



Universiteit
Leiden

The Netherlands

Safety and performance of high-risk medical devices: the role of real-world data

Hoogervorst, L.A.

Citation

Hoogervorst, L. A. (2025, November 18). *Safety and performance of high-risk medical devices: the role of real-world data*. Retrieved from <https://hdl.handle.net/1887/4283146>

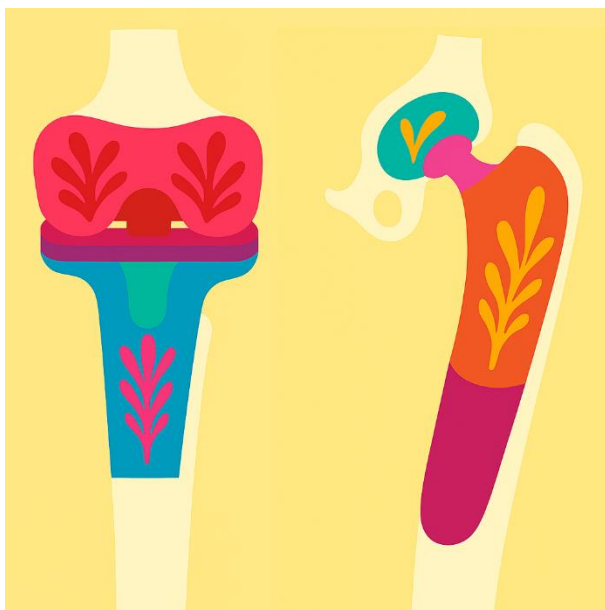
Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4283146>

Note: To cite this publication please use the final published version (if applicable).

Chapter 9 – Summary, implications for clinical and regulatory practice, and future perspectives



High-risk implantable medical devices such as hip and knee implants are essential for clinical care. Internationally, the use of real-world data sources to assess safety and performance of high-risk medical devices has increased in recent years as shown by the increased number of publications using registry and administrative data sources, aiming to strengthen the evidence of these medical devices. Real-world data have the advantage of typically including all (i.e. unselected) patients receiving a medical device rather than a selection of patients as enrolled in clinical trials.¹ Despite the increase of supporting evidence on performance of medical devices on the market, safety incidents relating to high-risk medical devices still occur.²⁻⁵ This raises uncertainty on the quality standards applied for approving these high-risk medical devices for market access.⁶

The aim of this thesis was to investigate how real-world data of implantable high-risk medical devices – particularly from arthroplasty registries and safety notices – can be used to i) ensure high-quality evidence regarding safety and performance of high-risk medical devices, and ii) develop methods to evaluate the safety and performance of high-risk medical devices. In this chapter, the main findings across all studies of this thesis are summarised with respect to these two aims, the implications of our findings for clinical and regulatory practice are discussed, and recommendations for future practice and research are given.

Quality of real-world evidence to assess the safety and performance of high-risk medical devices

The results from the studies in this thesis emphasise two key points with respect to the quality of real-world data and evidence on safety and performance of high-risk medical devices:

- Large heterogeneity exists in data collection as well as methodological characteristics of medical device (e.g. arthroplasty and cardiovascular) registries, which determine the quality of registry data as well as definitions.

Reporting these characteristics is crucial to interpret the data, and highlights the need for consensus on a minimum dataset to be (publicly) reported by registries.

- A multifaceted approach in safety signal detection is needed. Registries were shown not to capture all hip and knee implants with safety concerns, but neither did safety notices reports, thereby emphasising the importance for combining several data sources.

Large heterogeneity in registries reporting on characteristics that determine quality of data and analysis

Among the 26 European arthroplasty registries (i.e. hip and knee implants) and 20 European cardiovascular registries (coronary stents and valve repair/replacement) that were identified and reviewed, large heterogeneity was shown in structures of the registries, methodology for data collection, definitions of variables, methodology of analysis, as well as the level of public transparency related to these characteristics (**Chapter 2**). The lack of uniformity in definitions hampers the ability to pool data or compare data across registries. This finding reflects a broader challenge previously identified in studies examining differences between definitions used by registries. For instance, a study assessing the quality of cardiac registries across all subspecialties of cardiac care found that many registries provided explicit similar definitions for only a limited number of variables.⁷ Moreover, this study found that the quality of the registries was highly variable across registries, supported by using a validated registry grading system. For arthroplasty registries, similar findings were found in a previous study, which also identified “substantial heterogeneity in definitions regarding revision, readmissions, and complications, as well as methodology for analysis”.⁸ The study in **Chapter 2** shows that the lack of consistency across registries in data collection methods extends beyond different definitions of variables and outcomes, it also pertains to data validation and funding sources, which can influence

data quality and thereby hamper the ability to pool and compare data across registries.

Data completeness is essential for any medical device registry to ensure the representativeness of data for daily practice and to support clinical and regulatory decision making. **Chapter 2** shows that procedure-level completeness was reported by a slight majority (65%) of arthroplasty registries, with 11 out of 13 registries meeting the 95% completeness threshold recommended by the International Medical Device Regulators Forum (IMDRF).⁹ As for cardiovascular registries this was worse, as none mentioned procedure-level completeness. These findings are in agreement with previous studies.^{7,10} These results are important because failure to capture all patients receiving a medical device (i.e. lack of procedure-level completeness) may introduce selection bias.⁹

Next to procedure-level completeness, it is also important to capture all (relevant) outcomes. For instance, in orthopedics, failure to capture all revision procedures could result in underestimating the actual revision risk in daily clinical practice. Such an underestimation consequently provides false reassurance of safety and performance assessment for specific implants. Consequently, clinicians may be inclined to select these implants for their patients, assuming they perform well. Similarly, regulators may struggle to detect (early) signs of medical device failure or other safety concerns if data are based on a favourable selection of patients or incomplete reported outcomes.

The need for a minimum dataset to judge the quality of registry data

One of the key advantages of registries is that they aim to include the vast majority of patients receiving a medical device rather than a selection of patients participating in clinical trials. Nevertheless, the evidence generated from registry data can still be improved. The main improvements needed are: i) agreement on a minimum set of

variables that all registries should publicly report to provide information needed to judge the quality of registry data when they are used in medical device safety surveillance, and ii) consistency in definitions of variables and outcome measures across registries, to ensure valid comparisons of data between registries as well as enable pooling of data. The latter not only for research, but also to facilitate early detection of rare safety issues as well as adverse events in seldom used implants for specific indications.

Chapter 3 therefore conducted a Delphi study aiming to achieve consensus on such a minimum dataset to assess the quality of registry data and analyses needed to judge safety and performance of medical devices. Experts from different backgrounds were involved (i.e. healthcare professionals, methodologists, registry experts, regulators, and assessors of notified bodies) to define a minimum dataset during three rounds. The final set included 15 data quality items and 8 data analysis items. The experts considered “data completeness” as the most important data quality item and “the definition of the outcome” the most important data analysis item. This minimum dataset provides more specific guidelines for medical device registries to (publicly) report these items. As such, the dataset is intended to function as a standard reporting framework for registries. In addition, the items included in this minimum dataset were mapped to more generic principles of regulatory frameworks.¹⁰⁻¹² This resulted in a decision framework that can assist European regulators when assessing safety and performance of medical devices during market surveillance, as well as assist manufacturers when using registry data for post-market surveillance.

Existing frameworks have focused on usability and methodological principles for using registry data, based on international consensus among regulators, with input from experts (i.e. International Medical Device Regulators Forum (IMDRF)).^{13,14} Furthermore, the European Medicines Agency (EMA) and the Food

and Drug Administration (FDA) indicated relevance and reliability as key principles when using real-world data.^{10,15,16} In the United States, the coordinated registry networks framework aims to produce all the necessary evidence for regulators and other stakeholders by combining data from multiple data sources.^{12,17} The latter framework includes several items related to data quality, focusing on relevance, coverage, data completeness, and data verification. However, it misses several important variables, which are included in our decision framework (**Chapter 3**), such as reporting funding sources, as well as items related to analysis of data to evaluate safety and performance of the medical device such as defining outlier performance. Reporting on funding sources can help users to assess potential conflicts of interest that may influence data reporting and thereby quality of the evidence, where analyses focused on “defining outlier performance” provides a clearer criterion for identifying medical devices performing significantly worse than expected. Thus, our decision framework adds relevant information and specified a minimum dataset for each domain (data suitability for regulatory question, data governance, data quality, and data analysis) within relevance and reliability as guiding principles.¹⁸⁻²²

Implementing the minimum dataset covering these four domains (data suitability for regulatory question, data governance, data quality, and data analysis) will support the collaboration between registries, thus enabling comparisons across registries as well as data pooling. The International Society of Arthroplasty Registries (ISAR) has already taken the initiative to ask all of their member registries to report on their website several of the items in the minimum dataset, which is an important first step towards implementation.²³ The decision framework proposed in **Chapter 3** may already assist regulators when assessing the safety and performance of medical devices for market surveillance as well as manufacturers when using registry data for post-market surveillance. Future research may advance our framework by developing criteria on what constitutes sufficient and good quality

evidence, as registries can score low on certain aspects but high on others. Until then, regulators, assessors of notified bodies, and manufacturers are recommended to focus on the items deemed most important by experts when judging the quality of the data e.g. to rely on data from registries with at least 95% data completeness.

The need of a multifaceted approach in safety signal detection

While registries may signal problems related to some medical devices, for instance by implementing outlier detection procedures, **Chapter 4** explored the extent to which some safety concerns may be missed by focusing only on revision risk as the outcome of outlier detection. We showed that incorporating an additional data source (i.e. publicly released safety notices) identified medical devices with safety concerns that would not have been identified using registries' outlier data analysis alone: a significant proportion (approximately 25%) of total knee implants with safety concerns were identified uniquely by safety notices reports. This finding highlights that a critical limitation exists, when relying only on registries' outlier procedures for identifying safety signals of implants, as it will only identify those safety concerns resulting in revision surgery. Safety notices issued by manufacturers or published at Ministries of Health or regulatory agencies websites capture issues that registries do not capture, such as complications related to medical device components or wrong labelling of implant packaging, adverse events other than revision risk (e.g. patient reported outcome measures) for an orthopedic implant. Although, relying only on these safety notices would also miss safety concerns since about 20% of the knee implants with safety concerns were only identified by registries (i.e. outlier implants). This stresses the importance of integrating multiple data sources, enabling a more comprehensive and reliable approach to safety and performance monitoring of medical devices.

The International Medical Device Regulators Forum (IMDRF) defines safety signal detection as “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”.¹⁴ Safety signal detection for medical devices, such as knee implants, aims to identify potential risks related to the product's use as soon as possible. A review evaluating the current state of knowledge and examining the existing tools used for medical device safety signal detection found; i) description of safety signals (“coding”) is heterogeneous; ii) no agreement on preferred methods for signal detection exists, and iii) no golden standard for signal detection has been established.²⁴ Hence, the authors stated that a global dataset of medical devices should be created using automatic reports from national and/or regional databases. In the absence of such a database, an automated web scraper tool such as the Coordinating Research and Evidence of Medical Devices (CORE-MD) post-market surveillance tool could be helpful in gathering all safety notices for a class of medical devices and classify them according to a medical device problem.²² This CORE-MD tool was used to successfully identify a larger number of total knee implants with safety concerns than only using registry revision data (**Chapter 4**). Although this CORE-MD tool failed to identify all implants with safety concerns. This indicates that safety notices and registry outlier data measure different aspects of safety and performance of total knee implants, stressing the importance of a multifaceted approach combining multiple data sources.

The previously mentioned coordinated registry networks framework from the United States could be a potential pathway to achieve this multifaceted approach, requiring patient-level linkage of data.^{12,17} The recently published European Health Dataspace Act enables such a multifaceted approach, as it will: i) empower individuals to access, control and share their electronic health data across borders for the healthcare delivery (primary use of data); ii) enable the secure and trustworthy reuse health data for research, innovation, policy-making, and regulatory activities (secondary use of data), and iii) foster a single market for electronic health record (EHR) systems, supporting both primary and secondary use.²⁵

In **Chapter 4** the occurrence of a safety concern as the starting point for analysis is discussed, either from the perspective of safety notices or outliers in registry data. However, this does not answer the question what percentage of all total knee implants currently on the market do not have any safety concerns reported in any of these data sources. To evaluate this, 30 randomly selected hip and knee implants were evaluated for the presence of safety concerns: Ten implants did not have any safety concerns identified across the four different data sources (**Chapter 5**). No safety concerns were identified on manufacturers' websites for any of the implants, which could be due to commercial interests of the manufacturers, or agreement with other stakeholders to publish the safety concern. As discussed earlier, relying on a single data source results in the omission of some implants with publicly disclosed safety concerns, again underscoring the importance of a multifaceted approach integrating multiple data sources to ensure a comprehensive safety assessment.

Methods used to evaluate the safety and performance of medical devices

The performance of medical devices is a dynamic interaction between the surgeon (i.e. surgical skills and surgical teamwork), the medical device itself (e.g. materials), and the patient (e.g. comorbidities) receiving the medical device. These factors in combination will determine the overall outcome for patients. If any of these factors are suboptimal, they may lead to worse performance or even failure of the medical device. In Europe, regulators have the responsibility to ensure that medical devices with safety concerns are withdrawn from the market. To take action, any poor performance of a medical device must therefore be directly related to the medical device itself, its' design or the materials used, rather than other influencing factors such as surgical skills or patients' comorbidities.

Several methods are used to assess the safety and performance of orthopedic implants using registry data, which compare implants to: i) the best-performing implant; ii) the average performance of other comparable implants, and iii) absolute thresholds by using objective-performance-criteria.²⁶⁻³⁴ However, these methods are mainly applied within a registry rather than investigating whether a given implant consistently performs worse across all registries, which would be expected if the reason for worse performance would be caused by the implant itself rather than by other factors. In **Chapter 6**, we present an external validation of the most frequently used objective-performance-criteria in orthopedics (i.e. the Orthopaedic Data Evaluation Panel (ODEP)). Using data of nine national and regional orthopedic registries, it was shown that total hip implants with an ODEP-rating had lower cumulative revision risks than total hip implants without an ODEP-rating, suggesting that ODEP-ratings could be useful for implant selection. For total knee implants this could not be assessed since only 13% of all total knee implants reported by registries could be matched to an ODEP-rating. In **Chapter 6** is discussed that total hip implants with a higher ODEP-rating did not consistently have lower cumulative revision risks in registries than total hip implants with lower ODEP-ratings. The latter indicates that the ODEP-rating cannot be used to distinguish between better and worse performing total hip implants. Furthermore, it was observed that the assigned ODEP-rating varied widely across registries, i.e. total hip implants receiving an ODEP-rating based on the revision risk of one registry would not necessarily receive the same ODEP-rating based on data from another registry or based on the pooled revision risk across all registries. This variability implies that assigned ODEP-ratings do not necessarily apply to the performance of a specific total hip implant in different countries. This variation in performance (i.e. revision risk) may be explained by several factors, such as variations in patient demographics (i.e. case-mix), difference in indications for primary and revision surgery between countries, and variation in definition of revision across registries.^{8,35-40} This highlights the need

for transparency on data sources on which the ODEP-ratings are based. Also, since ODEP-ratings varied widely between registries, regulators and notified bodies should only withdraw implants from the market when they show consistently low ODEP-ratings across multiple registries.

Recent literature has emphasised the problem of camouflage when evaluating the performance of orthopedic implants, meaning that the performance of a specific implant-design variant is concealed when evaluation is done at the level of the implant name, which may contain multiple variants of an implant with slight iterations of the designs, which do not necessarily have the same performance.^{41,42} To prevent camouflage, ODEP reviews implants at the product-code level so that different variants under the same brand name have their own code (i.e. unique device identifier).⁴³ However, because registries do not publicly report revision risks at the level of the implant variant, the large majority (87%) of total knee implants in **Chapter 6** could not be matched to an ODEP-rating. To solve this matching-problem, and thus prevent camouflage, registries should collect revision risks at the product-code level instead of collecting it solely on implant brand name. Publishing such detailed information in registries' annual reports would likely be infeasible due to the immense number of pages involved. Online tools, like StatLine (an electronic database of Statistics Netherlands), can address this issue by allowing users to generate their own tables and graphs.⁴⁴ Another potential solution to prevent this matching-problem, is to request registry data for research purposes at the implant product-code level (similar to ODEP) instead of medical device name.

As stated earlier, the performance of an implant is also related to characteristics of patients and therefore the case-mix of the patient population in which the implant is used. Therefore, **Chapter 7** analysed the similarities and differences in patients receiving specific total knee implants across four national registries

to better understand possible differences in safety and performance for the same implant. A limited number (7%) of total knee implants were used in only one registry, thus data pooling across registries is feasible for the majority of total knee implants. Pooling data across registries results in earlier detection of any safety concerns, while also increasing the statistical power, which is particularly valuable for low frequency implants (e.g. tumor prostheses). However, heterogeneity between registries remains a concern, mainly due to differences in registry methodologies, included outcomes, patient case-mix, and definition of outcomes used. Relying solely on data obtained from individual registries could limit sample size, yet may provide more consistent data, allowing for clearer interpretation of results.

In **Chapter 7** the differences in patient characteristics (patient age, body mass index (BMI), gender, and diagnosis of osteoarthritis) between registries for patients receiving the same total knee implant is discussed. Considerable differences in patient characteristics were found, although only a small amount of these differences were deemed to be clinically relevant. Particularly differences in diagnosis (i.e. percentage of patients with osteoarthritis) were found to be both statistically and clinically relevant for most total knee implants, which indicate that countries differ in the diagnosis for which a specific total knee implant is most commonly used. The observed variation in patient characteristics across registries is in line with previous studies identifying variations in pre-operative characteristics of patients undergoing total knee implants across countries.⁴⁵⁻⁴⁷ However, these studies analysed all total knee implants combined, rather than analysing specific total knee implants indicating that **Chapter 7** adds a more detailed analysis. Such an implant level analysis provides more relevant information for clinicians, by understanding which specific total knee implant is more frequently used in specific patient groups.

Registry data are not only used to identify outlier performance of specific implants, but also used as quality tools assessing hospital performance regarding the outcomes for orthopedic patients (**Chapter 2**).⁴⁸ This feedback is mostly based on all-cause revision risks, and therefore does not indicate how to improve care as revision surgery depends on a multitude of factors.⁴⁸ Using data from the national Dutch arthroplasty register (LROI), in **Chapter 8** the between-hospital variation in revision risks after primary shoulder arthroplasty is discussed. Considerable between-hospital variation in revision risks following primary shoulder arthroplasty was found. But, statistical reliability of ranking hospitals performance was low, indicating that most of the observed variation in revision risks was due to chance rather than reflecting true differences in hospital performance. A previous study has shown better (i.e. moderate) statistical reliability of ranking hospitals after primary total hip and knee arthroplasty, likely explained by the higher (annual) procedural volumes for both primary total hip and knee arthroplasty compared to primary total shoulder arthroplasty.⁴⁸ Even though (overall) revision risks are an important clinical outcome and an international accepted outcome of treatment failure, (overall) revision risks are not a suitable measure for comparing hospital performance after primary shoulder arthroplasty, due to the relatively low number of events (**Chapter 8**). This indicates that other more frequently occurring outcomes should be used for hospital performance comparisons, such as patient-reported outcomes and hospital readmissions. However, these measures also have their disadvantages e.g. where patient-reported outcome measures can capture patient-perceived improvements in function and quality of life following arthroplasty surgery, their use for hospital performance measurement is hindered by relatively low response rates. A recent study showed that such a selection of patients has likely resulted in overestimated improvements of patient-reported outcome measures after total hip and knee arthroplasty.⁴⁹

Future perspectives

Although the value of highly complete registries is paramount, the studies of this thesis showed that registries can be further optimised, thus providing even more reliable data needed to assess performance and safety of implants, which will be of value to multiple stakeholders (e.g. manufacturers, regulators, patients, and manufacturers). Some of these challenges and shortcomings regarding the quality, transparency, and utility of registry data for assessing the safety and performance of high-risk medical devices used on the market are discussed in this thesis. In the following section, relevant future directions of research are discussed.

Combining medical device registry data using federated network analysis

The coordinated registry networks developed in the United States sound promising, and could also be implemented in the European Union due to the recently introduced European Health Dataspace Act. Such a federated network analysis conducts a pooled analysis of medical performances across multiple data sources. Given the significant amount of data collection of medical device registries, a federated network analysis will offer the potential to create real-world data regarding the safety and performance of specific medical devices based on both quantitative data (e.g. registries capturing data on patient reported outcome measures) as well as qualitative data (e.g. safety signals reporting medical device recalls). However, this federated network analysis is only achievable if all databases perform their analyses at the same level (e.g. medical-device level) to allow for data linkage and pooling and have harmonised their definitions and outcomes.

Overcoming challenges due to stricter clinical evaluation requirements set in the European Medical Device Regulation

Replacing the European Medical Device Directive (MDD), the European Union Medical Device Regulations (MDR 2017/745) came into effect on May 2021 and introduced stricter requirements for

providing clinical evidence of high-risk medical devices such as orthopedic implants.^{6,50,51} Medical device manufacturers are required to establish systematic methods to monitor their devices used on the market by consistently collecting, recording, and analysing data on their safety and performance.⁵¹ Under the Medical Device Directive (MDD), manufacturers could claim conformity for their new medical devices, but with the Medical Device Regulation (MDR) they must demonstrate performance and safety for each new variant, even if this new device is almost “equivalent” to the device already used on the market. This indicates that for each newly introduced medical device safety and performance data have to be collected, although the manufacturers do already have safety and performance data on the predecessor that can be used to estimate the likely performance of the new variant.

The stricter European Medical Device Regulation aims to improve patient safety, but it also causes challenges such as the higher administrative burden and increased financial costs for manufacturers.^{50,52} Consequently, manufacturers might withdraw certain medical devices from the European market to reduce this burden. This could be particularly relevant for less frequently used medical devices (e.g. tumour prostheses and elbow implants), for which the financial costs and time needed to collect sufficient clinical evidence on the safety and performance may outweigh the market benefits. As a result, this could discourage manufacturers from collecting sufficient clinical evidence as well as the development of new less frequently used medical devices, ultimately limiting the availability of these medical devices for specific patient populations.

Bridging studies used in pharmacological research (“a supplementary study conducted in the new region to provide pharmacokinetic, pharmacodynamic, and/or clinical data on efficacy, safety, dosage, and dose regimen to enable extrapolation of clinical trial data from the foreign region to the new region”) could be beneficial for manufacturers to provide evidence regarding the

safety and performance of new medical devices. Bridging studies are already used in the evaluation of medicines to get market approval for the same medicine in a different region, but are not used (yet) in the medical device field.^{53,54} Developing a methodology for bridging studies, where an estimation of the safety and performance of the new medical device is made based on the available safety and performance data of the old variant, might prevent the challenges manufacturers currently encounter in collecting sufficient clinical evidence for their new devices. If such a methodology is validated to be used for (high-risk) medical devices, the need to collect additional clinical data could be reserved for those devices where there remains uncertainty about the safety and performance, thereby reducing the burden for manufacturers and ensuring access to safe devices for patients. As stakeholders (e.g. regulators and assessors from notified bodies) may have different views on what constitutes acceptable uncertainty in clinical evidence, achieving consensus on accepted (un)certainty is crucial, and would enhance consistency and predictability in the decision-making by stakeholders regarding the acceptable benefit-risk ratio of new medical devices. One potential method to achieve such a consensus is a Delphi study, involving all relevant stakeholders in the regulation of medical devices. Transparency and collaboration between all stakeholders is key to improve safety and performance of implants for patients.

References

1. Klein P, Blommestein H, Al M, et al. Real-world evidence in health technology assessment of high-risk medical devices: Fit for purpose? *Health Econ.* 2022;31 Suppl 1(Suppl 1):10-24
2. Pijls BG, Meessen J, Tucker K, et al. MoM total hip replacements in Europe: a NORE report. *EFORT Open Rev.* 2019;4(6):423-9
3. Porte PJ, Smits M, Verweij LM, et al. The Incidence and Nature of Adverse Medical Device Events in Dutch Hospitals: A Retrospective Patient Record Review Study. *J Patient Saf.* 2021;17(8):e1719-e25
4. Gausden EB, Puri S, Chiu YF, et al. Mid-term survivorship of primary total knee arthroplasty with a specific implant. *Bone Joint J.* 2023;105-b(3):277-83
5. Krouwer JS. More Focus is Needed to Reduce Adverse Events for Diabetes Devices. *J Diabetes Sci Technol.* 2022;16(2):498-9
5. Krouwer JS. More Focus is Needed to Reduce Adverse Events for Diabetes Devices. *J Diabetes Sci Technol.* 2022;16(2):498-9
6. Fraser AG, Nelissen R, Kjærsgaard-Andersen P, et al. Improved clinical investigation and evaluation of high-risk medical devices: the rationale and objectives of CORE-MD (Coordinating Research and Evidence for Medical Devices). *Eur Heart J Qual Care Clin Outcomes.* 2022;8(3):249-58
7. Dawson LP, Biswas S, Lefkovits J, et al. Characteristics and Quality of National Cardiac Registries: A Systematic Review. *Circ Cardiovasc Qual Outcomes.* 2021;14(9):e007963
8. van Schie P, Hasan S, van Bodegom-Vos L, et al. International comparison of variation in performance between hospitals for THA and TKA: Is it even possible? A systematic review including 33 studies and 8 arthroplasty register reports. *EFORT Open Rev.* 2022;7(4):247-63
9. Lu H, Cole SR, Howe CJ, et al. Toward a Clearer Definition of Selection Bias When Estimating Causal Effects. *Epidemiology.* 2022;33(5):699-706
10. Food and Drug Administration (FDA). Use of Real-World Evidence to Support Regulatory Decision-Making for Medical Devices. Available: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/use-real-world-evidence-support-regulatory-decision-making-medical-devices> (last accessed 28 April, 2025)
11. National Institute for Health and Care Excellence (NICE). NICE real-world evidence framework. Available: <https://www.nice.org.uk/corporate/ecd9/chapter/overview> (last accessed 28 April, 2025)
12. Sedrakyan A, Aryal S. Maturity framework and select approaches for developing Coordinated Registry Networks (CRNs): Medical Device Epidemiology Network (MDEpiNet) supplement. *BMJ Surg Interv Health Technol.* 2022;4(Suppl 1):e000148

13. International Medical Device Regulators Forum (IMDRF). Tools for Assessing the Usability of Registries in Support of Regulatory Decision-Making. Available: <https://www.imdrf.org/sites/default/files/docs/imdrf/final/technical/imdrf-tech-180327-usability-tools-n46.pdf> (last accessed 28 April, 2025)
14. International Medical Device Regulators Forum (IMDRF). Methodological Principles in the Use of International Medical Device Registry Data. Available: <https://www.imdrf.org/sites/default/files/docs/imdrf/final/technical/imdrf-tech-170316-methodological-principles.pdf> (last accessed 28 April, 2025)
15. European Medicines Agency (EMA). Guideline on registry-based studies - Scientific guideline. Available: <https://www.ema.europa.eu/en/guideline-registry-based-studies-scientific-guideline#topics> (last accessed 28 April, 2025)
16. European Medicines Agency (EMA). Patient registries. Available: <https://www.ema.europa.eu/en/human-regulatory-overview/post-authorisation/patient-registries> (last accessed 28 April, 2025)
17. Sedrakyan A, Marinac-Dabic D, Campbell B, et al. Advancing the Real-World Evidence for Medical Devices through Coordinated Registry Networks. *BMJ Surg Interv Health Technol*. 2022;4(Suppl 1):e000123
18. Guilhaume C. A tool to assess the registries quality: The Registry Evaluation and Quality Standards Tool (REQueST). *European Journal of Public Health*. 2021;31(Supplement_3)
19. European Society of Cardiology (ESC). BigData@Heart. Available: <https://www.escardio.org/Research/Big-Data-Heart> (last accessed 28 April, 2025)
20. European Society of Cardiology (ESC). BigData@Heart. Available: <https://www.escardio.org/Research/Big-Data-Heart> (last accessed 28 April, 2025)
21. Network of Orthopaedic Registries of Europe (NORE). Minimal datasets. Available: <http://nore.efort.org/minimal-datasets> (last accessed 28 April, 2025)
22. European Network of Cancer Registries (ENCR). ENCR Recommendations 2022 Updated recommendations for a standard. Available: <https://policycommons.net/artifacts/3358432/encr-recommendations-2022-updated-recommendations-for-a-standard/4157085/> (last accessed 28 April, 2025)
23. Lübbeke A, Hoogervorst LA, Marang-van de Mheen PJ, et al. Arthroplasty registries at a glance: an initiative of the International Society of Arthroplasty Registries (ISAR) to facilitate access, understanding, and reporting of registry data from an international perspective. *Acta Orthop*. 2025;96:116-26
24. Josep Pane KMCV, Dorian Villegas, Laura Gamez, et al. Challenges Associated with the Safety Signal Detection Process for Medical Devices. *Med Devices (Auckl)*. 2021
25. European Commission. European Health Data Space Regulation (EHDS).

- Available: https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space-regulation-ehds_en (last accessed 28 April, 2025)
26. de Steiger RN, Miller LN, Davidson DC, et al. Joint registry approach for identification of outlier prostheses. *Acta Orthop.* 2013;84(4):348-52
 27. Department of Health and Aged Care. Prostheses List - Guide to listing and setting benefits for prostheses. Available: <https://www.health.gov.au/sites/default/files/documents/2020/06/prosthesis-list-guide.pdf> (last accessed 28 April, 2025)
 28. Poolman RW, Verhaar JA, Schreurs BW, et al. Finding the right hip implant for patient and surgeon: the Dutch strategy--empowering patients. *Hip Int.* 2015;25(2):131-7
 29. Chubb HA, Cornish ER, Hallstrom BR, et al. Early Benchmarking Total Hip Arthroplasty Implants Using Data from the Michigan Arthroplasty Registry Collaborative Quality Initiative (MARCQI). *Orthop Res Rev.* 2021;13:215-28
 30. Orthopaedic Data Evaluation Panel (ODEP). What is ODEP? Available: <https://www.odep.org.uk/> (last accessed 28 April, 2025)
 31. Keurentjes JC, Pijls BG, Van Tol FR, et al. Which implant should we use for primary total hip replacement? A systematic review and meta-analysis. *J Bone Joint Surg Am.* 2014;96 Suppl 1:79-97
 32. Deere KC, Whitehouse MR, Porter M, et al. Assessing the non-inferiority of prosthesis constructs used in hip replacement using data from the National Joint Registry of England, Wales, Northern Ireland and the Isle of Man: a benchmarking study. *BMJ Open.* 2019;9(4):e026685
 33. Wyatt M, Frampton C, Whitehouse M, et al. Benchmarking total hip replacement constructs using noninferiority analysis: the New Zealand joint registry study. *BMC Musculoskelet Disord.* 2021;22(1):719
 34. Nelissen RG, Brand R, Rozing PM. Survivorship analysis in total condylar knee arthroplasty. A statistical review. *J Bone Joint Surg Am.* 1992;74(3):383-9
 35. Schmerler J, Harris AB, Srikumaran U, et al. Body Mass Index and Revision Total Knee Arthroplasty: Does Cause for Revision Vary by Underweight or Obese Status? *J Arthroplasty.* 2023;38(12):2504-9.e1
 35. Schmerler J, Harris AB, Srikumaran U, et al. Body Mass Index and Revision Total Knee Arthroplasty: Does Cause for Revision Vary by Underweight or Obese Status? *J Arthroplasty.* 2023;38(12):2504-9.e1
 36. Bigham WR, Lensing GS, Walters MM, et al. Outcomes of Total Knee Arthroplasty Revisions in Obese and Morbidly Obese Patient Populations. *J Arthroplasty.* 2023;38(9):1822-6
 37. Spicer DD, Pomeroy DL, Badenhausem WE, et al. Body mass index as a predictor of outcome in total knee replacement. *Int Orthop.* 2001;25(4):246-9
 38. Bayliss LE, Culliford D, Monk AP, et al. The effect of patient age at

intervention on risk of implant revision after total replacement of the hip or knee: a population-based cohort study. *Lancet*. 2017;389(10077):1424-30

39. Perdisa F, Bordini B, Salerno M, et al. Total Knee Arthroplasty (TKA): When Do the Risks of TKA Overcome the Benefits? Double Risk of Failure in Patients up to 65 Years Old. *Cartilage*. 2023;14(3):305-11

40. Schmerler J, Bergstein VE, ElNemer W, et al. The weight of complications: high and low BMI have disparate modes of failure in total hip arthroplasty. *Arthroplasty*. 2024;6(1):9

41. Wilton T, Skinner JA, Haddad FS. Camouflage uncovered: what should happen next? *Bone Joint J*. 2023;105-b(3):221-6

42. Phillips JRA, Tucker K. Implant brand portfolios, the potential for camouflage of data, and the role of the Orthopaedic Data Evaluation Panel in total knee arthroplasty. *Bone Joint J*. 2021;103-b(10):1555-60

43. European Commission. Unique Device Identifier - UDI. Available: https://health.ec.europa.eu/medical-devices-topics-interest/unique-device-identifier-udi_en (last accessed 28 April, 2025)

44. Centraal Bureau voor de Statistiek (CBS). Information about StatLine. Available: <https://www.cbs.nl/en-gb/figures/statline/information-about-statline#:~:text=StatLine%20is%20the%20electronic%20database,for%20finding%20and%20showing%20data> (last accessed 28 April, 2025)

45. Lingard EA, Katz JN, Wright EA, et al. Predicting the outcome of total knee arthroplasty. *J Bone Joint Surg Am*. 2004;86(10):2179-86

46. Ackerman IN, Dieppe PA, March LM, et al. Variation in age and physical status prior to total knee and hip replacement surgery: a comparison of centers in Australia and Europe. *Arthritis Rheum*. 2009;61(2):166-73

47. Franklin PD, Miozzari H, Christofilopoulos P, et al. Important patient characteristics differ prior to total knee arthroplasty and total hip arthroplasty between Switzerland and the United States. *BMC Musculoskelet Disord*. 2017;18(1):14

48. van Schie P, van Steenbergen LN, van Bodegom-Vos L, et al. Between-Hospital Variation in Revision Rates After Total Hip and Knee Arthroplasty in the Netherlands: Directing Quality-Improvement Initiatives. *J Bone Joint Surg Am*. 2020;102(4):315-24

49. van Schie P, van Bodegom-Vos L, Zijdemans TM, et al. Linking Patient-Reported Outcome Measure Scores to Adverse Event Data to Gain Insight into Overestimation of Postoperative Patient-Reported Outcome Measure Improvement After Total Hip Arthroplasty and Total Knee Arthroplasty Due to Selective Nonresponse. *J Arthroplasty*. 2025;40(1):45-52

50. Kearney B, McDermott O. The Challenges for Manufacturers of the Increased Clinical Evaluation in the European Medical Device Regulations: A Quantitative Study. *Ther Innov Regul Sci*. 2023;57(4):783-96

51. Melvin T, Torre M. New medical device regulations: the regulator's view. *EFORT Open Rev.* 2019;4(6):351-6
52. Maresova P, Rezny L, Peter L, et al. Do Regulatory Changes Seriously Affect the Medical Devices Industry? Evidence From the Czech Republic. *Front Public Health.* 2021;9:666453
53. Zeng D, Pan Z, Lin DY. Design and analysis of bridging studies with prior probabilities on the null and alternative hypotheses. *Biometrics.* 2020;76(1):224-34
54. Wang T, Cao X, He Y, Chen X. Innovation drug approvals based on a bridging study: from concept to practice. *Transl Breast Cancer Res.* 2022;3:2