IMDRF International Medical Device Regulators Forum

FINAL DOCUMENT

Title:	Methodological Principles in the Use of International Medical Device Registry Data
Authoring Group:	IMDRF Patient Registries Working Group
Date:	16 March 2017

Kimby M Barton

Kimby Barton, IMDRF Chair

Table of Contents

Preface

The document herein was produced by the International Medical Device Regulators Forum (IMDRF), a voluntary group of medical device regulators from around the world. The document has been subject to consultation throughout its development.

1.0 Introduction

The International Medical Device Regulators Forum (IMDRF) Registry Working Group was created with the purpose of developing:

(1) Essential principles for linking electronic patient, device and outcome registries and/or related data repositories or identifiers such as Unique Device Identifiers (UDIs), including the principles behind data access, security, informatics formats, governance and other key areas related to global regulatory applications for medical device evaluation; and

(2) Essential principles related to optimal methodologies for analysis of heterogeneous data sources applied to medical device safety signal detection, performance and reliability.

This document focuses on the task described in (2). In doing so, the document leverages the essential principles behind data access, security, informatics formats and other key areas related to global applications for medical device information described in (1). It applies the medical device registry definition from (1) and introduces the methodological concept of international Coordinated Registry Networks (iCRNs) to maximize the potential of data captured in the international registries.

This methodological document also builds on the IMDRF Common Data Elements (CDE) for Medical Device Identification document. The CDE effort outlines the common data elements for medical device identification that may be used through regulatory activities or process (pre-market, and post-market), including the future possibility of electronic regulatory submission of device identification information and covers the harmonization of terms and their definitions (IMDRF CDE).

Much of the material presented and developed here was preceded by multi-stakeholder work advanced by the U.S. National Medical Device Registry Task Force. In particular, the methods discussed there should apply reasonably directly to the international setting. This document highlights aspect of those considerations that differ in important ways from the national setting (Medical Device Registry Task Force).

2.0 Scope

This document provides information and guidance on:

- International coordination in methodologies that would add value to multiple international stakeholders including regulators;
- Methodological principles in the clinical evaluation of performance/effectiveness and safety across the device lifecycle using international Coordinated Registry Networks (iCRNs);
- Methodological principles in signal detection via iCRNs.

The focus will primarily be on implantable therapeutic devices, as this area represents highest risk devices with most registry activities and opportunity to reach consensus.

3.0 References

1. A Report from the Medical Device Registry Task Force & the Medical Devices Epidemiology Network. Recommendations for a National Medical Device Evaluation System. Strategically Coordinated Registry Networks to Bridge Clinical Care and Research, August 2015. Available at: http://www.fda.gov/downloads/aboutfda/centersoffices/officeofmedicalproductsandtobacco/cdrh/cdrhreports/ucm459368.pdf. Accessed August 6, 2016.

2. Global Harmonization Task Force. Clinical Evaluation. Study Group 5. N2:2007 (GHTF/SG5/N1:2007). Available at: http://www.imdrf.org/docs/ghtf/final/sg5/technical-docs/ghtf-sg5-n2r8-2007-clinical-evaluation-070501.pdf. Accessed July 26, 2016.

3. ISO 14971:2007; http://www.iso.org/iso/catalogue_detail?csnumber=38193

4. IMDRF Common Data Elements for Medical Device Identification. Available at: http://www.imdrf.org/docs/imdrf/final/consultations/imdrf-cons-cde-mdi-150708-2.pdf. Accessed August 9, 2016.

5. OECD (2013), Life expectancy at birth (indicator). Available at: https://data.oecd.org/healthstat/life-expectancy-at-birth.htm. Accessed August 6, 2016.

6. Wild D, Ermenco S, Mear I, Martin M, Houchin C, Gawlicki M, Hareendran A, Wiklund I, Yee L, von Maltzahn R, Cohen L, Molsen E. Multinational Trials - recommendation on the translations requires, approaches to using the same language in different countries, and the approaches to support pooling the data: the ISPOR patient-reported outcomes translation and linguistic validation good research practices task force report. 2009, *Value in Health*, 12(4):430-440. Available at: http://onlinelibrary.wiley.com/doi/10.1111/j.1524-4733.2008.00471.x/full

7. OECD (2016), "Non-medical determinants of health", *OECD Health Statistics* (database).
DOI: Available at: http://dx.doi.org/10.1787/data-00546-en. Accessed on 01 August 2016

8. Morales P, Cronenwett J, Thatcher R. Registry Assessment of Peripheral Interventional Devices (RAPID). Developing a minimum core dataset for total product life cycle device evaluation across multiple data sources: a step toward establishing a National Evaluation System for Health Technology for peripheral intervention devices. *Endovascular Today* - In press. Vol. 15, No. 8.

9. Sunil V. Rao, Connie N. Hess, Britt Barham, Laura H. Aberle, Kevin J. Anstrom, Tejan B. Patel, Jesse P. Jorgensen, Ernest L. Mazzaferri, Sanjit S. Jolly, Alice Jacobs, A Registry-Based Randomized Trial Comparing Radial and Femoral

Approaches in Women Undergoing Percutaneous Coronary Intervention *JACC: Cardiovascular Interventions*, 2014, Volume 7, Issue 8, Pages 857-867.

10. O'Malley AJ, Normand S-LT, Kuntz RE. Application of models for multivariate mixed outcomes to medical device trials: coronary artery stenting. *Statistics in Medicine*, 2003;22:313-336): Available at: http://onlinelibrary.wiley.com/doi/10.1002/sim.1337/epdf

11. Sedrakyan A, Campbell B, Merino JG, Kuntz R, Hirst A, McCulloch P. IDEAL-D: a rational framework for evaluating and regulating the use of medical devices. BMJ. 2016 Jun 9;353:i2372. doi: 10.1136/bmj.i2372

12. Idea, Development, Exploration, Assessment, Long-term (IDEAL) Retrieved from: http://www.ideal-collaboration.net/. Accessed August 9, 2016

13. Wang C, Dominici F, Parmigiani G, Zigler CM. Accounting for uncertainty in confounder and effect modifier selection when estimating average causal effects in generalized linear models. *Biometrics*, 2015, 71(3):654-665: Available at: http://onlinelibrary.wiley.com/doi/10.1111/biom.12315/epdf

14. Methodological guidelines and recommendations for efficient and rational governance of patient registries. The PARENT cross border PAtient REgistries iNITiative. Submitted for publication.ec.europa. Available at: http://ec.europa.eu/health/ehealth/docs/patient_registries_guidelines_en.pdf

15. Poloniecki J, Sismanidis C, Bland M, Jones P. Retrospective cohort study of false alarm rates associated with a series of heart operations: the case for hospital mortality monitoring groups. *BMJ : British Medical Journal*. 2004;328(7436):375. doi:10.1136/bmj.37956.520567.44. Available at: http://www.bmj.com/content/bmj/328/7436/375.full.pdf

16. Medical Device Alert. Available at: http://www.jisrf.org/pdfs/mediacl-device-alert.pdf. Accessed July 28, 2016

17. Australian National Joint Registry. Available at: https://aoanjrr.sahmri.com/en. Accessed August 1, 2016.

18. IMDRF Registry Essential Principles. Available at: http://www.imdrf.org/docs/imdrf/final/consultations/imdrf-cons-essential-principles-151124.pdf. Accessed August 3, 2016.

19. Del Trigo M, Munoz-Garsia A, Wijeysundera H, Nombela-Franco L, Cheema AN, Gutierrez E.Serra V, Kefer J, Amat-Santos IJ, Benitez LM, Mewa J, Jiménez-Quevedo P, Alnasser S, Garcia del Blanco B, Dager A, Altisent OAJ, Puri R, Francisco Campelo-Parada F, Dahou A, Paradis JM, Dumont E, Pibarot P, Rodes-Cabau J. Incidence, timing, and predictors of valve hemodynamic deterioration after

transcatheter aortic valve replacement: multicenter registry. *Journal of the American College of Cardiology* 67.6 (2016): 644-655. Available at: http://content.onlinejacc.org/article.aspx?articleid=2491128

20. Sedrakyan A, Paxton E, Graves S, Love R, Marinac-Dabic D. National and international postmarket research and surveillance implementation: achievements of the International Consortium of Orthopaedic Registries initiative. *J Bone Joint Surg Am*. 2014 Dec 17;96 Suppl 1:1-6. Available at: http://jbjs.org/content/96/Supplement_1/1.abstract

21. Macpherson GJ, Brenkel IJ, Smith R, Howie CR. Outlier analysis in orthopaedics: use of CUSUM: the Scottish Arthroplasty Project: shouldering the burden of improvement. *J Bone Joint Surg Am*. 2011 Dec 21;93 Suppl 3:81-8). Available at: http://jbjs.org/content/93/Supplement_3/81.abstract

22. Shewhart W. The application of statistics as an aid in maintaining quality of a manufactured product. *Journal American Statistical Association.* 2012 1925;20:546–548

23. Spiegelhalter DJ. Funnel plots for comparing institutional performance. *Stat Med* 2005;24:1185–1202).

24. de Steiger FR, Hang JR, Miller LN, Graves SE, Davidson DC, FRACS, Five-Year Results of the ASR XL Acetabular System and the ASR Hip Resurfacing System An Analysis from the Australian Orthopaedic Association National Joint Replacement Registry. *J Bone Joint Surg Am,* 2011 Dec 21; 93 (24): 2287 -2293. http://dx.doi.org/10.2106/JBJS.J.01727

25. GHTF/AHWG-GRM/N1R13:2011 - GHTF Regulatory Model: Available at: http://www.imdrf.org/docs/ghtf/final/steering-committee/technical-docs/ghtf-sc-n1r13-2011-ad-hoc-regulatory-model-110413.pdf

26. Krucoff M, Sedrakyan A, Normand ST et al. Bridging Unmet Medical Device Ecosystem Needs with Strategically Coordinated Registry Networks: An Overview of Recommendations of the National Medical Device Registry Task Force. *JAMA. 2015;314(16):1691-1692:* Available at: http://jama.jamanetwork.com/article.aspx?articleID=2431082

27. U.S. Food and Drug Administration. Center for Devices and Radiological Health, Strengthening Our national System for medical device Postmarket Surveillance: Available at: http://www.fda.gov/downloads/AboutFDA/CentersOffices/CDRH/CDRHReports/UCM301924.pdf

28. Council for International Organization of Medical Sciences: Available at: http://www.cioms.ch/index.php/publications/available-publications?task=view&id=27&catid=54

29. Nakatani T, Sase K, Oshiyama H, et al. Japanese Regisrty for Mechanically Assisted Circulatory Support (J-MACS): First Report. (Manuscript in submission).

30. Kirklin JK, Cantor R, Mohacsi P et al. First Annual IMACS Report: A global International Society for Heart and Lung Transplantation Registry for Mechanical Circulatory Support. *J Heart Lung Transplant* 2016;35:407-412. Available at: http://www.sciencedirect.com/science/article/pii/S1053249816000498

4.0 Definitions

Clinical evaluation: The assessment and analysis of clinical data pertaining to a medical device to verify the clinical safety and performance of the device when used as intended by the manufacturer (GHTF/SG5/N1:2007).

Lifecycle: All phases in the life of a medical device, from the initial conception to final decommissioning and disposal (ISO 14971:2007).

Medical Device Registry: Organized system with a primary aim to increase the knowledge on medical devices contributing to improve the quality of patient care that continuously collects relevant data, evaluates meaningful outcomes and comprehensively covers the population defined by exposure to particular device(s) at a reasonably generalizable scale (e.g. international, national, regional, and health system)". We think that such revised definition might better reflect the use of medical devices registry data to increase the quality of medical care.

Signal detection: The process of determining patterns of association or unexpected occurrences that have the potential to impact patient management decisions and/or alter the known benefit-risk profile of a device.

5.0 Vision

I. We envision international harmonization of medical device registries analytical methodologies via international Coordinated Registry Networks (iCRNs) based on demonstrated best practices;

II. While not all countries will contribute registry data to every device evaluation, all countries will benefit from the global collaborative;

III. The collaboration should be based on a systematic agreed upon process for sharing and evaluating data/findings from medical device registries amongst;

IV. All registries will agree on pre-specified analyses and collaborative sharing of the outputs with each other and the regulators;

V. A standing IMDRF registry working group should exist to facilitate this process.

6.0 Factors Contributing to Between-Country Variation in use and Outcomes Associated with use of Medical Devices

Several key characteristics contribute to differences among countries in both the use of medical devices as well as their associated outcomes. These include but are not limited to:

MARKET ENVIRONMENT: Because not all medical devices are available in all countries at the same time, the length of market experience will vary across countries. Moreover, medical device adoption will differ as a function of the extent of device reimbursement, the potential population size exposed to the device, and the number of medical device competitors currently in the market.

INTRINSIC AND EXTRINSIC ETHNIC FACTORS: Specific characteristics of the populations differ across countries, even for the same indication. Intrinsic factors include genetic information, body mass index, body composition, and other ethnic features; extrinsic factors involve aspects shaped by the cultural and behavior climate such as medical practice patterns, diet, and other environmental conditions. For example, the OECD reports that life expectancy at birth ranges from 56.8 years in South Africa to 83.4 years in Japan; infant mortality rates from 1.8 per 1000 live births in Finland to 40.9 in India (OECD, 2013) and in 2011, total fat in grams/capita/day from 87.3 in Japan to 171.5 in Austria. Additionally, extrinsic factors could influence some device performance outcomes more than others, such as patient reported outcomes (Wild et al., 2009).

REGISTRY CHARACTERISTICS: Key features that may vary across countries include granularity of data, degree of coverage or completeness of the market (e.g. full census, partial census, sample), duration of longitudinal follow-up, attrition rates, data privacy standards, regulation, ability and level of information exchange, and adherence to external standards (OECD, 2016). In addition there are well documented variations in consistency of data element terms and definitions, variation in data quality and the degree of use of standard data validated against master data sources. Recent example of MDEpiNet PASSION/RAPID project is a good illustration of how informatics principles were used to develop Common Data Elements (CDE)s which created an opportunity and link the CDEs to IMDRF CDEs, and integration of Device Identifier of UDI and standard GUDID data into the registries (Morales, 2016).

MEDICAL DEVICE REGULATION REQUIREMENTS: Requirements for assessment of clinical data in general show significant differences among major regions on a global level. For methodological principles applied to registry data these differences do not have a major impact as they can easily be implemented in the various processes by global regulatory bodies.

As an example, the demonstration of equivalence between medical devices in Europe will be subject to change as a result of legislative revision. This will provide greater detail with regards to the access to data and the clinical, technical and biological requirements needed to establish equivalence. This is likely to impact on feasibility and economic considerations in establishing registries to collect post-market data as a part of the overall conformity assessment.

"Grandfathering" as is applied in some regions is becoming less justifiable as a result of advances in technology and standards of care. As a consequence the vast majority of

devices will be required to produce clinical evidence on their own product. Registries can be an important source for these clinical data, if they are of "regulatory grade" i.e. they meet the requirements concerning data quality or governance and manufacturers take this into consideration when developing clinical trials designs and marketing strategies.

HEALTH CARE DELIVERY SYSTEMS: Differences in health care delivery systems could also be a contributor to between country variability. For example, when patients move from one health care system to another, the capture of their long-term data can be impacted. In addition, the differences in payment reimbursements within various health care systems can impact the availability of devises and the treatment options for the patients.

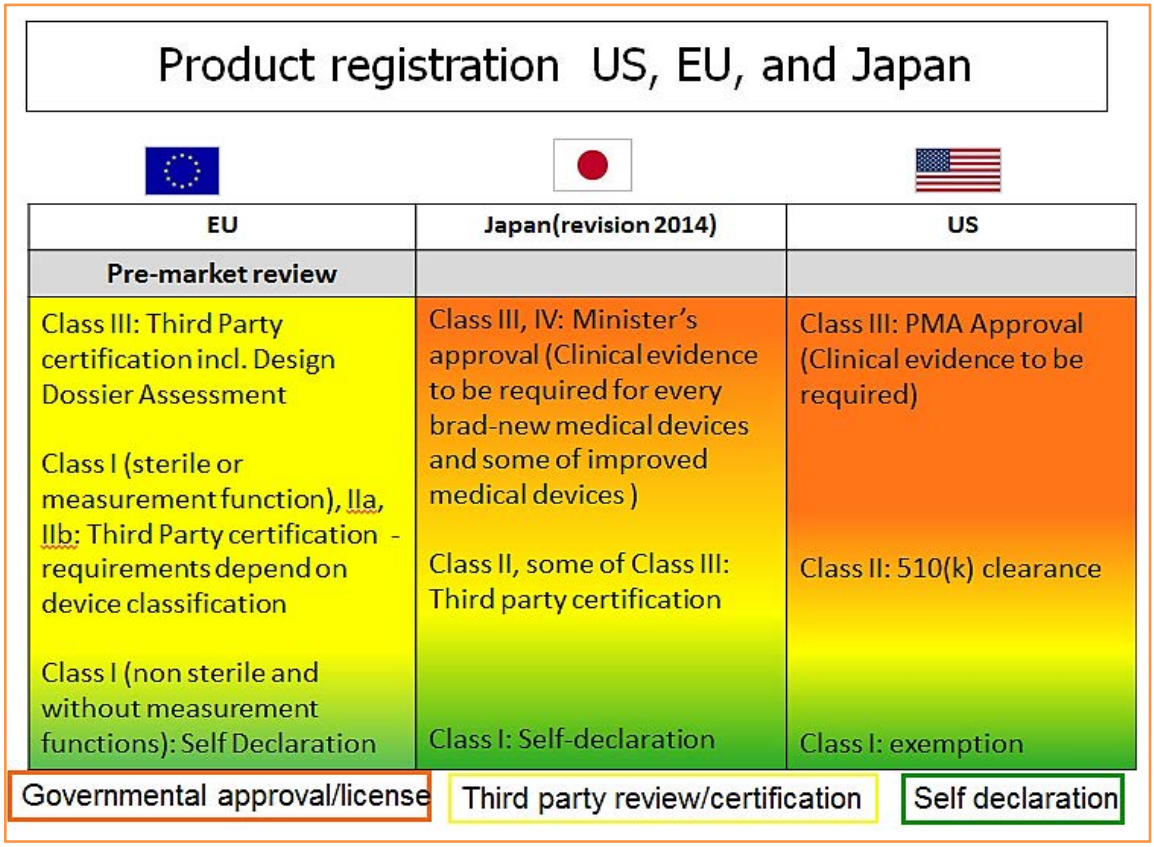

Table 1. Product registration US, EU, and Japan

7.0 Clinical Evaluation of Performance/Effectiveness and Safety using International Registry Data

7.1 **Context and Methodological Considerations**

Exploiting international registries will enhance the availability of evidence related to the total product lifecycle of a medical device (Figure 1). This document focuses on analytical considerations arising from leveraging information collected from around the world (Stage 3, Figure). We do not describe general approaches associated with confounding when clinical data arise from observational studies or other routine design features. We emphasize analytical issues related to the variation arising due to the inclusion of multi-national data.

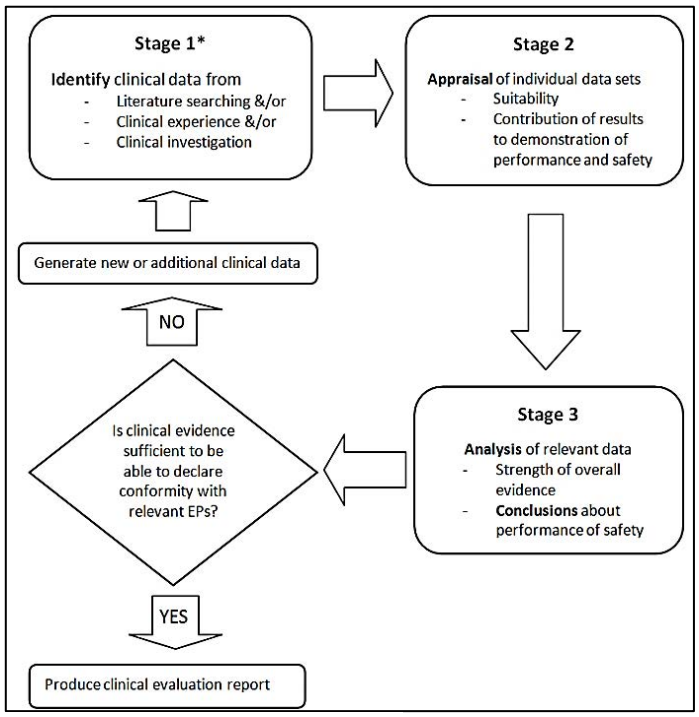

Figure 1. Stages of Clinical Data

HETEROGENEITY VERSUS HOMOGENEITY: The analysis of data obtained from country-specific registries involves combining heterogeneous medical device information with the goal of extracting performance and safety parameters. Even when data definitions and collection strategies are completely harmonized across registries, differences in device outcomes will likely persist due to both systematic and random between-country variation. Systematic factors include intrinsic and extrinsic features of region-specific populations. To the extent possible, higher priority should be given to biological response-related performance and safety outcomes rather than to outcomes more prone to social, life style, care delivery differences. However, even accounting for patient clinical features, and intrinsic and extrinsic factors, random, unexplained between-country differences will remain.

POOLABILITY: The extent to which registry-specific information is combinable or "poolable" will lie between two extremes: (a) no pooling and (b) (unadjusted) complete pooling without accounting for country-specific features. The *no pooling* option assumes that nothing can be learned about a medical device using data collected from another country about that device. The *complete pooling* option assumes that patient-level information from all countries provide information about exactly the same device effectiveness parameters, essentially treating all data as if arising from a single country. Both extremes (in their pure form) are unlikely; however, there are some situations in which pooling may be closer to one of these extremes or the other. There could be interactions unique to one or two countries that would prohibit pooling. In the case where there are limited registries for the given condition/device then the extremes would apply. Non pooling of the data but looking at each countries registry data alone as a stratified analysis may be useful as well.

EXCHANGEABILITY: Assuming information arising from international registries is poolable, the type of statistical dependence among the observations within and between the registries must be determined based on assumptions. Exchangeability is the degree to

which observable medical device outcomes are similar, and consequently, describes the dependence in a probability distribution of all the outcomes. Assuming independence within registry conditional on a country-specific effect is a reasonable initial assumption for combining international registries data. This implies, for example, assuming that revision rates following implant with a particular total hip device, adjusting for patient-specific confounders within the i^{th} country, has a country-specific mean rate μ_i, implying that the rates could differ across countries after adjusting for patient differences. To acknowledge this variation, it is assumed that the country-specific rates arise from a distribution with a common mean and variance. Thus, the country-specific rates are completely exchangeable across countries. The relative magnitude of the between-country variance in the outcomes to the total variation in the outcomes quantifies the degree to which estimates in one country "borrow" information from the ensemble of countries. As a framework in which to perform these calculations, we could extrapolate data based on similarities (or differences) between country specific factors. Such an approach would be facilitated if we had access to a metric for assessing how similar the data was across countries in question.

Nesting Trials in Registries

SAFE PCI

The Study of Access site for Enhancement of PCI for Women (SAFE-PCI) prospectively compared radial access and femoral access in 1787 women undergoing PCI. It was the first registry-based randomized trial conducted in the U.S. The authors indicated that "..this registry-trial infrastructure provided several efficiencies in site selection, data collection, and site workload (Sunil et al, 2014)."

MINIMUM NECESSARY DATA FOR ANALYSIS: Every effort should be given to adherence to minimum necessary standards for personal information in order to mitigate risk. This principle should be translated to the agreements on necessary analysis, and the data needed for those calculations. For example, learning curve-associated revision rate may just need time-to-event and sequence of procedures, but a similar analysis based on real-time events (recall, introduction of new devices, new publications, etc.) would need exact dates. Therefore, in the harmonization of endpoint analyses/definitions, risks to exposure of private information must be mitigated, even more than the general prohibitions against transfer of private information over national borders.

7.2 **Methodological Opportunities**

The methodological features of international medical device registries provide numerous opportunities to learn continuously about device performance through the product's entire life-cycle. These opportunities are afforded by the degree of similarity or exchangeability in the data. Essential performance evaluations include assessments of long-term medical device outcomes, comparative effectiveness estimates generally, and for performance evaluations involving rare outcomes in particular. Specific examples follow.

CHARACTERIZING LEARNING CURVES: Temporal variation in market entry dates of medical devices and geographic variation in characteristics of medical device users across

countries will reduce uncertainty in time to "steady state." This benefit is afforded by the availability of multiple opportunities to observe "first" use of new medical devices across different countries in relation to a subsequent broader set of medical professional users. Understanding learning effects could potentially influence future training requirements worldwide.

WIDENING INDICATIONS: Variability in both the intrinsic and extrinsic ethnic features associated with international registry data will undoubtedly include information about device performance in different indications. Such populations will arise due to geographic variation in market entry date of devices as well as to local practice patterns. While any one registry may lack sufficient numbers of patients to characterize medical device performance adequately, pooling country-specific registry information can increase the effective sample size for such populations, thereby reducing the uncertainty of device outcomes in new populations. Clinical data from countries with on-label populations can be leveraged to inform other countries. However, approvals in different countries may have different indications for use (e.g. different intended use populations (e.g., disease, race, sex/gender) as well as different intended uses (treatment, adjunct, relief of symptoms) in the approved labeling. These factors can also complicate the pooling.

DETERMINATION OF OBJECTIVE PERFORM ANCE CRITERIA & PERFORMANCE GOALS: Like widening indications, variability in intrinsic and extrinsic factors will enhance the creation of objective performance criterion (OPC) for medical device. By having more variety in patients one is more likely to capture the true underlying variance in the effect measure, leading to greater accuracy of the effect measure. In addition using samples from multiple countries allows for a greater pooled sample size and types of patients to be included in the analysis. O'Malley and colleagues demonstrated the construction of OPCs for bare metal coronary stents in different patient types (O'Malley et al., 2003).

JOINT ACTION PARENT

The EU- -funded Joint Action PARENT (Patient Registries Initiative) undertook an analysis of exchangeable neck hip arthroplasty implant. One goal was to identify potential limitations for a regular multinational signal detection network in a pilot project.
Main topics identified were:
Potentially relevant datasets should be identified in advance to avoid delays in case of an incident. Basic information on potentially relevant data sources should be available (e.g.., contact data, basic information on pathologies and devices covered, outcome parameter). More information available at:
Methodological guidelines and recommendations for efficient and rational governance of patient registries.

IDENTIFICATION OF SUBGROUP EFFECTS: In pre-specified subgroups, it is reasonable to assume that information about medical device performance in a particular subgroup of interest is related (but not identical to) information about medical device performance in other subgroups. For example, the comparative restenosis rate for a particular drug eluting stent relative to another stent may differ among diabetic patients with ST-elevated MI, diabetic patients with a non-ST elevated MI, non-diabetic patients with ST-elevated MI, and non-diabetic patients with non-ST-elevated MI but these rates should be related in some way. If a particular country has small numbers of patients within particular subgroups, borrowing information from similar subgroups from other countries will increase the precision associated with each particular country's subgroup estimate. In addition to examination of pre-specified factors, the availability of international registries will support hypothesis generating subgroup effects through the use of newer big data methods (Wang et al., 2015).

NESTING RANDOMIZED TRIALS IN REGISTRIES: In addition to analytical strategies associated with international registry data, access to such registries facilitates faster accrual of subjects to participate in clinical trials, thereby shortening the duration of the trial. The infrastructure available within a registry may also be used to identify subjects to participate in comparator arms in experimental studies. Efficiency gains could be realized through statistical matching or other design strategies by analyzing data found in registries prior to randomization (see SAFE PCI).

COMPARATIVE EFFECTIVNESS APPLICATION: International Consortium of Orthopedic Registries (ICOR) is an example of using distributed health data system with harmonized data definitions and data extraction followed by combing the data using innovative methodology across multiple national orthopedic registries. The coordinating center communicates with registries that apply standardized SAS syntax to their data and send summaries from each registry to coordinating center. This structure of the system as a decentralized distributed network helps overcome issues related to security, operations,

IDEAL (D) Recommendations related to registries

Stage 0 (pre-clinical)	Standards for publication/registration of preclinical data need to be established for devices. Many devices enter the market after this stage (see stage 4).
Stage 1 (first in human)	Reporting of first in human use integrated into a process by which Devices are patented and regulated. This can be the start of Registry for breakthrough technologies such as trans-catheter valves.
Stage 2a and 2b (prospective developmental and exploratory studies)	Device iterations mostly occur at Stage 0-1, but problems with device insertion/activation may require iteration. Quality control and learning curve estimation are important. Continuation of registry form stage 1 for breakthrough technologies or initiation of new registry for existing technologies can help initiate these studies. Studies ideally conducted in experienced centres to minimize risks of harm.
Stage 3 (assessment via RCT or alternatives)	Registry is an ideal infrastructure for initiating clinical trials: "Nested" RCTs possible within registries. Regulators should reach consensus on an international set of principles for deciding when an RCT should be required.
Stage 4 (long-term study)	Registries valuable and may begin much earlier (see stages 1-2. This is particularly important for "nth-of-a-kind" products that enter practice after stage 0. For first-of-a-kind devices, registries should ensure controlled introduction.

legal, and those related to patient privacy. It has major advantage of strengthening estimation by borrowing information from multiple registries. The analytic method of ICOR to combining survival curves is a flexible and robust approach to comparative effectiveness as it allows evaluation of risk change over time, determination of interactions and the risk factors (see IMDRF Registry Essential Principle).

RECENT TOTAL PRODUCT LIFECYCLE (TPLC) APPROACHES WITH STRONG EMPHASIS ON REGISTRIES. One of the approaches to TPLC is the IDEAL-D framework (Sedrakyan et al., 2016). The IDEAL(D) framework builds on prior efforts by an international expert consensus group that initially described the studies and reporting requirements for developing evidence for new surgical procedures: starting form first in man through widespread adoption. There are 4-5 stages within this Framework (IDEAL stands for Idea, Development, Exploration, Assessment, Long-term study). The Idea stage focuses on the "first-in-man" use of new technology. In the next Development stage, inventors modify the technology and in the Exploration stage other users get involved addressing technical details, indications, operator learning curves and quality control. In the Assessment stage operators collaborate on a definitive study of the new technique. Finally, Long term study (Stage 4) is needed to detect late and rare side effects, "indication creep" and performance variation. The application of this realistic framework for devoices as IDEAL(D) has a strong emphasis on registries (see the box).

8.0 Signal Detection

8.1 Context and Methodological Considerations

Single and aggregate reports and 'root cause analyses' are useful for identifying unexpected major harms. For example, the ASR artificial hip failure was recognized by MHRA in collaboration with clinicians based on case series reports with unique failure features (Medical Device Alert). However, only systematic processes will ensure continuous evaluation of implants to determine comparative performance and differences between them. Many important considerations, such as comparisons of rates of events between distinct sets of devices, are best addressed on the basis of summary measures rather than by informal aggregation of individual anecdotes. By shifting the focus from individual reports towards systematic summary analyses, we can exploit the power of registries to detect strong signals. For example, the Australian National Joint Registry found higher rates of implant revision for the entire class of metal on metal implants particularly those that are larger than 36mm (Australian National Joint Registry).

Registry fitness for use in the regulatory setting would depend on the type of the registry. For example, some registries are case-based and could not provide the rates in the absence of complementary data). However, most registries that would fit the IMDRF definition of medical device registry are very good data source to provide rates of events (IMDRF Registry Essential Principle). Registries can also help overcome regulatory limits related to sample size requirements for legacy post approval studies by allowing infrastructure that is already embedded in the health care delivery system to serve as a venue for addressing important regulatory questions, thus obviating the need for stand alone, large postmarket cohort and offering an opportunity for effective, less costly nesting of premarket data collection.

DELAY IN SIGNAL VERIFICATION: There is potential delay before information from international registries will be summarized, vetted and discussed with the regulator. Still, from the efficient regulation perspective the summary information provided by registries provide more complete picture than anecdotal and potentially biased information available immediately from single reports or passive surveillance systems.

DEVICE and COMPONENTS ISSUES: In the instance of implants with multiple components that can be used to create 'custom' implants, the international registries will be able to determine the 'mix and match' process and allow for investigations of how components impact both effectiveness and safety.

There are limitations to a registry approach that need to be recognized. For example, in the field of orthopedics, the device removal or replacement might not happen due to advanced age, comorbidities, patient refusal, or financial burden on patients. In addition much more extensive surgery might occur due to unavailability of certain components due to recall but this information may not be captured in the registry.

In addition the lack of use of a standard device identifier (the DI of the UDI) that can be validated against a vetted data source (in the US it would be GUDID) and the lack of a standard codeset to identify devices impacts the analysis of device outcome information both within and across countries

Signal Detection

J-MACS

Japanese Registry for Mechanically Assisted Circulatory Support (J-MACS) is a prospective registry harmonized with the United States registry (INTERMACS), now evolving into the network of international registries (IMACS). One of the examples of its signal detection is "transition to back-up mode", pointed out by the J-MACS event adjudication committee. The manufacturer has immediately taken actions with the regulatory agency (PMDA), to find out the root cause (open circuit of the driveline) and fixed it.

8.2 **Methodological Opportunities**

HARMONIZATION OF TERMINOLOGY: For a purpose of IMDRF convergence it is important to employ a consistent methodological vocabulary and set of processes for the implementation of signal detection. For example, 'outlier performance' is a term often used by international registries to describe a signal detection process that many regulators consider. From a process harmonization perspective the signal detection process can pre-specify actions that regulators will take for known issues or concerns. We recognize that harmonization of terms and processes is evolving over time, and we may need to consider updating data analysis, data collection and leverage linkages to other data sources.

PERIODIC UPDATES OF DATA CAPTURE: Registries should be flexible enough to allow for periodic update of data capture driven by the gaps in evidence. For one example, in the instance of hip replacement, over time new events such as soft tissue reaction and peri-prosthetic fractures were initially identified outside of the context of the registry. Based on those findings, the additional data field was added to the UK NJR

registry to learn more about soft tissue reaction. In addition, the linkage to other data sources allows better examination of the peri-prosthetic fractures (sometimes captured in a separate registry. These patient outcome terms and associated device problem terms should be codified and included as common data elements in recognized standard vocabularies available for use across National and International health IT systems, including patient records and registries (Registry Essential Principles).

For another example, after randomized controlled trials and evidence evaluation from registries, the Transcatheter Aortic Valve Replacement (TAVR) procedure received widespread adoption with several hundreds of thousands of inoperable and high risk patients treated with the procedure worldwide. Several years post introduction, international registries began identifying an increased incidence of valve hemodynamic deterioration (VHD) and decreased valve leaflet mobility in a subset of patients receiving valves possibly associated with both TAVR and SAVR. The initial risk factors identified included lack of coagulation therapy, larger BMI, and smaller size valve (Del Trigo et al, 2016). These international registry findings helped direct further research to better identify the risk factors associated with use of these devices.

Note that while we may need to add fields to registries periodically to address emerging questions, we should also consider when such updates can provide for removal of data fields for questions that have been adequately addressed. These data elements should be part of core recognized data standards and data code sets. Such an addition will make it easier to maintain a sustainable system.

POSITIVE SIGNALS: When viewing signal detection as a method to highlight positive as well as negative findings, and for use with efficacy and safety measures, it clarifies the possibility for the use of signal detection to contribute to benefit/risk assessments. Such assessments would fit naturally into a signal detection framework, especially with respect to determining benefits that became more apparent as additional data were made available over time from a variety of sources, such as that assembled from multiple national registries. It is straightforward to see how the process of exclusion (discussed later in terms of signal detection taxonomy) could be used to rule out specific levels of harms of interest. However, one could also leverage the same machinery to rule out lower values of efficacy (so that performance was demonstrated to be higher than a particular threshold or value). Such findings could potentially be used to update changes in labeling for a device that reflected the refinement in knowledge of performance of the device.

NEW VS MATURE DEVICES: There may be distinct considerations at play when examining a registry of an early innovation device versus one of established devices (as might become available with the proliferation of a given technology). Concerns for signal detection may readily differ depending on the relevant maturity of devices in the market place. For example, we might be more concerned with short term quantification of harms for first in kind devices, but be more concerned with potential time shifting effects (such as unanticipated results stemming from minor iterative changes) for well-established devices. With a larger class of related devices, we have the opportunity to look for differential performance, but this is not even possible for first in kind devices.

CAMOUFLAGING**:** This effect is an important consideration when evaluating devices. Considering class/attribute level versus the individual product level evaluation is a matter of debate. In general, evaluating a device on the individual level is complicated due to sample size limitations and might not be aligned with the philosophy of product development. Minor variations in implant design are often performed for marketing purposes to offer 'diverse' products while from performance perspective they are all comparable. Hence, class/attribute level assessments are the first step in signal detection process. However, in some instances the incremental changes that are considered 'benign' might lead to a much higher rate of device failure for the specific iteration of the device. Missing these effects within attribute level comparison is the essence of the camouflaging problem.

This effect illustrates the importance of performing signal detection (outlier) analysis at the level of device identifier model/size cluster level in addition to implant attribute (classification) level. This process also helps to revise attributes and come up with new ones based on real life experiences.

PROVIDER EFFECT: Provider effect is another key consideration and needs to be taken into account to make sure effects (camouflaging, class) are not limited to a few generally underperforming surgeons (e.g. volume). The registries, regulators and professional societies should set some criteria to parse out when it is a provider vs device effect. Specifically if there are known extrinsic or intrinsic factors that impact the success of procedures/devices, this may need to be some weighted or adjusted threshold.

PRESPECIFICATION OF THRESHOLD VALUES: It is key that pre-specification of threshold values for various signal detection methods is provided, in particular when planning for sharing of information from analyses conducted in parallel on registries from different countries. This will enable meaningful and timely sharing of potential signals across data sources. Pre-specifying the risk level for signal detection based on negotiation or accumulated strong evidence (e.g. OR, RR, HR 2.0) helps to achieve harmonization. There might be different requirements set for early entrants versus established products in the market, component versus entire device replacement. The disparate methods themselves may also be better served by employing different types of thresholds. For processes that are relatively stable and are based on all available information up to a certain time point we may be less tolerant of sharp variation than we would be for other processes that only considered the much smaller set of data available within a moving window of time (such as the most recent two months). Finally, there might be different threshold levels for different applications. Note that agreement of a common threshold value in a particular area here is advocated explicitly to provide a sharing of comparable information; it is not intended that such a shared threshold value must be adopted uniformly by all participating regulatory authorities for their own regulatory actions.

An additional point about threshold values is that they very often are considered in the context of a frequentist testing scenario where one might be concerned with the probability of controlling the error rate across an entire experiment (such as type I error). This familiar framework can breakdown in cases of long-term surveillance, in that the "experiment" across which one is attempting to control error rates is poorly defined. In these scenarios, there are strong arguments for considering a shift to a Bayesian paradigm where one instead

might focus on a continually updated posterior probability (of benefit or harm, say), rather than on the p-value from a multiply-repeated testing procedure.

8.2. Examples of Tools that are used in Registries

8.2.1 Cumulative Sum of Outcomes (CUSUM) Methodology

Cumulative sum of outcomes (CUSUM) methodology allows determination of excessive rates of failures or adverse events of implants. CUSUM is a sequential statistical analysis methodology with graphic application. It allows on line identification of changing device failure or surgical complications. There are various methods in use today. For example, a likelihood-based scoring method of calculation of CUSUM is use d by the Scottish \Orthopedic registry described as part of ICOR series (Sedrakyan et al., 2015). If the device failure rate is close to or below average, the CUSUM will remain close to zero. Outlier device or surgeon status is identified at the point set in advance and is named the prediction limit. Setting the statistical thresholds at agreed upon levels helps to balance the risk of failure against that of false alerts. Setting a prediction limit is not an exact science, and changing the statistical criteria will change outlier identification. It is a classic example of the tradeoff between sensitivity (finding potential issues) and specificity (avoiding spurious findings). Hence the results should be interpreted as a signal that does not yet mean a bad implant or a bad device. One of the advantages of CUSUM method is ability to track both surgeons and introduction of the new device to evaluate the surgeon-device 'package'. For example CUSUM allows tracking of outcomes of high volume surgeon with changes in practice over time and determination periods of outlier performance that was in fact linked with introduction of new implants (see figure from Scottish Arthroplasty Registry- part of ICOR series) (Macpherson et al., 2011).

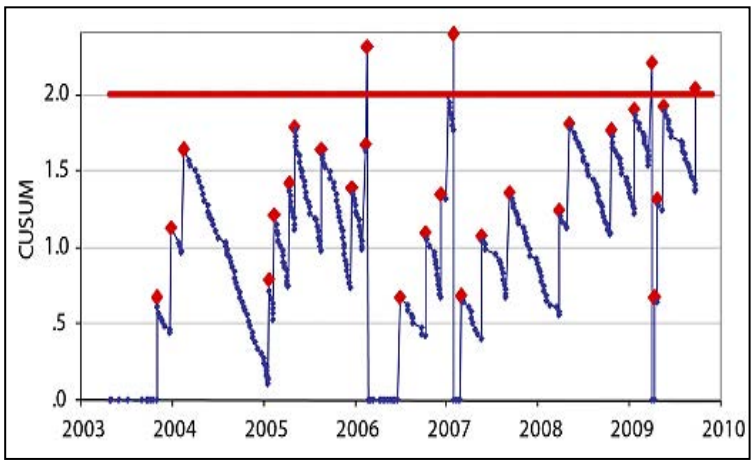

Figure 2. CUSUM chart for a high-volume surgeon with a complex practice, showing an increase in outliers corresponding with the introduction of a new implant (Macpherson et al., 2011).

8.2.2 Cumulative Revision Rate over Time

Depicting an unadjusted **cumulative revision rate over time** after implantation of the device is a simple but powerful technique allowing identification of outlier implants when compared to overall or group average. The method also allows calculation of accompanying 95% pointwise confidence intervals using various methods. For example, the Australian orthopedic registry process identified the ASR hip as outlier device using this method followed by Cox proportional-hazard modeling to calculate the hazard ratios and adjust for age and sex in order to conduct a comparative analysis of revision rate between groups (de Steiger et al., 2011).

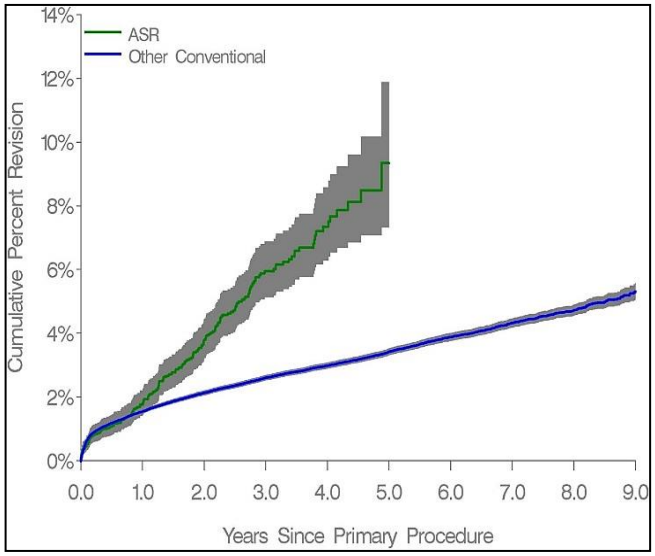

Figure 3. Cumulative revision rate of primary conventional total hip arthroplasties (de Steiger et al., 2011).

8.2.3 Funnel Plots

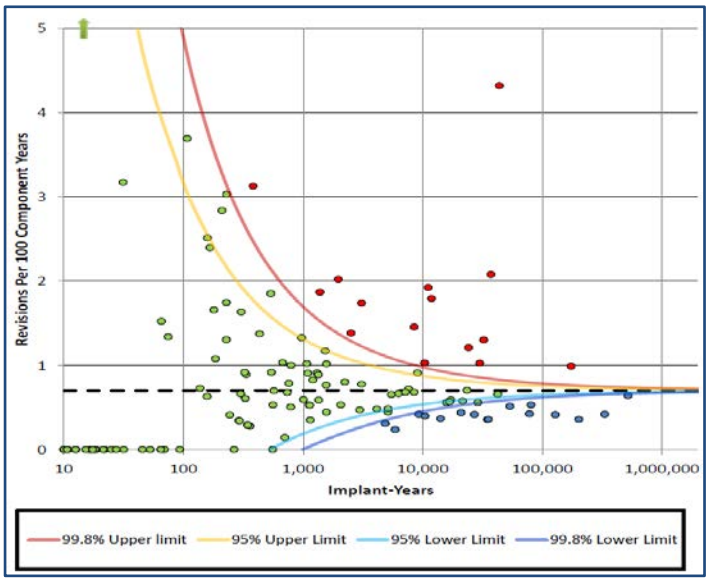

Figure 4. Revision Per Hundred Component Years

Another graphical approach (Figure 4.) is that of **funnel plots**, which are based on application of Shewhart Charts in medicine (Shewhart, 2012).They are example of Statistical Process Control (SPC) methodology with valuable graphical display to showcase the distribution of the data. Through the use of funnel plots, it is possible to compare the observed events (e.g. specific device failure) against the national average within the population (Spiegelhalter, 2005).

In the hypothetical Figure 4 from the UK National Joint Replacement registry each circle represents one device and the x-axis denotes the number of device implants combined with number of years followed up for a particular device (or volume) tracked by registry for that device. The y-axis represents the "true" event-rate (unobserved). Devices falling above 95% or 99.8% control limits (set in advance) for risk are deemed as outliers. For various "true" event rates around the gold-

The Data Extraction and Longitudinal Trend Analysis (DELTA) approach is a variation of CUSUM approach also offering real-time medical device safety surveillance. For example, DELTA methodology has been successfully applied retrospectively and demonstrated the feasibility of an early warning detection system for faulty Fidelis ICD leads. The DELTA network was utilized in a prospective propensity-matched cohort analysis of 7 newly-introduced cardiovascular devices, using clinical data captured in the Massachusetts PCI database from 2003 to 2007. For this project, the NCDR CathPCI registry was used as the data collection tool. The DELTA system identified issues in 3 out of 21 safety analyses that triggered sustained alerts in 2 implantable devices. Patients receiving a Taxus Express2 drug-eluting stent experienced a 1.28-fold increased risk of post procedural myocardial infarction (2.87% vs 2.25% for those receiving alternative drug-eluting stents) (Recommendations for a National Medical Device Evaluation System).

standard rate, funnel plot shows which devices can be called "outlier". Details of the calculation of funnel plot values for standardized ratio data are as follows:

Assume a standardized ratio SR = O / E based on an observed count O and expected count E, where E is defined as number expected on the basis of the product of national device failure rate and the patient time at risk. We assume an expected or target ratio t. A square root transformation is applied to both the standardized ratio (y), and the target (t):
$y = \sqrt{O}\ (O/E)$
$t = \sqrt{O}(t)$
The standard error(s) is given by $s^2 = [1/(4*E)]$
Thus, the unadjusted transformed Z-score is: $z = (y - t) / \sqrt{(s^2}$

Like several other methods, this approach is heavily dependent on assumptions about equivalent underlying risk. If there is heterogeneity in the underlying risk (as might occur with differing standards of care across sites, differing expertise of operators, or differing disease progression among patients between sites), then departures outside the limits may be more reflective of issues with the assumptions broadly, than with issues in the performance of those points outside the limits. Also, for funnel plots in particular the performance may be somewhat suspect for extremes in the x-axis. Values close to zero (very little use) have extreme variability, while high values of the x-axis can contribute an excess proportion to the determination of what constitutes an outlier (since that determination is based on all of the data).

When considering how to screen registry data across sources in a coordinated fashion, it can be very helpful to think about what patterns of potential signals are of interest, and what data might be made available to investigate those patterns. Signal detection approaches can generally be categorized into four broad groups: separation, heterogeneity, exclusion, and deviation. This taxonomy for signal detection was previously introduced in the white paper "Recommendations for a National Medical Device Evaluation System".

Table 2. A Taxonomy for Signal Detection

Type	Purpose	Example
Separation	Identify divergence between two devices	Is the adverse event rate following drug eluting stenting different from that following bare metal stenting?
Heterogeneity	Determine if and when one process differs from a collection of processes	If and when does the average post-implant infection rate for Surgeon A differ from the average infection rate for all surgeons in the country?
Exclusion	Determine when a signal is sufficiently refined that a threshold value may be excluded, even if the process is relatively constant	When does the average hospital mortality following implantation of a left ventricular assist device in Hospital A exceed 15%?
Deviation	Determine if and when a single process leaves a pre-defined area of acceptability	If and when does the incidence of inappropriate shocks by implantable cardioverter/defibrillators leads exceed x?

Note that differing signal detection approaches may allow one to compensate for differences observed across separate countries. For example, it may be that different reporting requirements might be expected to yield differing absolute rates of reported events for the same device across different countries. However, if the proportional difference in reporting rates was constant on an additive scale (possibly obtained via log transformation if it was originally constant on a multiplicative scale), then the difference in rates between two competing devices might be expected to be comparable across multiple countries. In this example, separation might be more justifiable than heterogeneity, as the later might be more reflective of country differences than true device differences.

Alternatively, one could conceive of scenarios where aligning around a similar signal process could necessitate application of differing (absolute) threshold values. For example, suppose that differing countries had differing base mortality rates, but they were interesting in comparing departures from the (separate) base mortality in a consistent fashion. One might construct a monitoring plan that allowed for a departure of plus or minus 5% from baseline mortality, so that the width of the acceptability corridor was constant across countries even when the underlying rates differ. This would be a variant of a signal detection process that has previously been used to examine deaths as linked to numbers of operations.

9.0 General Recommendations

9.1 Recommendations regarding international coordination in methodology that would add value to multiple international stakeholders including regulators include:

a. Leveraging IMDRF work already in progress (unique device identification adoption, creation of standard common data elements, defining a code set for patient and device problems associated with adverse events) to reduce variation in the data being exchanged between registries would improve data analysis accuracy and signal detection;

b. Advancing coordination in addressing important questions that are difficult to resolve within a single country registry;

c. The proposed international methodological pilots could be a vehicle for further convergence of methodological approaches

9.2 Recommendations regarding methodological principles in clinical evaluation of performance, effectiveness and safety across the device lifecycle, including signal detection, using international Coordinated Registry Networks (iCRNs) include:

a. The process should exist by which important information and data (on either a summary level or observation level) will be shared in a structured fashion by regulators across multiple countries. This process should be agreed upon before analyses are performed.

b. Where appropriate, registry structure should be leveraged to efficiently answer questions that would have historically been addressed via more resource intensive legacy tools (e.g. 522 studies in the US, PASS studies in EU);

c. Registries should be exploited to facilitate the conduct of clinical trials both premarket and postmarket;

d. Separation within and between country variation for analysis is necessary in order to ensure effective individual and international decision making**.** Explicit modeling to help determine factors influencing the within- and between-country variability would be useful;

e. Pre-specification of analyses that could drive regulatory decisions is essential. Beyond the direct specification of analyses, effort should be devoted towards construction of a verification and reproducibility plans for findings from the analyses driven by models;

f. Further consideration should be given to assessing optimal role of spontaneous adverse event reporting in the context of iCRNs.

10.0 Pilot Projects

The following methodological pilot projects could be a vehicle of addressing important regulatory questions:

a. Pooling Data for Regulatory Decisions in International Coordinated Registry Network (iCRN)

b. Statistical Approaches for Informing the Device Total Product Life Cycle Internationally

c. Piloting PASSION/RAPID Global Case Report Form for Peripheral Vascular Devices in International CRN as Infrastructure to Nested Clinical Trial

Pooling Data for Regulatory Decisions in International Coordinated Registry Network (iCRN)

Methodology-specific pilot: theoretical derivations, simulation-based summaries, and empirical approaches to characterizing the validity of pooling assumptions and the coherence of comparisons, determination of a minimum number of observations required, and approaches to representing uncertainty of the strengths of relationships in the context of label extensions, signal detection, and clearance of predicate devices.

1. Disease/device focus	Applicable to any condition or device (see the examples of proposed priority areas)
2. Immediate research question(s)	What is the validity of pooling assumptions made in the context of international CRNs? What types of devices and populations can be compared? What is the minimum number of observations required for label extensions or clearance of predicate devices? How can uncertainty of the strengths of relationships be best represented? How can big data techniques (e.g., data mining, machine learning) be utilized for signal detection?
3. Stakeholders engaged	Regulators, industry, academia, patients, payers, patient representatives from example device area and iCRN component registry owner representatives.
4. Existing international resources leveraged	Methodology illustrated using existing international registries such as ICOR, ICVR, ICOBRA etc.
5. Efficiencies	Study results will indicate how to develop more efficient

promoted	(statistical efficiencies) estimates for regulatory inferences.

Statistical Approaches for Informing the Device Total Product Life Cycle Internationally

The availability of international registries may allow for shifting some premarket device data collection requirements to the postmarket setting, this shift requires the use of valid and reliable data elements that reflect the outcomes of interest in well-defined populations. Approaches for using international Coordinated Registry Network (iCRN) data to provide: (a) important long-term device performance information for mature devices; (b) solid intelligence to help improve the device; and (c) evidence on which patients are the best candidates for a device require assessment and illustration

1. Disease/device focus	Applicable to any condition or device (see the examples of the proposed priority areas)
2. Immediate research question(s)	How comparable are data elements and definitions between various data sources and clinical trials? Can patient reported outcomes be utilized to assess device benefit? How can stakeholder preferences be factored into the benefit/risk assessment?
3. Stakeholders engaged	Regulators, industry, academia, payers, patient representatives from example device area and CRN component registry owner representatives.
4. Existing international resources leveraged	Methodology illustrated using existing international registries and consortia such as ICOR, ICCR, ICVR, ICOBRA. Stakeholder utility banks could be constructed and leveraged for future device assessments.
5. Efficiencies promoted	Study results will indicate how to develop more efficient ways to assess data elements in various data sources including patient reported outcomes and preferences.

Piloting PASSION/RAPID Global Case Report Form for Peripheral Vascular Devices in International CRN as Infrastructure to Nested Clinical Trial

The MDEpiNet PASSION (Predictable and Sustainable Implementation of National) Registries RAPID program successfully developed data collection tool consisting of 100 well specified Common Data Elements (CDEs) including the Device Identifier (DI) pf the UDI and data elements from the US Global Unique Device Identification Database based upon IMDRF. Via transparent and well documented multi-stakeholder engagement this effort has established the international infrastructure that can be used for both premarket and postmarket clinical studies and surveillance.

1. Disease/device focus	Peripheral vascular devices (see the examples of proposed priority areas)
2. Immediate research question(s)	Can international CRNs be successfully used for nesting clinical trials? Demonstrate the ROI in utilizing the global case report form

		and international infrastructure?
3.	**Stakeholders engaged**	Regulators, industry, academia, payers, patient representatives from example device area and CRN component registry owner representatives.
4.	**Existing resources leveraged**	Methodology illustrated using existing national or international registries in peripheral vascular space
5.	**Efficiencies Promoted**	Study results will indicate how to develop more efficient estimates for regulatory inferences.

Note: Examples of proposed priorities where international coordination could be helpful include but are not limited to: anaplastic large cell lymphoma associated with breast implants; safety of fenestrated and chimney abdominal aortic aneurysm devices; evaluation of gender differences in hysteroscopic sterilization devices; differential revision rates associated with exchangeable/modular necks for modular hip replacements; and facilitation of nested clinical trials in the area of medical devices.