Understanding Evidence about the Value of Medical Technologies

Prepared for AdvaMed by:

Michael D. Miller, MD

May 2017

AdvaMed
Advanced Medical Technology Association

Introduction:
This report is designed to help policy advisors and analysts understand the types of information that can inform clinical and coverage decisions for medical technologies. The goals of this paper are to illuminate the complexity of medical technologies, to demonstrate the need for appropriately matching methods for generating evidence about value to each technology and clinical situation, and to elevate thinking about how such evidence is created, analyzed, and used. [Note – this report is intended to be a companion to "A Framework for Comprehensive Assessment of Medical Technology: Defining Value in the New Health Care Ecosystem" and uses some of the same nomenclature and categorization, including Figure 1. It is also a companion for the AdvaMedDx Report" A Framework for Comprehensive Assessment of the Value of Diagnostic Tests".]

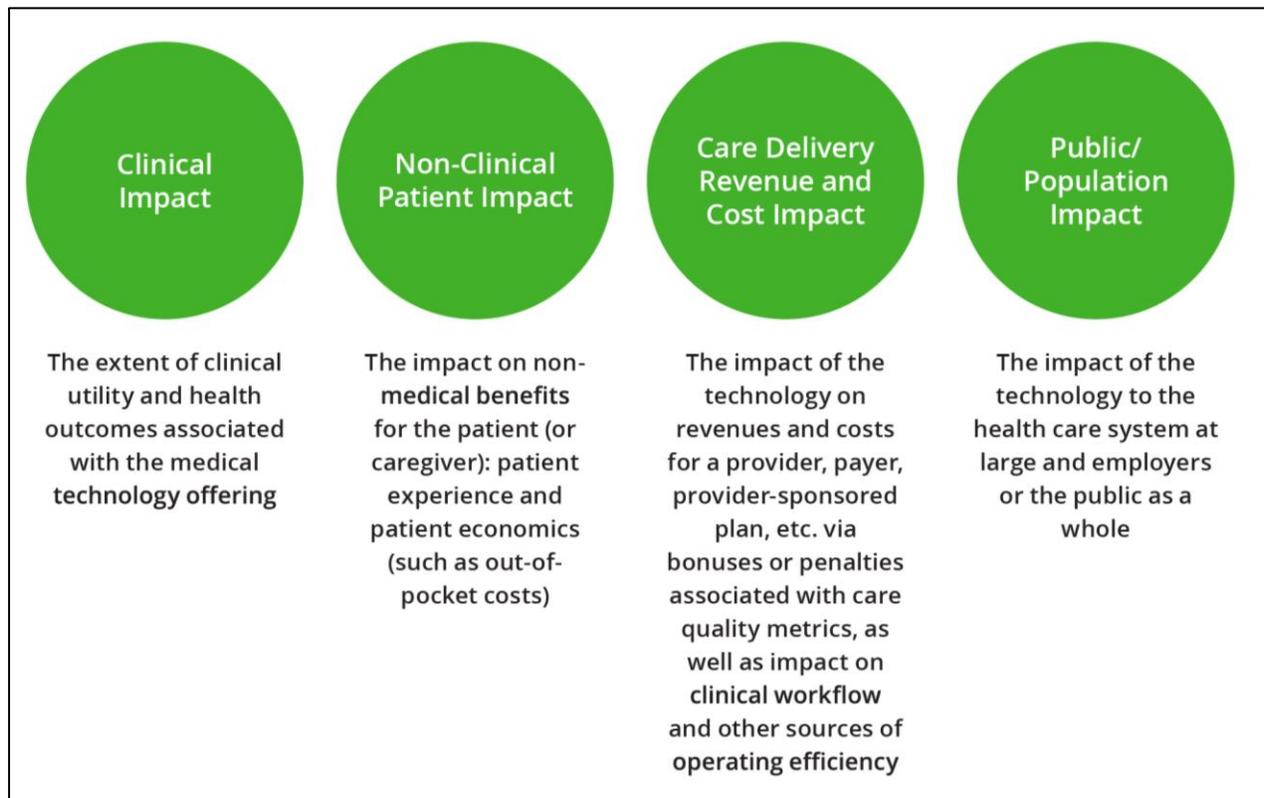**I. Context for This Paper**
With the accelerating shift in health care payment methodology from the volume of services delivered to the value of that care, all stakeholders – payers, providers, clinicians, patients, policymakers, and employers – are increasingly scrutinizing their choices and demanding quality evidence about the comparative benefits versus the cost of medical technologies.

To meet the requirements from different stakeholders, medical technology companies must increasingly present more information about the value proposition of their products and services that is specific to each stakeholder's concerns. Rising expectations about value propositions are leading all stakeholders to think beyond the traditional information about clinical and safety attributes required for the regulatory process to include: clinical impact, non-clinical patient impact, care delivery revenue and cost impact, and impact on the public and society more broadly. (See Figure 1 below.)

In meeting these rising expectations, medical technology and diagnostics companies must demonstrate each element of their value propositions by providing stakeholders with high quality evidence. This information must be both appropriate for the technology and to the stakeholder so that coverage, utilization, and payment decisions for innovative products are truly value driven. However, the desire for evidence of value must be balanced against the costs and time involved in collecting data, and how that process can delay the availability of novel technologies that address important unmet clinical needs. Responding to the expanding expectations for value information, medical technology companies have been developing and providing such information to stakeholders – particularly for high-cost technologies that have been subject to the greatest scrutiny – within the privacy and contractual limitations from regulatory agencies, and public and private payers.

This paper is designed to help inform discussions about those issues so that evidence about medical technologies can be developed - and used efficiently and effectively - by all concerned stakeholders to improve care for patients and to maximize the value medical technologies provide to patients and society.

Figure 1. Aspects of Value for Medical Technologies[1]

Clinical Impact	Non-Clinical Patient Impact	Care Delivery Revenue and Cost Impact	Public/ Population Impact
The extent of clinical utility and health outcomes associated with the medical technology offering	The impact on non-medical benefits for the patient (or caregiver): patient experience and patient economics (such as out-of-pocket costs)	The impact of the technology on revenues and costs for a provider, payer, provider-sponsored plan, etc. via bonuses or penalties associated with care quality metrics, as well as impact on clinical workflow and other sources of operating efficiency	The impact of the technology to the health care system at large and employers or the public as a whole

Expanding the range of evidence demonstrating the value of medical technologies available to payers (and other stakeholders) as they shift their focus from volume to value builds upon the traditional challenges of evidence generation within the medical technology industry.[2] Those traditional challenges start with the great diversity of medical technologies: from implantable orthopedic and cardiovascular devices, to minimally invasive surgical instruments, to imaging and radiation therapy equipment, to diagnostic tests. The extremely broad range of types and uses of technologies requires that different methods be used to demonstrate how various types of technologies create value. Meeting the expanded data and evidence needs of different stakeholders for showing value also requires more varied processes for generating evidence related to value.

For producing that evidence, there are ranges of options for both how it is generated and the specific types of data that go into the evidence. To efficiently provide high quality insights for different situations, the most appropriate and relevant options need to be selected. As stakeholders seek to understand value and appropriate use, it is also very important that they

---

[1] Derived from "A Framework for Comprehensive Assessment of Medical Technologies: Defining Value in the New Health Care Ecosystem, co-developed with Deloitte Consulting LLP, May 2017.

[2] Price R, Long G. Challenges in Developing and Assessing Comparative Effectiveness Evidence for Medical Technology. In: Greenberg PE, Birnbaum HG, editors. Decision Making in a World of Comparative Effectiveness Research: A Practical Guide. Singapore: Springer Nature; 2017 (In press).

recognize the very different characteristics of medical technologies compared to biopharmaceutical products, and how those differences determine the appropriate methods for collecting and analyzing evidence.[3]

This paper discusses several types of evidence – and the data supporting that evidence – that can be used to describe the value of medical technologies, and the relevance of each type of evidence in different value assessment situations and to different stakeholders. Finally, some general recommendations – including increased collaboration between medical technology manufacturers, payers, and providers – are presented for how to address challenges involving the generation and use of evidence about the value of medical technologies.

Definitions of Data and Evidence

While often used interchangeably, the words "data" and "evidence" have subtle differences: "Data" is the basic information (numbers or observations) obtained in an inquiry, while "evidence" is the information that is relevant for reaching some conclusion or insight, as in "the evidence pointed to the accused being innocent of the crime." Or looked at another way, data – as the more granular concept – is information from individuals, while evidence is that data aggregated and analyzed. For the purposes of this paper, the two are closely tied together since it is the relevant data from different studies that forms the evidence about value.[4]

II.   Challenges of Evidence Generation within the MedTech Industry

Data and evidence for medical technologies are used for different purposes, including regulatory approval, coverage and payment policies, and clinical guidelines or guidance. While the data used in those areas can certainly overlap, the amount and types of evidence different stakeholders may want or require are very different. Thus, to meet the needs of all stakeholders, generating the appropriate data and evidence can be very challenging for companies and researchers, and it can similarly be challenging for certain users who want specific types of data and analyses. For example, while a regulatory agency will likely want to see the entirety of the data generated by clinical studies of all types, payers – looking for evidence of clinical benefit and cost implications – will want detailed summaries of those same clinical studies along with analyses related to direct, indirect, and offsetting costs, as well as expectations about the size of affected populations. Clinicians will also want summaries of clinical studies, but may pay additional attention to types of patients, subpopulations, particular

---

[3] For example see: "Medical Device Clinical Trials – How Do They Compare with Drug Trials?" (http://mastercontrolinc.blogspot.com/2014/02/medical-device-clinical-trials-how-do.html?source=n3w5meddev accessed 4/16/17), and "Outlining the key differences between medical device and drug trials." (http://www.clinicaltrialsarena.com/news/medical-devices/outlining-the-key-differences-between-medical-device-and-drug-trials-to-determine-how-device-trials-can-best-be-run-effectively-4644599 accessed 4/16/17)

[4] " Big Talk about Big Data: Discourses of 'Evidence' and Data in British Civil Society", William L Allen, September 2014 (http://2plqyp1e0nbi44cllfr7pbor.wpengine.netdna-cdn.com/files/2014/10/Big-Talk-about-Big-Data-Final-Report2.pdf), and https://www.ucl.ac.uk/sts/staff/clarke/data-evidence_distinction

clinical situations, as well as information and evidence generated outside of clinical studies done for regulatory or payer purposes, including case reports, and real-world evidence. Clinical guidelines may also provide insights to clinicians about the processes and procedures for learning how to use new medical technologies, and supporting or confounding factors related to their healthcare environment. All stakeholders should be interested in patient preferences, risk tolerances, and quality of life.

Several factors should be considered in making decisions about the types of studies for generating evidence about value, including the diversity of medical technology, how medical technologies are an integral part of delivering complex care, rapid innovation cycles for many technologies, the diverse roles of technologies in making treatment selection and intervention decisions, and the applicability and appropriateness of different methods for generating evidence.

A. **Diversity of medical technology**: As noted above, medical devices and diagnostics encompass diverse groups of products ranging from implantable orthopedic and cardiovascular devices, to minimally invasive surgical instruments, to imaging and radiation therapy equipment, to *in vitro* diagnostics. These devices also vary widely in their levels of complexity, and degrees of risks and benefits for patients. Given this diversity, a "one size fits all" set of guidelines, principles, or a specific checklist for evidence generation, would be unsuitable and impractical. Conversely, evidence standards need to reflect different medical technologies' clinical use, level of complexity and risk, innovation time-cycle, the feasibility of data collection, and ultimately the aspects of value important to specific stakeholders and how the evidence will be used, e.g. regulatory approvals, coverage and payment, and clinical guidance.

B. **Integral part of complex care processes**: One challenge for developing evidence of value for many medical technologies is the way those technologies inherently produce clinical benefits and overall value only while embedded in complex processes of patient care, (such as surgical procedures), or as part of a diagnostic workup that guides decisions about therapeutic options. This is one way that medical technologies are different from biopharmaceuticals. Specifically, a clinician's expertise and experience, and the organization and operation of the clinic or institution can have a significant impact on clinical and economic outcomes. In addition, improvements in clinician experience or expertise – or changes to the healthcare delivery organization's operations or standardized practices – can lead to changes in outcomes or efficiency.[5] This learning and experience effect can confound evaluations of individual technologies, and comparisons between interventions – particularly when comparing medical technologies, biopharmaceuticals, clinician services, and patient or life style activities. This real-world variability complicates the processes of evidence collection and analysis for describing the value of a technology.[6]

---

[5] See, for example, David M. Cutler and Robert S. Hickman, "Technological Development and Medical Productivity: Diffusion of Angioplasty in New York State," Working Paper 9311, National Bureau of Economic Research, October 2002.
[6] For example, different studies have shown experienced surgeons produce better outcomes (See "Surgeon Experience and Clinical and Economic Outcomes for Shoulder Arthroplasty" JB&JS, 12/2003 http://journals.lww.com/jbjsjournal/Abstract/2003/12000/Surgeon_Experience_and_Clinical_and_Economic.8.aspx accessed

C. **Rapid innovation cycles**: After a medical device comes to market – and clinicians and patients gain experience in real world settings – improvements may continue to be made to the device itself. Examples of medical technologies that go through such rapid innovation cycles include software updates and modifications to a device's materials or structures. In addition, a medical technology's value proposition may change when additional uses are developed – potentially in combination with other technologies. The innovation cycle for new technologies is typically as short as 18-24 months. Over time, those incremental and cumulative innovations can alter clinical and cost-effectiveness, but they can also complicate the process of developing evidence that demonstrates value.

Therefore, as improvements to a technology are made (e.g. upgrades) – and as experience is gained with the changes – the initial assessments may be found to have underestimated its effectiveness, and the conclusions about its value may become out-of-date. This lag in updating the value of medical technologies to account for such changes can be minimized or shortened as healthcare systems increasingly use advanced IT and analytical systems that enable them to do more granular and timely data collection and analyses. Such systems should be able to embrace the rapid learning cycle of certain medical technologies and conduct appropriate analyses that are specific to their clinical situations and populations. And as advanced healthcare delivery systems are accepting more risk for clinical and economic outcomes, those delivery systems will have increased incentives to conduct such rapid updates to their value assessments for medical technologies. These capabilities – potentially integrated with data from companies and multiple payers – will also be important for supporting bundled payments and other "alternative payment models" where multiple stakeholders are sharing financial risks over timeframes that are appropriate for the medical condition and the value provided by the specific technology.

D. **Diverse roles of technologies in treatment selection and intervention process**: Diagnostic and imaging technologies (as a subset of medical technologies) present their own special analytic challenges. The core value of a diagnostic technology lies in how it enables improved clinical decision-making and therapy selection – which is distinct from the value of the underlying therapeutic intervention itself. That is, the diagnostic and the therapy act synergistically to produce value to patients because without the diagnostic, the therapeutic selection would be less precise and potentially result in worse outcomes or more adverse effects. Thus, diagnostics may have different evidence standards for efficacy since they are part of the process leading to a diagnosis or for making a clinical decision. For example, while there is increasing excitement about the value of precision therapies targeting specific genetic or metabolic mechanisms (particularly for cancers), the sequence of events enabling that value starts with a precise diagnostic test. Similarly, diagnostics that can provide more rapid results – or are available at the bedside or clinic (such as a handheld

---

4/16/17, and "Effect of Hospital Volume, Surgeon Experience, and Surgeon Volume on Patient Outcomes After Pancreaticoduodenectomy: A Single-Institution Experience" JAMA Surgery, 7/2010 http://jamanetwork.com/journals/jamasurgery/fullarticle/406128 accessed 4/16/17), as well as Atul Gawande's overview article "The Learning Curve," http://www.newyorker.com/magazine/2002/01/28/the-learning-curve, accessed 4/16/17

ultrasound device) – may provide additional value in terms of reduced time to making clinical decisions or increased access to more precise diagnostics.

E. **Applicability and appropriateness of different methods for generating evidence:** Randomized controlled trials (RCTs) are considered the "gold standard" of evidentiary support, and RCTs are almost always used for biopharmaceutical products, typically using double blinded protocols. When it is appropriate to use them, RCTs have the advantage of being designed to minimize (or eliminate) bias from the patient or researcher because they are both unaware as to who is receiving the active intervention and who is getting a placebo. This blinding reduces (or eliminates) one potential confounding factor in evidence generation. RCTs also can be designed with enough participants to provide statistically significant results assuming that a certain degree of benefit or "treatment effect" is found in the trial, which provides a level of certainty to all stakeholders about the benefits seen in the trial.

While RCTs are used when possible for medical technologies, that is not always feasible or ethical. Moreover, since RCTs tend to be one of the most costly and time-consuming forms of research, other methods may yield equivalent evidence for decision-making sooner and with less resources. For example, from a practical design standpoint, it is not always possible to apply randomization and double blinding methods for both patients and clinicians, or the small number of potential patients may not lend itself to an RCT structure. Furthermore, from an ethical perspective there are serious issues and risks associated with RCTs, such as exposing patients to sham surgeries from which they will likely not benefit.

For generating evidence about any medical intervention or innovation, the basic principle should be that the process of collecting evidence should be the one that is best at answering questions related to safety, effectiveness, and utility that patients, clinicians, payers, and regulators are seeking to answer. This will lead to the use of combinations of types of evidence that create timely and nuanced insights into medical technologies and their evolving value that are appropriate for the evidence needs of different decision makers. Examples of situations where certain methods for generating evidence are potentially more appropriate than others include:

- When the historical course of an illness is well-established, that may provide an adequate comparator group for an intervention if the study is developed to minimize bias;
- Where the intervention yields a clear clinical change that has previously been documented, observational studies may provide sufficient evidence;
- When there are long-term durability issues (for example, with an implanted mechanical device such as an artificial joint), part of the evidence may come from bench studies that can replicate years of use in a much shorter time; and
- Retrospective analysis of large real world evidence data sets may provide adequate support for clinical use or reimbursement of off-label applications of a technology.

The factors described above are important for understanding the most appropriate and efficient ways to collect the best evidence for medical technologies, and these issues are described in more detail in Section III. It should also be recognized that just as medical technologies can be very different from one another, as a group, they are also different from biopharmaceutical products. As such, applying the same standards for evidence to medical technologies and biopharmaceuticals is inappropriate. Some of those innate differences and how they apply to generating evidence related to value are presented in the box below.

General Characteristic Differences* Between Medical Technologies and Biopharmaceuticals, and Implications for Generating Evidence to Assess Value		
	Medical Technologies	**Biopharmaceutical Products**
Size and Scale:	• Macroscopic	• Microscopic
MOA:	• Mechanical (can involve physiological processes such as in diagnostic tests)	• Physiologic, often with systemic effects • At the cellular and subcellular level
Production:	• Engineered and manufactured • Sometimes programmed • Some aspects can involve chemical processes or nano-scale features	• Chemically synthesized – either de novo or starting from natural products – or grown in cell cultures or whole organisms
Relevance for Evidence Generation:	While both engineering and scientific research are based upon careful processes, measurement, and validation, biomedical science operates at the molecular level, where despite the explosion of knowledge, there are still many uncertainties and unanticipated outcomes from even well characterized molecules. In contrast, there is much greater certainty about the mechanism of action for medical technologies. Therefore, larger and more extensive testing is needed for biopharmaceutical products than for medical technologies. Moreover, in many cases, medical technology has local or mechanical effects that can result in permanent changes (e.g., replacement of a joint), rather than a drug's systemic effects, which may be reversible after the drug is stopped. And further, diagnostics may only interact with samples (such as blood) that are taken from an individual, and not interact directly with the patient.	
Speed of Innovation:	• Fast to very fast – with new versions and features added	• Slow and stable (Developing biosimilars is faster than generic drugs, but is still very significant)
Relevance for Evidence Generation:	Because innovation cycles for medical technologies can be very fast, requirements for extensive, long-term evaluation of medical technologies will often only produce fully applicable data about out-of-date products, not what is currently being used or has just been developed. Regulatory requirements reflect this rapid innovation cycle.	
Connection to Systems of Care:	• Highly connected • Can be dependent on user's experience or skills	• Less connected, although certain disease areas with complex regimens (such as oncology) have more

		connection
Learning Curve:	• Can be very significant.	• Generally much less.
Relevance for Evidence Generation:	Because the benefits a medical technology may be related to the systems of care in which it is deployed or the experience of the clinician using the technology, these factors must be considered when collecting evidence of effectiveness. In contrast, in most cases, the effectiveness of biopharmaceutical products is independent of the experience of the prescribing or delivering clinician, unless delivery is part of a procedure or through a medical device, such as during anesthesia or surgery. Complex biopharmaceutical regimens requiring multiple drugs, extensive monitoring, or frequent dosing adjustments can also have significant learning curves, as well as having their effectiveness and safety be more connected to the systems of care in which they are used.	
Patient Use and Interactions:	• Often implanted by a clinician (e.g., artificial joint) • Can be used directly by patients (e.g., knee brace) • Patients may have direct interactions with some diagnostics (e.g., MRI machines), and Health IT system, (e.g., telemedicine)	• Patients take medicines themselves (so-called self-administered medicines such as pills or certain types of injections), or the medicines are administered by a clinician.
Relevance for Evidence Generation:	Similar to the learning curve for clinicians and the importance for the systems of care in which a medical technology is used, how a patient interacts with a medical technology is an important consideration for how evidence is gathered about a medical technology. In many ways, the patient is an extension of the system of care in that they are the end recipient if not the end user. Even if patients do not directly handle the technology (such as a joint implant) patient characteristics can be important factors in determining the clinical and economic outcomes. For example, patient selection can be a factor in generating evidence if individual characteristics may determine their ability to properly use the technology, just as individual characteristics can influence a patient's adherence to a medication.	
* This table describes "general characteristics" of medical technologies and biopharmaceutical products. There certainly are exceptions to these generalities, and combination drug-device products include characteristics of both.		

III. **Types of Evidence and Guidance for Their Appropriate Use for Medical Technologies**
There are different research methods for generating evidence to demonstrate the value of specific medical technologies, whether done as a pivotal trial for regulatory purposes or in a post-approval setting to support coverage and payment, or additional uses. For any of those situations, the study type should align with the goals for the use of the evidence and the relevant questions from various stakeholders and decision makers. Similarly, the evidence – and how it is generated – should match the risks to patients, the practical limitations of evaluating the technology in a study, and other real world considerations for the use of the technology and conducting studies.

The two general categories of patient-centered evidence that can be collected are clinical and cost, however, each has many variants. [7] For example, clinical evidence can include mortality, disease progression, patient reported metrics, or many other measures of overall health. Similarly, evidence about cost can include direct spending or avoided costs, secondary costs or savings related to testing or services needed or avoided, or patient-related costs for transportation or lost work. And for both categories (and all the variants), the time-frame for the collection of the evidence is also crucially important, i.e. disease stability for a month or a year, and costs over the course of a hospitalization, or over a year, or over several years.

Overall, regardless of whether data are collected prospectively or retrospectively, the types of evidence that are collected must be directed toward the specific questions to be answered about the medical technology. Those questions vary according to the characteristics of the technologies and how the data will be used by patients, clinicians, payers and regulators for their respective decisions about care, access, and reimbursement. Unrealistic expectations about what types of evidence should be collected for a type of technology, or the metrics of the evidence collected, can lead to compromised patient access to care improving technologies.

Different processes for data collection, their appropriate uses, and some examples are described below:

RCT (Randomized Controlled Trial)
A RCT is a study in which similar people are randomly assigned to two (or more) groups to test a specific treatment or technology. One group (the treatment group) receives the intervention being tested, the other (the comparison or control group) has an alternative intervention, a dummy intervention (placebo), or no intervention at all. The groups are followed to determine the efficacy of the intervention. Outcomes are measured at specific times, and any differences in response between the groups are assessed statistically. This method is used to reduce bias and other variables beyond the intervention. RCTs can involve single or double blinding (i.e., clinicians and study participants are unaware as to which individuals are in the experimental and control groups), and other design characteristics to reduce bias and enhance the strength of evidence generated, e.g., cross-over design involves each group receiving the experimental therapy or control and then switching to the other after a predetermined amount of time.

RCTs are a method for collecting evidence in a prospective way that minimizes biases from researchers and helps control for variables beyond the technology being evaluated. While there are limitations for the applicability of RCTs depending on the technology (see section II), RCTs are used as evidentiary support when it is practical and ethical. From a trial design perspective, RCTs should be used when it is possible to design a "placebo" or control group of the trial, when large enough populations exist to reach a significant trial sample size, and when the time to conduct and analyze a trial will not be overwhelmed by the inherent innovation time-cycle of the technology. And of course, from an ethical perspective, RCTs should be used when the

---

[7] In general, studies that look at clinical evidence can be considered "clinical-effectiveness studies," while those that look at cost are considered "cost-effectiveness studies."

placebo or control will not do more harm to the patient beyond their current disease/condition – although in some situations a currently accepted technology (i.e., standard of care), can be used as a comparator in an RCT rather than a placebo. In general, RCTs cannot usually be used for technologies in clinically emergent situations; technologies with rapid innovation cycles; implanted devices where randomization might be challenging and informed consent more difficult; and very large or capital expensive diagnostic or imaging technologies where side-by-side comparison studies would be impractical, although use v. non-use studies when ethically acceptable could be possible.

Examples of medical technologies where RCTs could be appropriately used include:
- ***Diagnostic software from Electrocardiograms***
  RCTs of diagnostic software can compare the interpretations generated by the software to those of expert clinicians. Evidence from RCTs of diagnostic software has been used for regulatory approval. [8] Secondary data collection about speed and costs can also be built into such RCTs to support payment and coverage decisions.

- ***Ankle sprain treatments***
  RCTs have been used to compare different options for treating acute ankle sprains.[9] Studies of this type would be used for regulatory approval as well as coverage or payment decisions, and clinical guidelines. Because there could be significant patient interaction with the device (such as the ability to take it off for bathing) collection of data about patient preferences and real-world experience could also be built into these studies, which could be important for health plans where patient satisfaction is calculated as part of the quality scores for reimbursement purposes. A confounding factor for this type of study could be clinicians' skills - particularly for the customized splints. However, analysis of the learning curve for individual clinicians (e.g., time, clinical outcomes, patient satisfaction) could also be a useful data collection point as part of this type of study.

- *Additional clinical areas where RCTs could be used to evaluate medical technologies would include cardiology (such as stents to open arteries), and orthopedic implants. In those cases, the treatment group would be compared to the standard of care.*

Observational Studies
Observational studies include prospective or retrospective studies in which the investigator observes the natural course of events with or without control groups. One difference between an observational study and a RTC is the lack of randomization between the experimental and control groups. And while the individuals generating the evidence or data in an observational study (e.g. reading X-rays) may be unaware of what interventions the individual received, the participants themselves are not blinded to their interventions.

---

[8] For example, software to "read" an electrocardiogram generated by an ECG machine provides an initial diagnosis and has led to standard guidance from the FDA (see https://www.fda.gov/MedicalDevices/ucm073942.htm)

[9] For example, see "Treatment of severe ankle sprain: a pragmatic randomised controlled trial comparing the clinical effectiveness and cost-effectiveness of three types of mechanical ankle support with tubular bandage. The CAST trial" Health Technology Assessments, 3/2009 https://www.journalslibrary.nihr.ac.uk/hta/hta13130/#/abstract, accessed 4/16/17. In addition, a 2012 "Update on Acute Ankle Sprains" in The American Family Physician Journal provides an overview of different treatment options and approaches http://www.aafp.org/afp/2012/0615/p1170.html, accessed 4/16/17.

Observational studies may produce equally or more relevant data in a timelier manner for medical technologies compared to what RCTs can deliver. For example, observational studies may be relevant for generating evidence of value for medical technologies when:

- The population is too small to develop an appropriately powered RCT
- The benefits can be appropriately evaluated by comparison to a historical control group or through a matched pairs approach
- Blinding and non-randomization are inappropriate or unethical
- Treatment adherence varies among different technologies
- Clinicians have different levels of training that may affect patient care outcomes
- Evaluating long-term durability with mechanical tests

Observational studies may also be seen by some stakeholders (or for some uses) as superior to RCTs because they generate "real-world" evidence and can include several different designs: retrospective and prospective studies, cohort studies, case-controlled studies, and cross-sectional studies. Some major types of observational studies are described below.

Prospective Cohort Observational Study (Longitudinal with Comparator Group)
An observational study with two or more groups (cohorts) with similar characteristics. One group receives a treatment or technology, and the other group does not - or receives a different therapy. As noted above, the major difference between this design and an RCT is the lack of randomization between the two cohorts. The study follows their progress over time from when they receive the intervention, and records are reviewed to collect data for analysis at multiple intervals. Unlike an RCT, observational studies can have greater sources of bias that can make the results less certain. For example, a patient or clinician by knowing they are receiving a specific treatment may change their behavior, which could alter the outcomes.

Examples:
- ***Wound healing***
  Prospective observational studies could evaluate different technologies for wound healing. For example, negative pressure treatments v. normal would dressings.[10] (Additional types of treatments for hard to heal wounds can include hyperbaric oxygen or ultrasonic debridement.) This type of study would also be amenable to collecting information about patient experience since wound healing can potentially be distressing and painful. Clinicians and patients would not be blinded to what type of treatment being used, and this could also be a clinical area sensitive to the clinician's experience. The data from such studies could be used for regulatory approvals, as well as for coverage and payment decisions since similar studies could evaluate differences in utilization of other healthcare services.[11] However, because patient variables would have to be carefully accounted for (such as their underlying medical conditions, and the location and cause of wounds) this could lead to specific

---

[10] See "Evaluation of chronic wound treatment with the SNaP wound care system versus modern dressing protocols," https://www.ncbi.nlm.nih.gov/pubmed/20885246, accessed 4/16/17, which compares usual dressings with negative pressure treatments to enhance wound healing.
[11] For example see "A comparison of diabetic foot ulcer outcomes using negative pressure wound therapy versus historical standard of care," https://www.ncbi.nlm.nih.gov/pubmed/17651226, accessed 4/16/17

12

> *subpopulation or causal factor decisions for both regulatory approvals and coverage rules.*

- *Another example is the testing for STDs in the emergency room using rapid diagnostic tests.[12]*

Prospective Observational Studies Using Registries

Collecting evidence using patient registries is another form of prospective cohort study. Registries use organized systems to prospectively collect uniform data (clinical and other) to evaluate specified outcomes for a population defined by a particular disease, condition, or exposure, and to serve predetermined scientific, clinical, or policy purposes. Registry-based studies do not include a prospective comparator group, so they are often best used for long-term tracking and evaluation of clinical effectiveness and safety. However, registry-based studies can carry significant long-term maintenance costs, which must be a consideration for how and under what circumstances they are established, i.e., for what technologies and what data elements are collected. As discussed above, the evidence to be collected should be matched to expected uses. Registries can also combine patient data and archived medical samples, which has led to breakthrough understandings about disease diagnosis, progression, and treatment, especially in oncology. Overall, as a 2014 Agency for Healthcare Research and Quality noted, "Medical device registries play an increasingly important role in bridging the gap between device performance in clinical trials and their use in routine practice over time."[13]

Because a registry's data elements are established when the registry is created, retrospective studies using a registry are limited by the data that have been included in the registry – or can be obtained from secondary sources such as electronic medical records. In addition, registries are often used for technologies that are implanted, have serial numbers or can otherwise easily be tracked, and particularly for following patients over the lifetime of the device even as they see multiple clinicians or move geographically.

Examples:
- ***Pacemakers***
  Studies using registries of pacemakers allow for tracking the effectiveness of the device, and understanding and demonstrating – under what conditions – the device may improve clinical outcomes by preventing certain cardiac events such as arrhythmias or death, as well as long-term complications.[14] Registry studies can be used to expand or accelerate regulatory approvals and clinical indications. And to the extent that the registry data can be linked to other clinical data - for example in connection with electronic medical records (see below) - it can be used to evaluate utilization and cost considerations related to outcomes.

---

[12] See "Impact of rapid diagnostic testing for chlamydia and gonorrhea on appropriate antimicrobial utilization in the emergency department," http://www.dmidjournal.com/article/S0732-8893(16)30346-7/abstract?cc=y, accessed 4/16/17. Such testing can lead to more appropriate treatments, fewer adverse effects (such as infertility), and lower costs.

[13] Introduction to "Registries for Medical Devices," Ch23 in "Registries for Evaluating Patient Outcomes," 3rd edition, AHRQ, 2014 Apr., Report No.: 13(14)-EHC111 (https://www.ncbi.nlm.nih.gov/books/NBK208640/)

[14] For example see "Complication rates associated with pacemaker or implantable cardioverter-defibrillator generator replacements and upgrade procedures: results from the REPLACE registry," Circulation 2010, https://www.ncbi.nlm.nih.gov/pubmed/20921437, accessed 4/17/17

- *Joint Replacements*
  Studies using registries for joint replacements can be used to track longevity of the implant and how frequently problems arise or revisions/replacements are needed.[15] These types of studies can be used for guiding clinical use, such as patient selection for certain types of devices. And to the extent that the registry data can be linked to other clinical data - perhaps in connection with electronic medical records (see below) - it can be used to evaluate utilization and cost considerations related to uses of the different types of joint replacements. An additional way registries could be used prospectively would be to evaluate different characteristics of the procedure for implanting the device or pre- or post-operative patient care and how those factors affect the performance of the joint replacement and the patient.

Retrospective Observational Clinical Studies Using Medical Records
Using patients' medical records as evidence about the clinical (or other) outcomes from medical interventions is by definition a retrospective study since the medical records would be reviewed after the intervention had been delivered to the patient.

Medical records are a rich source of data about patients' actual interventions and outcomes, but typically lack specific cost information (although utilization can act as a proxy for costs). Information from medical records can be challenging to aggregate and analyze if it is: not all in electronic formats, in different formats, in non-standardized formats (e.g., free text), or is spread across multiple clinician and provider systems. In addition, the accuracy of all medical records information may be variable, with the propagation of inaccurate information (or failures to update information) being facilitated by cut-and-paste functionalities of electronic record systems.

Example:

- ***Physiological Monitors in Intensive Care Units (ICUs)***
  With increased electronic connection of medical records and physiological monitoring devices, such as blood gas analyzers, medical records can be used to retrospectively assess how physiological parameters affect clinical outcomes compared to a pre-existing database, i.e., to evaluate the technology against the null hypothesis, or the principle that "not everything that can be measured matters."[16] This type of evidence could be used to evaluate cost effectiveness, including additional resources used by certain patients as well as time motion studies to evaluate how such monitoring effects the efficiency of ICU staff. This information would then be useful for guiding purchase and usage decisions by hospitals and health systems, as well as staffing and even architectural design of hospitals or other healthcare environments.

---

[15] For example see, "The American Joint Replacement Registry—the first 5 years," Arthroplasty Today, March 2017, http://www.arthroplastytoday.org/article/S2352-3441(17)30007-9/abstract, accessed 4/17/17.
[16] For example see "Multiparameter Intelligent Monitoring in Intensive Care II (MIMIC-II): A public-access intensive care unit database," Critical Care Medicine 2011, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3124312/, accessed 4/171/17; and "Using electronic health record collected clinical variables to predict medical intensive care unit mortality," Annals of Medicine and Surgery, November 2016, http://www.sciencedirect.com/science/article/pii/S2049080116301315, accessed 4/17/17

Retrospective Observational Studies Using Cost Data

Data from hospitals, doctors' offices and clinics, and other providers can be used to provide information relative to the comparative effectiveness (clinical and cost) implications of medical technologies and outcomes. Such data can be a robust source of financial information, and is often able to provide additional information, including diagnostic information, therapeutic, treatment and pharmaceutical interventions, and health care provider and institutional information. However, this type of data may often be less granular than clinical information gained through the electronic health record or clinical trial, and depending on the data base and information stream may have reporting delays of several months before it is sufficiently aggregated, reconciled, and available for analysis.[17] This limitation is also found in many retrospective chart reviews and less automated data capture systems utilized in randomized clinical trials.

Example:
- *The treatment of diabetic ulcers is one example for how data can be used to understand how costs are produced from different services and care settings.[18]*

Case Studies

Case studies compare one or a few patients to the known natural history of patients with the condition or clinical situation being evaluated.[19] Case studies are typically done to highlight unusual situations that may provide useful insights for additional study, or to provide information about rare situations where more formalized RCTs or prospective types of studies would be impractical because of the small number of patients. In addition, case studies or small series can provide real-world context for the use of certain medical technologies. (Also see Patient Reported Evidence below.)

Additional Types of Evidence or Data from Research Studies

Two additional sources of data or evidence about medical technologies are Patient Reported Evidence (PRE), and the syntheses of multiple studies and sources of data into a new or more comprehensive conclusion about a medical technology. This latter activity can be conducted as a statistical meta-analysis of prior studies, or as the consensus of experts in the clinical area. While meta-analyses typically combine results from similar types of studies (e.g., RCTs or Observational), consensus statements evaluate all types of evidence, but may weigh each differently. It should also be noted that Patient Reported Evidence can be collected in all types of studies from RCTs to Case Studies. Each of these are briefly discussed below:

Patient Reported Evidence (PRE)

Report of the status of a patient's health condition that comes directly from the patient, without interpretation of the patient's response by a clinician or others. The importance of PRE is that it

---

[17] "Big data" in healthcare is often used as an all-encompassing term that includes the aggregation of clinical and cost data, which holds the promise for providing more granular analyses of healthcare technology.

[18] See "A cost analysis of diabetic lower-extremity ulcers," Diabetes Care, 2000, http://care.diabetesjournals.org/content/23/9/1333.short, accessed 4/17/17, and See "Burden of Diabetic foot Ulcers for Medicare and Private Insurers," Diabetes Care, 2013, http://care.diabetesjournals.org/content/early/2013/10/29/dc13-2176.full-text.pdf, accessed 4/19/17

[19] Occasionally a case study can be done prospectively when an unusual clinical situation presents itself, and the clinicians/researcher has the foresight to collect data and other information from the outset.

incorporates aspects of the patients' perspectives not captured as more formal clinical or economic evidence. Examples of this type of PRE can include ability to work, sleep quality, mobility, and emotional responses (such as pain or anxiety) to their use of or exposure to specific technologies. Unlike more structured methodologies for collecting evidence, PRE may be more qualitative information (or subjective assessments assigned quantitative metrics) rather than quantified data based upon specific reference criteria such as a clinical lab blood measurement. However, PRE can also be collected as part of other types of research such as RCTs and prospective clinical studies. In addition, PRE can be used to support clinical guideline decisions as well as coverage and payment decisions, particularly as part of quality of care metrics that clinicians or health systems are held accountable for when they are in risk or gain sharing arrangements. And lastly, PRE can be a useful method for gaining insight into both the extent of patient related indirect costs (such as transportation or work loss), and the significance that patients place upon those factors.

Meta-Analyses & Consensus Statements
Synthesis of multiple studies or types of information can provide additional insights about a medical technology to paint a more nuanced picture of the evidence, and potentially help fill in areas that are important to stakeholders but which individual studies have not provided sufficient evidence. As described in more detail below, meta-analyses can provide quantitative insights that may be of use to a broad array of stakeholders, while the primary utility of consensus statements is to guide clinician's shared decision making processes with their patients. However, while there are formalized processes for both meta-analyses and consensus statements they are not without controversy. For example, the statistical significance of meta-analyses may be misleading because of differences across the individual studies for patient selection criteria or the how the medical technology is used. Similarly, consensus statements are dependent upon the experience of the experts providing their input to the process.

> **Meta-Analyses:** A method that uses statistical techniques to combine results from different independent studies and obtain a quantitative estimate of the overall effect of a particular intervention or variable on a defined outcome. Thus, it is a statistical process for pooling data from many clinical trials to produce a stronger conclusion than can be provided by any individual study.

> **Consensus Statements:** Synthesis of many types of information by experts in a specific field based upon both the available data and their collective experiential wisdom in the clinical or technical area using processes where different types of evidence are weighted, and individuals' expertise and perspectives are collectively, aggregated, synthesized, and reported in structured formats.

Conclusions About Using Different Types of Evidence to Describe Attributes of Value
Analyses of medical technologies can – and very often should – involve aggregating and integrating evidence from many separate studies (e.g. meta-analyses), and involve modeling or estimating long term effects – recognizing that rapid innovation can invalidate such long-term assessments. These processes may involve selecting or differentially weighing different types of evidence, but generally recognize that randomized controlled trials (RCTs) are the gold standard

for evidence generation. However, as discussed above, it is appropriate to use certain types of research studies to generate evidence about specific types of medical technologies, and further, the intended use of the evidence by specific stakeholders should guide what types of evidence are used in their analyses and decision making. In these processes of evaluating and using evidence, diminishing the importance of certain types of evidence simply because it is not from an RCT when the data has been generated from a research type that is an appropriate match for the technology and the intended use of the data is clearly problematic. This is particularly the case when the exclusion or de-weighting of evidence is based upon comparisons to studies of biopharmaceuticals or medical technologies from entirely different categories that have different novelty attributes, risks, or clinical applications.

IV.   <u>Conclusions and Recommendations</u>

Stakeholders are increasingly looking to make the best possible choices about allocating funds and other resources for improving outcomes for patients. In developing innovations, medical technology companies ensure that the value of those innovations is clear and well-understood. Increasingly, medical technology companies define, as early as the product design phase, a clear strategy for developing evidence. Evidence generation strategies are informed by discussions with payer, physician and provider stakeholders about their evidence expectations for the technology. Because many medical technologies are characterized by a continuous flow of incremental product improvements, a strategy for developing evidence is at the forefront of company priorities throughout a product's lifecycle, and then continuously after initial approval or clearance. Companies are also increasingly looking for patient perspectives as part of their evidence development strategies.

Similarly, payers and providers should recognize the range of legitimate evidence types, and have appropriate processes for incorporating that evidence into assessment methodologies to ensure high value technologies reach their patients. As healthcare delivery and financing in the U.S. moves rapidly toward greater accountability for economic and clinical outcomes – and increased risk sharing with technology companies – more collaborations will be needed among medical technology manufacturers, providers, and payers to align evidentiary support requirements and outputs so that patients have appropriate access and choices.

Determining which studies of what types are adequate and appropriate for including in a methodological process for evaluating a particular technology is a crucial decision since data inclusion or exclusion can dramatically alter the results of the evaluation – no matter which aspects of value for a medical technology are being evaluated. Such differences can lead to differences in coverage and access rules and restrictions. Another important factor in using and selecting methodologies for generating data and evidence to evaluate medical technologies is how different types of evidence are weighed relative to each other. Generally, evidence from RCTs is considered the most influential or important because of the potential for less bias and fewer potentially confounding factors. However, as has been discussed above, all types of evidence can be important for making appropriate access and utilization rules – particularly in situations when observational studies are the most appropriate evidence collecting process and RCTs are inappropriate or impractical. Therefore, dismissing non-RCT generated evidence is very problematic since it sets an unreasonable and unrealistic bar for many medical

technologies that would be detrimental to the quality of care for patients and to the long-term increase in value for the health care system.

Recommendations:

1. **Evidence and Data Generation**
    - Evidence collected to demonstrate the value of specific types of medical technologies must be appropriate for that technology and the needs of the relevant stakeholders.
    - To generate evidence, the research activities should be conducted as efficiently as possible to answer the relevant stakeholders' questions based upon the uses the evidence will serve.
    - The types of evidence developed or required for a medical technology must reflect the technology's clinical use, patient perspectives, and innovation and clinical user-experience time-cycles. In other words, the evidence generated should match the medical technology and its risks, expected benefits, uncertainties, differences from existing options, and other relevant factors, and similarly be aligned with the intended use of the evidence and the final assessments.

2. **Assessment of Evidence**
    - Methods for analyzing the evidence about medical technologies must be appropriate for both the type of evidence and the aspects of the technology. For example, analyses should not disregard evidence if it is not from controlled or blinded trials because the importance of incremental innovations often cannot be captured effectively or efficiently through such RCTs.
    - Both clinical and cost evaluations must be done in the context of specific patient populations, not as generalized approaches. For example, evaluations of a technology by a Medicare Advantage plan, a large urban public hospital, and a small rural hospital could all be very different.

3. **Decision-Making Using Evidence**
    - "One size fits all" approaches are not appropriate for demonstrating aspects of clinical or economic value for medical technologies.
    - Evaluations of cost implications for a medical technology should be conducted within the scope of the organization's patient populations, from specific stakeholder perspectives, (e.g., the patient), and within timeframes appropriate for the technology and the stakeholder. More comprehensive analyses of costs with a broader scope (e.g., national) may be appropriate for public policy uses, but should be conducted with the transparency and accountability required for government and regulatory activities.
    - Decisions about the appropriate types of evidence, and how they should be analyzed, should be done in collaboration among all key stakeholders so that the benefits of medical technology innovations can be available to clinicians, patients, and payers in efficient and equitable ways. In contrast, restrictions on coverage by payers because there isn't enough evidence (yet), or limitations on evidence

collected by manufacturers, (without specific rationale according to the concepts described in this document), should be discouraged since it will only lead to delays in access for patients and society.

Overall, the process of generating and using data and evidence that informs all stakeholders about the value of medical technologies should be done in a way that improves their working together to improve quality, lower costs, and improve outcomes, while ensuring that Americans have appropriate access to the best technologies that improve their lives – and our society – as quickly as possible.