A Framework for Evidence Evaluation and Methodological Issues in Implantable Device Studies

Art Sedrakyan, MD, PhD,* Danica Marinac-Dabic, MD, PhD,* Sharon-Lise T. Normand, PhD,† Alvin Mushlin, MD, ScM,‡ and Tom Gross, MD, MPH*

Abstract: Implantable medical devices (IMD) are frequently used in interventional medicine. There are a host of complex methodological issues to consider in conducting device studies. A general conceptual framework for evidence evaluation is needed to help investigators conduct comparative studies in this setting. It is known that clinical trials of implants require study design planning and creative execution that are quite different from those in pharmaceutical setting. Important study design issues such as randomization, masking and allocation concealment require unique approaches for each device. In addition, device comparative studies must cope with sources of variability different from pharmaceutical studies. These include operator learning curve effects, hospital-operator-patient interactions, and issues related to device technical characteristics. Observational studies of IMDs are particularly challenging. Selection of comparison groups, adjusting for confounding and addressing learning curve issues needs careful planning. We propose a general framework for IMD evaluation and provide an outline of the methodological issues that require further discussion. We hope this article will inspire and help to inform those interested in advancing comparative safety and effectiveness of IMDs and to plan and pursue future methodological work in this area.

Key Words: comparative studies, medical devices, methodology

Implantable medical devices, ranging from intraocular lenses to coronary stents, are now ubiquitous inside and outside clinical practice. Their use and complexity are increasing and tens of thousands of models are currently on the market in the United States. Indeed, the introduction of some devices may be transformative of clinical practice.1 Moreover, the public health burden posed by device-related adverse events and product problems in general may be significant, as evidenced by visits to emergency departments,2 in-hospital events,3 and reports through Food and Drug Administration’s nationwide passive surveillance system.4

For implantable medical devices (IMDs), the impact on public health is at least partially related to the technical complexity of the device combined with operator training demands, both adding to heterogeneity in clinical safety and effectiveness. Furthermore, it is recognized that the operator (user, interventionist) error, in addition to device design, plays an important role in device safety and effectiveness,5 and these combined sources of variation are part of the “device-operator interaction.” The device-operator interaction relevance to IMDs is in contrast to pharmaceutilicals where safety and effectiveness are more related to the systemic effects of the chemical compound itself.6

The evidence generation for IMD outcome improvement is planned in an environment often quite different than that for drug evaluations. For example, clinical trials of IMDs are conducted in a relatively unique environment defined by highly skilled operators and by high-volume implanting centers. In addition, the trial is methodologically challenging; the fundamental concepts of randomization, blinding, and allocation concealment are not easily implementable. The design of the trial requires care and innovative approaches.7 Observational studies of IMDs have their own unique challenges when compared with drugs. The key differences between device and drug studies are related to the interaction between the device, operator, and the interventional area in which it is used. Device mechanical characteristics, technical complexity, imaging, and clinical detail, such as anatomic lesions, are important for outcomes evaluations but often only partial information about these variables is available through commonly used data sources, such as clinical registries maintained by professional societies or administrative data sources. Additionally, data about ancillary care provided in the hospital and operating room is important but often not available.

In this article we provide a framework for, and discussion of, key factors that impact IMD safety and effectiveness. An additional goal is to highlight important methodological challenges in studying outcomes of IMDs, ranging from the
domain of investigational device clinical trials to postmarket “real-world” observational studies. Future articles will offer case studies and detailed discussion of relevant methodological issues.

THE FRAMEWORK

We adhere to an evidence-based medicine perspective which can be described as comprehensive, thorough, explicit and careful development and assessment of best evidence to make decisions in healthcare.8,9 Important factors that need consideration in designing studies and assessing the outcomes of IMDs are depicted on Figure 1. The figure separately depicts key factors in the randomized clinical trial (RCT) and observational study settings and we show that evidence from the initial studies (particularly when part of premarket application) has a very important impact on the adoption and real-world use of IMDs (both on-label or off-label). We describe the importance of methodological validity/quality of the studies, enrolled patient population, surgical characteristics, surgery-related factors, and hospital characteristics.

Selection of Comparator

The choice of the comparison group has critical implications and introduces particularly challenging issues if the comparison therapy is optimal medical management rather than another device. When device is compared with optimal medical management rather than another device, blinding and allocation concealment issues will be more pronounced in the RCT setting (see methodological challenges section).

Blinding and allocation concealment is not possible in the observational setting and more detailed data on factors that drive therapeutic choices (both physician and patient), cross-over and refusals are needed to design a high quality study. This is particularly important when a device is compared with best medical management. Additionally, because of surgical risks associated with the
implant, the device-to-optimal management comparative studies will require long-term follow-up to not focus on surgically related complications.

Characteristics of the Device
In the initial preclinical stage, detailed information is collected on IMDs, from mechanical to electrical to software properties and biomaterial effects, to characterize their safety and effectiveness. Device characteristics, among others, include mechanical performance and reliability of the device. For example, reliability assessments range from battery life duration to physical wear of the implant. Biomechanical preclinical testing involves identifying variables associated with device failure. For example, component position and size on the range of motion to impingement of hip resurfacing systems have been shown to result in varying rates of displacement, wear, and failure. In addition, the fact that IMDs, unlike drugs, continually evolve (even during the study) over time makes it more challenging to accurately assess safety and effectiveness.

Characteristics of the Operator
The operator’s interpretation of evidence and preference are critical factors for adopting a new technology in the real world. The operator’s skill and learning curve are factors intertwined with evidence for adoption and use particularly if the technology is challenging and there is a protracted learning curve. For example, learning how to implant left ventricular assist device may require more time to master and hence might not be adopted as fast as coronary stents where the learning curve is steep. In many large trials of IMDs, procedures are conducted by highly skilled operator-enthusiasts of the procedure. This also has important implications for designing studies and evaluating the evidence in real-world settings where user sophistication may vary greatly. Lastly, operator preferences may be driven, in part, by a variety of other factors including training, institutional or professional society standards, reimbursement considerations, and formulary decisions.

Characteristics of the Procedure/Intervention
Care has to be taken to understand, and take into account, various procedural factors related to IMDs in addition to the device itself and the operator. Documentation of surgical techniques/approaches, application of various imaging modalities, and areas treated (eg, in the case of transmyocardial revascularization or cardiac ablation) can be important to proper interpretation of relevant outcomes. In addition, it is important to recognize and account for procedural choices, and invasiveness of treatment, that are driven by disease severity.10

Characteristics of the Intervention Setting
Hospital and access factors are important to evaluate and determine the groups of patients that are able to receive IMDs (more applicable to observational evaluations). In this context insurance, geographic location and distance to provider are all important. Reasons for offering a particular implantable device technology should also be understood. There might be financial and organizational reasons behind the hospital’s choice to adopt a certain procedure. This can be related to its size, financial pressures, and other issues that may encourage them to offer something new (eg, perhaps to stay competitive).11

Characteristics of the Patient
From the initial trials the clinical community learns about specific patient populations that may or may not benefit from the technology. Thus, from study to study very specific and selected groups of patients who were “responsive” to therapy are enrolled into future clinical trials. However, when the device enters the market and gets adopted in real-world settings the pool of patients who receive these devices is much larger than those who were “responsive” to therapy in the trials and in some instances most of the use becomes “off-label.”12 The potential impact of this kind of diffusion of new technology might be substantial on outcomes of the IMD. Patient related factors such as comorbidities and disease severity are certainly part of the framework and will impact the use of IMDs, particularly when the alternative is more invasive surgery. For new device intervention, which is similar to existing device intervention in its invasiveness, once operators master the IMD’s technical and management procedures, the impact of disease severity on the use of the device becomes more limited. Of note is that real-world use may create substantial difficulties in real-world evaluations of IMDs as patient selection factors are not always easily collected or collectable in clinical databases and registries. Furthermore, certain patient outcomes important to the assessment of safety and effectiveness, such as quality of life measures, may not be systematically collected.

Selection and Measurement of Outcome
The choice of outcomes can substantially affect the strength of the evidence and adoption/real-world use of the IMDs. Clinically meaningful outcomes are preferred in the IMD studies. In some instances, the outcomes are surrogate endpoints, such as changes in inflammatory markers, and strong conclusions based on these surrogate endpoints can be very challenging. In other instances, due to small sample sizes, the outcomes are combined into indexes such as “combined critical end point” or “composite end point.” This can be misleading in some situations as it dilutes the effect of outcomes that move in opposite directions or in case of rare but catastrophic events. Further, the events are not necessarily related to each other physiologically or as consequences of a given IMD use. The use of disease-specific measures of health status via patient or physician administered questionnaires will likely become quite important as medical options, including devices, expand. However, with the challenges in patient and physician blinding in IMD assessments, additional design and analytical issues need to be in place to reduce bias in health status reports. In the observational studies that rely on administrative data there are even greater challenges such as lack of information on quality of life or functional status and use of codes which require validation to assess potential misclassification.
METHODOLOGICAL ISSUES RELATED TO COMPARATIVE STUDIES OF IMDS

Many methodological validity/quality components developed for pharmaceuticals do not necessarily apply to implantable devices. In addition, learning curve, small study size, and device modifications during the study are challenges that are relatively specific to IMDS as compared with pharmaceutical studies. We discuss major challenges related to device interventions in the following sections.

Blinding or Masking Is Difficult to Achieve

Double and sometimes even single masking/blinding does not apply to many RCT based IMD evaluations particularly when the intervention/device is compared with non-intervention/device-based therapy. While this problem is in many instances impossible or unethical to overcome (ie, sham surgery) there are methods that the investigator can use to diminish the effect of an inability to mask (blind) patients and operators. One way to do this in the RCTs is the application of robust allocation concealment. Allocation concealment is the process by which the treatment allocation sequence is implemented to keep clinicians (particularly executors of the study) and participants unaware of upcoming assignments. To accomplish this, use of centralized telephone systems and announcement of the allocation in the intervention room just before the surgery can help conceal the allocation of the device intervention from the operator and the patient, thus minimizing the likelihood of operator and patient related refusal to undergo initially assigned device intervention. Another method for reducing the impact of inability to mask operators and patients is masking outcome assessors that can be implemented in both RCTs and observational studies.

Inclusion/Exclusion Is Affected After Initial Intent/Randomization and Impacts Analytical Approach Including Intention-to-Treat Analysis

An intention-to-treat analysis is the corner stone in the analysis of RCTs. It refers to the comparison of outcomes among patient groups defined by the initial intent/randomization mechanism. It is an important component of methodological validity occasionally misinterpreted by some operators. Data analysis should be based on the initial treatment intent, not on the treatment eventually administered. In the context of device trials this applies to randomized patients who converted (cross-over) from one type of device to another either before or during the intervention.

For example, if the conversion occurs before the start of the intervention this might affect originally planned inclusion/exclusion criteria as well as intention-to-treat analyses. On the other hand if the conversion occurs during intervention it is mostly related to intention-to-treat analyses. An example of the former is a case of surgical inspection after surgical incision and inability to perform particular device implantation. In this situation original plan to deliver specific device is affected and this has an impact on the clinical equipoise and inclusion/exclusions that were defined in advance. An example of the latter is a deployment of an implant and due to unsatisfactory results use of the comparator device to “correct” the problem. The information on many of these conversions is often missing, excluded, or analyzed using as-treated rather than intention-to-treat. Thus, intention-to-treat analyses may have a different meaning in the context of a drug evaluation versus the evaluation of a device/surgery. In drug evaluation, cross-over in many instances is informative; it may, for example, occur as a result of ineffectiveness or side effects of the drug. Thus, not applying intention-to-treat analyses will bias the results. In IMD evaluations, cross-over might occur not as results of “side effects” or a “problem” but also after surgical “inspection” and inability to conduct the assigned intervention due to specific anatomic or pathologic findings on the operating table. Certainly, understanding this problem and assessing the applicability of this important methodological validity factor in IMD evaluations will help to minimize the risk of bias. The investigators need to carefully discuss intention-to-treat analysis plan in the study protocol and outline the solution based on the uniqueness of the context and the device.

Of note is that in most observational studies information on cross-over is missing and intention to treat analysis is virtually impossible. This potential bias should be part of the discussion/limitations in any observational study.

Learning Curve Issues Affect the Outcomes of the IMD

The operator’s learning curve can be steep, protracted, or anywhere in between and has a substantial impact on the outcomes. One can design a study to address the learning curve by embarking on a clinical trial of a device only after mastering the IMD procedures. In the experimental setting, one solution proposed involves conducting an “expertise-based” clinical trial. In this design, operators more experienced with a particular arm of the study conduct the procedure for that arm and vice versa. In this way, effects due to expertise with the device can be substantially reduced. However, because different sets of operators are involved in each arm of the trial, confounding due to operator has been introduced. Moreover, it will not be clear if all operators can master the procedure or what the actual learning curve would look like. Given that elite surgeons participate in most of the initial trials, the impact of learning curve, its shape, and interaction with other variables are difficult to evaluate in the RCTs of IMDS. These factors can only be evaluated in real-world observational studies such as registry-based studies.

The training and certification in IMD use are critically important as there may be a significant learning curve associated with use of certain IMDS. The complexity of the learning curve should not be used to justify rejection of the technology and the learning curve should be studied as it relates to the application and dissemination of evidence in comparative studies. Most of the initial IMD procedures in real-world are conducted by “fast learner” elite surgeons and enthusiasts of the procedure. It is not clear if all operators can safely use a new device and what amount of training is enough to ensure safe application of the technology. Varying levels of training/experience may be invoked depending upon the subspecialty of the operator doing the procedure. This has
been the case with stenting of carotid arteries by operators from varying specialties (radiologists, cardiologists, and neurosurgeons) that establish thresholds for proficiency based on background expertise related to their specialties.18

Traditionally, the learning curve is studied using volume-outcome relationship. Volume-outcome studies have demonstrated that increased surgical volume has an inverse relationship with the likelihood of poor outcomes such as complications, revision surgery, length of stay, and mortality.19–21 Some other studies have shown a volume threshold for procedures above which increasing volume is no longer associated with improved outcomes.22,23 Others have noted a tri-modal institutional learning curve (rapid initial phase, followed by declining success—representing new adopters, and then recovery to an improved steady state. Thus, in our opinion, there are 3 components related to volume-outcome relationship: (1) lifetime experience (operator’s volume), (2) operator’s annual volume, (3) hospital volume where operators practice. Most of the published literature addresses the annual and hospital volumes, and there is limited evidence on the impact of lifetime experience (volume) on the outcomes. We need more insightful ways to evaluate operator learning curve since accumulated knowledge is clearly limited in the IMD setting. Other factors aside from volume, such as diagnostic versus interventional procedures and institutional teaching status, need further assessment as well.

Small Sample Size Is a Reality in IMD Studies

A large adequately powered study of IMD to evaluate all important clinical outcomes would be ideal, but it is not a realistic scenario in most instances for several reasons. First is the speed at which the technology of the IMD changes. Second, the cost of the IMDs to the patient, as well as other trial-related costs is covered by the company and this can be burdensome for many relatively small companies that produce IMDs. Additionally, the trials are typically powered first and foremost for effectiveness outcomes, often related to mechanical and engineering aspects of the device, and more common clinical safety outcomes.

However, investigators need to make sure that even in small studies they carefully report all important clinical outcomes.

Methods that enhance the ability to precisely to estimate the safety and effectiveness of IMDs in the absence of a large number of participants are desirable. In some instances Bayesian methods and adaptive clinical designs might help to overcome the limitations related to small studies by incorporating prior information. Bayesian methods may help to bring many innovative devices to market sooner by formally including prior evidence and reducing the need to accumulate large amount of data.24 More research is needed to fully appreciate the advantages and limitations of these methods.

Systematic reviews of RCTs are frequently used to evaluate safety and effectiveness. Most individual studies of IMDs are underpowered to determine differences in clinical outcomes but often carefully record all clinical outcomes. This offers a great opportunity to conduct evidence synthesis when a number of studies are available. While there is a potential that some imbalance of patient characteristics will occur in individual studies due to small sample size, it is unlikely that imbalance will systematically favor one group or another, as these imbalances should cancel out when combining many underpowered trials.25 There are 2 important caveats with meta-analyses however. First, a meta-analysis is an observational study of studies, so that the usual threats to validity exist. Second, subgroup analyses to explore heterogeneity in treatment effects are not always considered (or possible to do) but they might explain substantial heterogeneity in effects such as those related to device modification or practice change/patient selection. Third, and importantly for comparative research, it is risky to rely upon clinical trial populations to provide a complete understanding of safety and effectiveness in unselected real-world populations.

IMDs in a Changing Environment: Device Modifications and Changes to Medical Practice During a Study

Device development is a very dynamic process resulting in frequent modifications of the existing device based on experience with the device. Such modification, if small, can be made after the completion of a study or they can sometimes even occur during the course of the study. For example, the angle of the struts in a coronary stent could be changed to reduce adverse events or the stent deployment device could be modified for better delivery. In these situations, the impact of the device change may have a somewhat predictable impact on outcomes. In other situations the impact of a device modification is not as clear and warrants further scrutiny and careful follow-up. Therefore, progressive device modifications over a particular time period have implications for comparative studies in that it may or may not be acceptable to pool patient-level data directly across pre- and postmodification time periods. In addition, given that IMDs are generally long-term implants; consequential device modifications of IMDs may raise issues of the direct relevance of findings from prior versions to ones currently implanted.

The Potential for Imbalances Among Treatment Groups Must Be Taken Into Account

Because of the multiple sources of factors impacting outcomes of IMDs, both randomized and nonrandomized designs must adjust for numerous factors. In RCTs, any imbalances are by chance alone and this topic is beyond the challenges discussed in this paper.

In observational setting, cohort designs using a concurrent comparison group are preferred. In the ideal situation there is prospective and consecutive patient enrollment and data collection and the study is hypothesis driven. However, most of the large comparative studies of IMDs are based on registry data. These data, while prospectively collected, are not necessarily collected with comparative questions in mind. In addition, enrollment is on a voluntary basis.

As the assignment of the IMD is not random but rather left to the judgment of the operator there is a concern about confounding by indication— that is, the operator makes
choices based on disease severity, frailty or other patient characteristics. One strategy to overcome the latter limitation is to exploit the natural variation in care in observational studies. Choice of therapies (including IMDs) is more likely to be related to preferences and training of individual operators and institutions rather than patient, disease-related factors, or the facilities available. There is substantial evidence in the literature that willingness of the operator to provide a procedure, rather than its appropriateness for the patient, may explain substantial variation in practice.\textsuperscript{26,27} One can take advantage of his natural phenomenon in designing a comparative cohort study and at least partially address the confounding by “indication.” If the selection of the IMD is clear at the extremes (based on very strong indication for one strategy over the other), in the middle range there is a substantial overlap between choices. While the assignment is still not random, robust study design based on overlap between choices and statistical analyses may help to reach valid conclusions about safety and effectiveness.

Adjustment for known and measured confounders is a research area receiving increasing attention. Most of the adjustment techniques deal with imbalances in prognostic factors between the study groups. As we discussed previously, the comparison of various IMDs is possible by taking advantage of natural variation in care. Only after clinical judgments are made about the inclusion and exclusion criteria can these methods be applied to deal with imbalances in prognostic variables. All available statistical methods that deal with these imbalances through some type of adjustment are only able to eliminate the bias related to unobserved prognostic factors if the unobserved factors are highly correlated with the measured prognostic factors.\textsuperscript{28} Judgments should be made at the design stage about a potential of serious bias related to unobserved confounding. No statistical method can eliminate unmeasured bias.

Several analytical methods are available to estimate the safety and effectiveness (and adjust for confounders), typically assuming all important confounders have been measured.\textsuperscript{28} These methods involve stratification, regression models, or a combination of the 2 using propensity scores.\textsuperscript{29,30} Each approach relies on a set of statistical assumptions which may or may not be appropriate in the particular setting. When it is felt that there is unmeasured confounding present beyond that accounted for in the collected information, another potential approach is that using instrumental variable based methods.\textsuperscript{31,32} An instrumental variables approach also relies on a

\begin{table}[h]
\centering
\caption{Summary of the Methodological Issues and Recommendations for Clinical Studies of Safety and Effectiveness Involving Implantable Medical Devices}
\begin{tabular}{|l|l|}
\hline
Methodological Challenges & Recommendations \\
\hline
Blinding/masking: double or single blinding/masking is difficult to achieve or is impossible (mostly RCT, OS) & Use robust allocation concealment methods (RCT) Use independent outcome evaluators (RCT, OS) \\
Timing of randomization: inclusion/exclusion is affected after randomization/initial intend (RCT, OS) & Carefully assess the potential for cross-over. (RCT, OS) Discuss applicability of intention-to-treat and limitation in all studies (RCT, OS) \\
Learning curve: mastering the use of the IMD may affect outcomes (RCT, OS) & Estimate comparative evidence using skilled operators for treatment arms (RCT) “Real world” applicability enrolment of operators with variety of skills is necessary (OS) Report learning curve, volume-outcome relationship and threshold volumes whenever possible (RCT, OS) \\
Small sample size: is a reality in many instances and the study is underpowered to detect smaller differences or evaluate less common outcomes (RCT, OS) & Carefully document and report all outcomes even if underpowered to detect any differences (RCT, OS) Conduct longer follow-up studies. Consider Bayesian methods (RCT, OS) Use systematic review if there is more than one study and conduct meta-analysis if possible (RCT, OS) \\
Covariate imbalances across treatment groups: risk of imbalances in patient, physician, and hospital characteristics is large in studies that lack randomization (OS, sometimes RCT) & Document all imbalances in patient characteristics (RCT, OS) Use robust statistical methods to adjust for known confounders. (RCT, OS) Discuss the choice of the method, justify assumptions, and list the advantages and the limitations (RCT, OS) \\
Device modifications: device may undergo progressive modifications during the study period (RCT, OS) & Define the characteristics of the device (when it changes during the study) and outcomes that make study still valid (RCT, OS) Carefully document all changes to device and determine data “poolability” from clinical experts and in vitro bench testing (RCT, OS) Determine “poolability” through statistical testing or through inclusion of component in analytical model (RCT, OS) “All comer” studies including consecutive patients and operators with variety of experience and training as these are likely users of the technology (RCT, OS) Efforts need to made to standardize definitions of the variables in all IMD settings (OS) Need a device identifier (OS) \\
\hline
\end{tabular}
\end{table}

Parentheses denote applicability: RCT-designates recommended for trials; OS-designates recommended for observational studies.

It is understood that recommendations may not always be practicable.
set of assumptions the plausibility of which depends on the clinical scenario, availability of data, and strength of the instrument.

Generalizability Is Affected When Less Sick Patients, Elite Operators Participate in the Trials, or When Various Definitions Are Used for Predictors and Outcomes

Generalizability is defined as the extension of research findings and conclusions from a study conducted on a sample population to the population at large.\(^{33}\) It is important to design studies that consecutively enroll patients and operators ("all-comer") and make every effort to not exclude patients based on disease severity or comorbidities. We also need more research to determine whether more inclusive observational studies of "real-world settings" can add value when trials are highly selective. This is particularly challenging when IMDs are used off-label, which they often are as part of the practice of medicine.

Use of various definitions in the trials is another factor that limits generalizability of the IMD studies. Standardization of data elements that will be followed by all investigators of IMDS for a specific device is a good practice. However, in some instances investigators still use various definitions for the outcomes and endpoints in the trials. More research is needed to understand variability in definitions and how this may impact the usefulness of comparative study results in IMD setting. This concern is more pronounced in observational studies as in addition to varying definitions there is the problem of using varying diagnostic methods for detecting outcomes of interest. Use of varying approaches may introduce bias in outcome ascertainment and made study results difficult to compare.

Future Needs: Novel Data Sources and Unique Device Identifiers

Recently new data sources are developed that rely on high quality clinical registries and their ability to link with other databases. Clinical registries constitute a great infrastructure for novel large-scale studies that link national procedural registries such as National Collaborative Data Registries (NCDR) with robust claims databases such as CMS' Medicare claims database.\(^{34}\) These studies demonstrate that nationally representative analyses of IMDS are feasible using clinically rich procedure level registries and claims-based administrative databases for follow-up. In combination, these 2 resources provide a powerful mechanism for studying postmarket use and outcomes of IMDS devices.

A problem unique to IMDS (and medical devices in general) is the absence of a unique device identifier. This is in contrast to drug studies where there is standardization of drugs through the use of a National Drug Codes which is a unique product identifier for human drugs available through a unique 3-segment number. The absence of unique device identifiers makes it extremely difficult, if not impossible, to assess manufacturer-specific product performance in the real world in many large healthcare databases (such as those maintained by health plans) or even in procedure registries.

SUMMARY AND RECOMMENDATIONS

In this article we present a conceptual framework for IMD study design and evidence evaluation and discuss factors impacting comparative safety and effectiveness of implantable medical devices. We discuss some of the most urgent and current methodological issues that are faced in RCTs and observational studies. In doing so, we have not reviewed an exhaustive list of issues or dealt with all of the methodological challenges; rather, we have highlighted several key areas. Our goal is to raise the awareness of the issues particular to IMD study design and evidence evaluation, and to provide an outline and general direction. Table 1 summarizes the key issues and recommendations. We hope this article will inspire and help to inform those interested in advancing comparative safety and effectiveness of IMDS and to plan and pursue future methodological work in this area.

REFERENCES


