Has the Time Come to Replace Randomized Controlled Trials With Real-World Evidence? A Case of Medical Devices

Katarzyna Kolasa, PhD, Kozminski University, Warsaw, Poland; **Bella Huasen, MBChB, MRCS, DipEd, FRCR**, Lancashire University Teaching Hospital, Manchester, England, UK

On April 5, 2017, the medical devices industry welcomed a new medical device regulation (EU 2017/745) that will go into effect on May 25, 2020. The key change is strengthening the importance of clinical evidence in the regulatory process.

Medical device regulation introduced a clinical evaluation report as a framework for generating relevant information with respect to the safety and performance of medical devices, as well as its evidence on clinical benefits. The clinical evaluation report is defined as a live document with regular updates based on insights generated during day-to-day experience with a given medical device. This report is to be based on the critical evaluation of the scientific literature and real-life data collection. The medical device regulation includes postmarket clinical follow-up to ensure continuous input into a clinical evaluation report. This process is largely new for device companies and clinicians." Its objective is to routinely collect and evaluate data regarding the utilization of medical devices in real-life clinical settings.

Medical device regulation introduced a periodic safety update report for class IIa and III. In addition to postmarket clinical follow-up data, periodic safety update reports encompass postmarket surveillance as well as benefit-risk analysis. Manufacturers are also required to feed into periodic safety update reports information regarding characteristics of treated patients. For class III and implantable devices, a periodic safety evaluation report will require almost yearly updates. The compliance with medical device regulations will be vital as the periodic safety update report, along with the vigilance report and other reports, will be used to populate the European Databank of Medical Devices (Eudamed). Its key focus is market surveillance, but more detailed regulations regarding the structure of Eudamed will be defined by the end of 2019.[1]

Fulfilling requirements of the clinical evaluation report, postmarket clinical follow-up, and the periodic safety update report, which fully rely on real-world data, will not be enough. Thus, medical device regulations introduced an additional criterion of further clinical investigations, limited to the implantable and class III devices. A "clinical investigation" might be interpreted as the need for a randomized controlled trial (RCT), although this is not explicitly stated. Among the endpoints listed, there is the intended purpose, performance, safety of the device, and clinical benefit. Clinical benefits are defined very broadly as "the positive impact of a device on the health of an individual, expressed in terms of a meaningful, measurable, patient-relevant clinical outcomes, including outcomes related to diagnosis, or a positive impact on patient management or public health."

Deliberation on incremental benefit or the need for a comparator control study has not been introduced." Medical device regulation does, however, mention a need for "consideration of currently available alternative treatment options for that purpose, if any."

In summary, medical device regulation has introduced multiple references to the need for real-world evidence while omitting the explicit requirement for RCTs. It provokes therefore a question about the future of RCTs in the process of medical device assessment. Can real-world evidence generation be a better choice than RCT in the evaluation of a medical device's clinical and economic benefits? There are multiple arguments that need to be weighed before any meaningful conclusions can be drawn. One can divide them into 2 groups related to pre- and postmarket launch phase.

Prelaunch Phase

Feasibility

On average, 18 months is the suggested life cycle of a medical device.[2] There are at least 2 reasons for such a short time horizon. First, medical devices can be developed for either therapeutic or diagnostic purposes, with this scope of use being changed during the clinical development. Second, unlike pharmaceuticals, medical devices are usually invented. Innovation originates

primarily from end users' insights, rather than laboratory exploration. Medical devices undergo constant "incremental" development based on clinicians' or patients' feedback. Thus, one can assume that any RCT for a medical device should take less than 2 years to complete, otherwise its results will be released when the device is already outdated. Taking into consideration the fact that the median time of a phase III trial for pharmaceuticals ranges from 3.8 to 7.2 years, the question can be asked, whether just 2 years is truly a feasible time horizon for an RCT.[3] This question relates to both efficacy (is it a sufficient follow-up period to assess treatment outcome?), as well as clinical aspects (is it a sufficient time period to assess safety of treatment?). One could allow for the evaluation of the medical device in real-life settings with the timed framework adjusted effectively to capture the full range of risks and benefits without the need for conducting RCTs.

The heterogeneity of studied patients may pose additional challenges to identify an optimal comparator. If available, one can leverage alternative data sources (such as existing registries or modeling techniques) to assess the incremental risks and benefits of given treatments against alternative treatment options. The adaptation of propensity score matching or difference-in-difference technique can ensure robust comparability of different patients' groups. Still, it does not allow for head-to-head analysis between such heterogeneous groups.

Ethics

It is a well-recognized ethical problem of placebo patients being left without active treatment. For the assessment of efficacy of medical devices, that challenge is even more profound compared to pharmaceuticals. There are some risks involved in simulating the intervention, such as anesthetic deployment and/or some surgical procedures, which may be required for both treated and placebo patients. The "standardizing" of the pre- or postoperative care of these patients can cause some disturbance as well. Blinding of participants, healthcare providers, or other caregivers in some cases may cause some risks to patients and/or be simply unrealistic. The centralized assessment of the main outcome can provide a solution, but it requires additional financial investments and tight organizational collaboration across different healthcare professionals. Real-world evidence generation is not free from ethical consideration either. Some arguments can be raised about the introduction of a new treatment prematurely before a robust level of evidence has been collected. The approach based on real-world evidence generation does mean introduction of health technology to clinical practice without the assessment of efficacy and safety on patient levels.

The clinical evaluation report is defined as a live document with regular updates based on insights generated during day-to-day experience with a given medical device.

Postlaunch Phase

End user experience

Different levels of end users' experience may lead to different levels of performance when carrying out interventions. Any RCT conducted before appropriate training and experience has been acquired, may not reflect the true clinical value of the new medical device. In such a case, an unfavorable assessment could reflect a poorly mastered technique rather than an ineffective device. One example is the analysis of 841 patients who underwent carotid endarterectomy performed by vascular or cerebrovascular neurosurgeons between January 2008 and December 2010. End users were categorized into low-volume surgeons with 40 or fewer cases per year and high-volume surgeons for higher numbers of patients treated. The complication rate of stroke and death was 6.9% for low-volume and 2.0% for high-volume surgeons (P=.001). Overall complications were 13.4% for low-volume surgeons versus high-volume surgeons 7.2% (P=.008).

The learning curve has its cost dimension as well. Another example can be the adaptation of difference-in-differences methodology to the study of total knee arthroplasties with and without bipolar sealer based on the PREMIER database in the United States. A comparison of 11,721 total knee arthroplasties and 6376 total knee arthroplasties with bipolar sealer performed in the same hospitals by surgeons with similar levels of experience in terms of number of procedures conducted in the past. The initial higher costs of bipolar sealer ($1335) were more than offset by subsequent cost savings in the second ($583) and third ($986) years post-adoption. In essence, the study provided evidence of how higher medical/surgical supplies costs can be compensated by efficiency gains such as shortened length of stay.[4]

Govindarajulu et al[5] found that learning curve models can be applied with generalized estimating equations and generalized linear mixed-effects to fit the data; however, the variability of institutional learning between different sites is likely to add to the error of most models. Overall, in the study of operator learning of a new mechanical thrombectomy device, the generalized estimating equations model tended to perform better. These models are assumed to be better applied during the vigorous initial clinical trials prior to US Food and Drug Administration approvals.[5]

To ensure an unbiased estimate of the clinical and economic benefits, it should be advisable to anticipate how long such learning phases are expected to take, and plan the timing of the assessment of a given medical device accordingly. Some examples of specific approaches for RCT can help to accommodate such challenges. An example in case can be the factorial RCT.[6,7] Since it requires a greater number of patients to be included compared to the standard randomized control study, the feasibility of such an approach may be challenging from the perspective of recruitment of study participants. The observational real-world studies on the other hand may provide more flexibility to control the impact of the end user experience. The previously mentioned difference-in-difference methodology can be a good example in case. It does, however, introduce some limitations with respect to the choice of healthcare professionals with regard to their experience with a ›

given medical device and alternative treatment options.[4]

Another issue that faces using RCT in medical device data collection, is when the clinician or the patient chooses not to participate in a new change. Equipoise may at times be the only process that justifies the test of alternative medical devices. If clinical teams believe their current medical device or technique is adequate, then they may not randomize patients to alternative strategies[12].*

Institutional Context

The clinical benefits of medical devices may be affected by the institutional factors as well. Given the fact that a result of a procedure is not only dependent on the medical device and surgical experience, but also on the complex circumstances in which the medical device is used.

The analysis of 1,377,118 patients eligible for laparoscopic abdominal surgeries in Japan between 2011 and 2013 revealed striking differences in the treatment results across more than 2000 hospitals included in the study. Not only facilities with a few cases, but also those with the highest case numbers constituted a high risk with regard to the patients' safety. It was concluded that the implementation of a new medical procedure into the clinical practice requires not only appropriate training of end users but also the implementation of safety standards.[8] Surely, it is difficult to estimate clinical benefits objectively without standardization of preoperative care (patients, hospital facilities, and equipment), perioperative care (duration of procedure, supplies), and postoperative care (assessment, follow-ups).[9] It may be even more difficult to define an appropriate comparator for the analysis. Finally, the lack of standardization of clinical care can make it challenging to define standard treatment outcome. It would be under the discretion of healthcare professionals, hence, differences across sites with respect to the reporting of treatment success.

Review of 42 studies of leadless pacemakers (pacemakers that are implanted directly into the patient's heart, avoiding the need for leads between the pacemaker and the heart, which are prone to infection) found some 2500 different individual outcomes reported.[10] It may be challenging to organize a unanimous protocol-driven RCTs if there are no standards of clinical practice across different healthcare providers utilizing the same health technology. Real-life observational study can, on the other hand, provide better understanding of suboptimalities in clinical practice and allow for evidence-based clinical guidelines generation. There are examples of such processes within the National Institute for Health and Care Excellence in the United Kingdom.[11]

Conclusions

It remains extremely challenging to define in "black or white" terms the best approach towards data collection for the assessment of new medical devices. The impact of end user experience and

Table 1. Summary of opportunities and challenges with both RCTs and RWE for medical devices.

	RCT *indicates randomized controlled trial*	**RWE** *indicates real-world evidence*
Opportunities	• Robustness in the assessment of efficacy and safety • Regime of RCT may accelerate the standardization and evidence-based utilization of medical devices in the clinical practice • Further adaptation of RCTs may increase the number of scientific publications about the value of medical devices and consequently, accelerate the uptake of innovation	• Real-life study of the benefits of medical devices provides insights into the challenges and opportunities with its use in the clinical practice • Real-life studies facilitate knowledge sharing about different methodological advancements allowing to control for end user experience • Real-life studies offer relative easiness with the engagement of healthcare providers and patients • The considerable flexibility with both retrospective and prospective real-life study framework allows for relatively easy recruitment of patients
Challenges	• Short life-cycle of medical devices provides limited time framework to study its value • Multiple devices used during a single procedure make it challenging to design an RCT that allows for assessing the efficacy of a given health technology • Ethical issues with sham comparator • Difficulties with avoiding bias of treatment effect estimation due to the learning curve effect • Lack of scientific advice, regulations and transparency on how to design RCT for a medical device. • Lack of healthcare standardization across clinical settings regarding the same health technology • Difficulties with the recruitment of homogenous patients' group (internal validity may be questionable)	• Difficulties with the assessment of the incremental value of a given medical device in cases where multiple technologies are used during a single procedure • Generalizability of study findings in the light of variety in treatment patterns across multiple healthcare providers (external validity) • Difficulties with defining an alternative treatment option for a given patient in a real-life settings • Limited possibility to study relevant endpoints in case appropriate data are not routinely collected in the clinical practice

institutional context makes it almost impossible to estimate unbiased efficacy of medical devices. One can wonder why we should search for it at all if the treatment outcome is so multidimensional in a real-life setting anyway. So far, it has been seen that real-world evidence is the more widely chosen approach for researching clinical benefits of medical devices than RCTs. Of the 215 clinical trials conducted, for 32 innovative medical devices, only 15% of them were RCTs.

There is some reluctance among healthcare professionals to study efficacy and safety of medical devices in the protocol-driven studies as well. A cross-sectional survey showed that 58% of orthopedic surgeons prefer to participate in expertise-based controlled trials compared to only 17% for conventional RCTs. Does this mean that real-world evidence will replace RCTs? The answer to that question remains unknown. It is, however, very clear that health economics and outcome research expertise is needed to guide both manufacturers and end users of medical devices in the organization of a robust approach towards data collection regarding the value of given health technology to patients, clinicians, and budget holders. The summary of opportunities and challenges with both RCTs and real-world evidence is further illustrated in Table 1.

In conclusion, it is not important which type of study framework is chosen, as long as the right research questions are posed, followed by an analysis plan that allows for appropriate information to be acquired from all potential data sources. After all, the ultimate goal is that this choice technique increases patient safety through proper scientific assessment of benefits and harms, both in the short- and long-term.

References

1. Regulation (Eu) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC.

2. Wilkinson J. *Medical Technology in Europe,* March 6, 2009. http://www.eucomed.org/~/media/7804F449C2154F8E9207E8E57B19DD4B.ashx. Accessed July 13, 2019.

3. Wong CH, Siah KW, Lo AW, Andrew WL. Estimation of clinical trial success rates and related parameters. *Biostatistics*. 2019;20(2):273-286.

4. Kuznietsova V, Woodward RS. Estimating the learning curve of a novel medical device: bipolar sealer use in unilateral total knee arthroplasties. *Value in Health*. 2018;21(3):283-294.

5. Govindarajulu US, Stillo M, Goldfarb D, Matheny ME, Resnic FS. Learning curve estimation in medical devices and procedures: hierarchical modeling. *Angiojet Stat Med*. 2017;36(17):2764-2785.

6. Montgomery AA, Peters TJ, Little P. Design, analysis, and presentation of factorial randomised controlled trials. *BMC Med Res Methodol*. 2003;3:26.

7. Giesbrecht EM, Miller WC. A randomized control trial feasibility evaluation of an mHealth intervention for wheelchair skill training among middle-aged and older adults. *PeerJ*. 2017;5:e3879.

8. Miyata H, Mori M, Kokudo N, Gotoh M, Konno H. Association between institutional procedural preference and in-hospital outcomes in laparoscopic surgeries. insights from a retrospective cohort analysis of a nationwide surgical database in Japan. *PLOS One*. 2018;13(3):e0193186.

9. Neugebauer EAM, Rath A, Antoine S-L, Eikermann M, Seidel D, Koenen, C. Specific barriers to the conduct of randomised clinical trials on medical devices. *Trials*. 2017;18(1):427.

10. Smith R. Improving the evaluation and regulation of medical devices. https://blogs.bmj.com/bmj/2018/09/25/richard-smith-improving-evaluation-regulation-medical-devices. Accessed April 10, 2019.

11. www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/nice-interventional-procedures-guidance. Interventional procedures guidance. Accessed April 10, 2019.

12. Neugebauer E, Troidl H, Spangenberger W, Dietrich A, Lefering R. Conventional versus laparoscopic cholecystectomy and the randomized controlled trial. Cholecystectomy Study Group. *Br J Surg*. 1991;78(2):150-154.

Additional information

The preceding article was based on an issues panel presented at ISPOR 2018. To learn more about the ISPOR Special Interest Group on Medical Devices and Diagnostics, go to www.ispor.org/specialinterestgroups.