would be wasteful and unethical. In practice, the study sample size is determined based on a number of design parameters and following 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 sample size. 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)
    III. Adaptive (see interim analysis and stopping rule topic)
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 1-sided versus 2-sided confidence intervals (or 1-sided vs 2-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 Specific Principles by Stage of Device

A. **New/Early Stage Device:**
    I. While the device is new, the clinical condition is likely not new; thus, provide outcome rates associated with the condition as described in the available literature
B. **Surveillance**:
    I. Specify the frequency and duration of assessments (link to adaptive design)
    II. Describe and justify the basis for selection of effect size for “alerts” and for “warning” about potential safety signals
    III. If the study is proposing larger effect sizes than has been observed for other devices, provide a clear rationale for the effect sizes

9.3 References or Supporting Literature

1. 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.

2. 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/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm073137.pdf
3. 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." 27 July, 2017, https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm513027.pdf
4. US Food and Drug Administration. "Guideline for Industry: Structure and Content of Clinical Study Reports, ICH E3." ICH Expert Working Group, July 1996; https://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm073113.pdf
5. Velentgas P, Dreyer NA, Nourjah P, Smith SR, Torchia MM, eds. Developing a Protocol for Observational Comparative Effectiveness Research: A User's Guide. AHRQ Publication No. 12(13)-EHC099. Rockville, MD: Agency for Healthcare Research and Quality; January 2013. www.effectivehealthcare.ahrq.gov/Methods-OCER.cfm. (Chapt 9).

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: Insulin Pen Needles: Effects of Extra-Thin Wall Needle Technology on Preference, Confidence, and Other Patient Ratings Aronson, Ronnie et al. Clinical Therapeutics, Volume 35, Issue 7, 923 - 933.e4

10. Study Registration

Registration of randomized trials is standard practice and is required by publication policies at major journals and by governmental regulations (see references 2 and 3 below). 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. There is less agreement on the value of registering observational study protocols although pre-specification is likely to enhance reproducibility. Therefore, while registration of randomized trials on sites such as clinicaltrials.gov is mandatory, pre-specification and publication for all studies is strongly encouraged.

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
    I. Registration of observational studies has been more controversial (see references 4-11 below)
    II. Rather than adopt a dichotomous view of registering all observational studies or none, registration of selected observational studies in a publicly accessible repository like clinicaltrials.gov will make the best evidence available, assure a high degree of transparency, and reduce ethical questions of conflict of interest

10.2 References or Supporting Literature

1. Berger ML, 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 2009; 12(8): p. 1044–1052.
2. Chavers LS, Fife D., Wacholtz M, Stang P, Berlin JA. Registration of Observational Studies: perspectives from an industry-based Epidemiology group Pharmacoepidemiology and Drug Safety 2011; 20 (10):1009-13.
3. ClinicalTrials.gov. ClinicalTrials.gov. 2010 [cited 2010 September]; Available from: www.clinicaltrials.gov.
4. Editors. The registration of observational studies--when metaphors go bad. Epidemiology 2010; 21(5): p. 607–609.
5. Epidemiology AC. Sense of the College: Should Observational Studies be registered? in 2010 American College of Epidemiology: San Francisco, 2010.
6. Feinstein AR. Scientific standards in epidemiologic studies of the menace of daily life. Science 1988; 242(4883): p. 1257–1263.
7. Gliklich RE DN. Registries for Evaluating Patient Outcomes: A User's Guide. (Prepared by Putcome DEcIDE Center [Outcome Sciences, Inc. dba Outcome] under Contract No. HHSA290200535I TO1.) 2007, Agency for Healthcare Research and Quality: Rockville, MD.
8. ICMJE. http://www.icmje.org/publishing_10register.html. 2010.
9. ISPE. Guidelines for good pharmacoepidemiology practices (GPP). Pharmacoepidemiology and Drug Safety 2008; 17(2): p. 200–208.
10. ISPOR. ISPOR Good Outcomes Research Practices Index. ISPOR Good Research Practices 2011 [cited 2011 June 10, 2011]; Available from: http://www.ispor.org/workpaper/practices_index.asp.
11. Lash TL. Preregistration of study protocols is unlikely to improve the yield from our science, but other strategies might. Epidemiology 2010; 21(5): p. 612–613.
12. Motheral B, et al. A checklist for retrospective database studies--report of the ISPOR Task Force on Retrospective Databases. Value in Health 2003; 6(2): p. 90–97.
13. Poole C. A vision of accessible epidemiology. Epidemiology 2010; 21(5): p. 616–618.
14. Samet JM, To register or not to register. Epidemiology 2010; 21(5): p. 610–611.

15. Takkouche B and Norman G. Meta-analysis protocol registration: sed quis custodiet ipsos custodes? [but who will guard the guardians?]. Epidemiology 2010; 21(5): p. 614–615.
16. Vandenbroucke JP. Preregistration of epidemiologic studies: an ill-founded mix of ideas. Epidemiology 2010; 21(5): p. 619–620.

11. Monitoring Plan

Monitoring clinical investigations **(Box 5)** is essential not only for the protection of human subjects, but also for the conducting 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 standards. Reasons for study monitoring include protocol compliance, adverse effects, treatment comparisons to stop trial (early if needed), data management to identify data errors or missingness, and study futility. Use of an independent Data Safety Monitoring allows confidential access to treatment-related bias and may not only ensure human subject safety but also reduce bias in study management.

11.1 General Principles to Follow

A. Monitoring Committees
   I. Describe the charge of the data safety monitoring committee, members and their expertise, frequency of meetings, and procedures
   II. Describe the process for data quality monitoring including members and how data issues will be resolved
   III. 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)

B. Interim Analyses
   I. Define operational committee interpreting interim analyses (Steering Committee, Data Safety Committee, etc.)
   II. Define purpose of any interim analyses (for early stopping for futility, for efficacy, for safety, for adaptive designs, or potential mid-course corrections)
   III. Describe and justify number and frequency of analyses
   IV. Provide a description of the stopping rule
   V. If interim analyses for surveillance studies are planned to be released, describe when these analyses will be conducted and what directive language will accompany the release
      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
   VI. 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 Specific Principles by Stage of Device

A. Surveillance:
    I. If pre-planned, surveillance data can be used to reduce sample size needed in pre-market studies (see FDA guidance on balancing both types of study)
        a. For instance, if a development program is likely to leave residual uncertainty with respect to safety (or effectiveness in actual practice), the requirements for pre-approval data might be reduced, in discussions with FDA early in the program, when there is a valid post-approval study planned during the pre-approval period
B. The post-approval data are then used to "offset" the uncertainty remaining at the time of approval

11.3 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; 39:384–389.
2. Chang, M. Introductory Adaptive Trial Designs: A Practical Guide with R, 2015.
3. Chow, SC and M Chang. Adaptive Design Methods in Clinical Trials, CRC Press, 2011.
4. Johnston KC, Hall CE, Kissela BM, Bleck TP, Conaway MR. Glucose regulation in acute stroke patients (GRASP) trial: A randomized pilot trial. Stroke. 2009; 40:3804–3809.
5. Mehta, C. R., and Pocock, S. J. Adaptive Increase in Sample Size when Interim Results are Promising: A Practical Guide with Examples. Statistics in Medicine, 2011; 30, 3267-3284.
6. The National Institute of Neurological Disorders and Stroke rt-PA Stroke Study Group. Tissue plasminogen activator for acute ischemic stroke. N Engl J Med. 1995; 333:1581–1587.
7. O'Brien, P.C., and Fleming, T.R. A multiple testing procedure for clinical trials. Biometrics, 1979; 35 549-556.
8. Palesch YY, Hill MD, Ryckborst KJ, Tamariz D, Ginsberg MD. The ALIAS pilot trial: A dose escalation and safety study of albumin therapy for acute ischemic stroke--II: Neurologic outcome and efficacy analysis. Stroke. 2006; 37:2107–2114.
9. Pocock, S.J. Group sequential methods in the design and analysis of clinical trials. Biometrika, 1977; 64, 191-199.
10. 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
11. 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." 7 November, 2013, https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm373766.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." 13 April, 2015,

https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm393994.pdf

13. US Food and Drug Administration. "Adaptive Designs for Medical Device Clinical Studies. Guidance for Industry and Food and Drug Administration Staff." 27 July, 2016, https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm446729.pdf
14. US Food and Drug Administration. "Adaptive Designs for Clinical Trials of Drugs and Biologics. Guidance for Industry." September 2018, https://www.fda.gov/downloads/drugs/guidances/ucm201790.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, V. L., Hall, C. E., et al. (2014). 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*, 9(2), 246–251.

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 is often less detailed than the final SAP. It must be approved prior to any analyses, and sometimes before the first patient is enrolled/randomized.

12.1 General Principles to Follow

A. Definition and justification of target population and study samples
    I. ITT sample (effectiveness)
    II. Safety Sample (safety)
B. Indicate the treatment of missing data, associated assumptions, and how validated
C. Definition/description of computation of derived variables
D. Define study success criteria
E. Provide statistical models and test for analyses of:

I. Primary, secondary, procedural, device, and safety outcomes
II. Interim analyses
III. Subgroup analyses

F. Provide a plan for adjustment for multiplicity of all endpoints, with possibly the exception of safety endpoints
G. Describe sensitivity analyses including the feature addressed and assumptions made
H. Provide example of tables and graphs
I. Describe and justify Interim analysis plan and impact on statistical design (type one error spending function, similar to previous section)
J. Pre-specify how learning curve effects will be handled

12.2 Specific Principles by Stage of Device

A. Late Stage or Surveillance:
I. If using an observational study, causal inference approach should be justified including
II. Choice of approach (modeling the treatment assignment mechanism only versus modeling both the treatment assignment mechanism and the outcome model)
III. Strategies to mitigate selection bias such as the use of machine learning approaches to condition on many potential confounders
IV. Strategies to minimize selective inference
a. For example, modeling the treatment assignment mechanism without the outcome or controlled procedures (not permitting the outcome to influence the treatment assignment mechanism) to model both treatment and outcome (e.g., machine learning)

12.3 References or Supporting Literature

1. Guidelines for the Content of Statistical Analysis Plans in Clinical Trials Gamble et al. JAMA. 2017;318(23):2337-2343.
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. London, England: European Medicines Agency; 1998
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