A Distributed Health Data Network Analysis of Survival Outcomes

The International Consortium of Orthopaedic Registries Perspective

Samprit Banerjee, PhD, Guy Cafri, PhD, MStat, Abby J. Isaacs, MS, Stephen Graves, MBBS, DPhil, FRACS, FAOrthA, Elizabeth Paxton, MA, Danica Marinac-Dabic, MD, PhD, and Art Sedrakyan, MD, PhD

Abstract: The International Consortium for Orthopaedic Registries is a multinational initiative established by the United States Food and Drug Administration to develop a health data network aimed at providing a robust infrastructure to facilitate evidence-based decision-making on performance of medical devices. Through the International Consortium for Orthopaedic Registries, individual data holders have complete control of their data and can choose to participate in studies of their choice. In this article, we present an overview of the data extraction process and the analytic strategy employed to answer several device performance-related questions in total hip arthroplasty and total knee arthroplasty. In the process, we discuss some nuances pertinent to International Consortium for Orthopaedic Registries data that pose certain statistical challenges, and we briefly suggest strategies to be adopted to address them.

In today's environment of growing concern for the safety and effectiveness of orthopaedic devices, the U.S. Food and Drug Administration (FDA) established the International Consortium for Orthopaedic Registries (ICOR)¹. The ICOR is a multinational initiative to develop a health data network^{2,3} aimed at providing a robust infrastructure to facilitate evidence-based decision-making on the performance of medical devices. National and regional registries are invaluable resources for medical device data and can address questions pertinent to all stakeholders, such as patients, clinicians, regulatory agencies, payers, and manufacturers⁴. Registries working together and combining their data substantially enhance their capabilities and impact. Although a centralized database is desirable, it is often not feasible or practical because of multiple security, proprietary, operational, legal, and patient-privacy barriers. In an effort to overcome these barriers, the ICOR provides a decentralized distributed health data network, in which individual data holders have complete control over their data and can choose to participate in studies of their choice. The ICOR focuses on certain device categories, harmonizes definitions, classifies exposures of risk, and obtains data summaries from individual registries.

In this series of articles, we focused on primary total hip arthroplasty and total knee arthroplasty and obtained data from

seven national and international participating registries (two from the United States and one each from Sweden, Italy, Spain, Norway, and Australia); see Table I for a complete list. The outcome of interest for both total hip arthroplasty and total knee arthroplasty was time to first revision surgery. Several key risk factors were considered for both total hip arthroplasty and total knee arthroplasty, as will be explained in the next section. Inclusion and exclusion criteria and classification of these risk factors were harmonized by the ICOR investigators. A standardized data extraction code was developed by the coordinating center and was distributed to participating registries to be run on their databases. Data summaries obtained from each registry were then combined to evaluate the risk of revision surgery in an advanced meta-analytic framework.

Our implementation of total hip arthroplasty and total knee arthroplasty risk evaluation under the ICOR framework can be considered as an advanced meta-analysis of survival outcomes. The most common approach in such scenarios is to combine (meta-analyze) hazard ratios across studies or registries to evaluate the comparative treatment effect (e.g., device A compared with device B) within each registry and to pool these estimates across registries with appropriate weights. As treatment effect is measured within each registry, this method has

Disclosure: One or more of the authors received payments or services, either directly or indirectly (i.e., via his or her institution), from a third party in support of an aspect of this work. None of the authors, or their institution(s), have had any financial relationship, in the thirty-six months prior to submission of this work, with any entity in the biomedical arena that could be perceived to influence or have the potential to influence what is written in this work. In addition, no author has had any other relationships, or has engaged in any other activities, that could be perceived to influence or have the potential to influence what is written in this work. The complete **Disclosures of Potential Conflicts of Interest** submitted by authors are always provided with the online version of the article.

Disclaimer: The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding source, the National Center for Advancing Translational Sciences, based in Rockville, Maryland.

Fig. 1

The ICOR distributed analysis process for total hip arthroplasty in which the ICOR Coordinating Center creates analysis programs and distributes them to six registries (A), registries run the programs on-site and submit resulting Nelson-Aalen survival estimates for all possible covariate combinations (B), and the center uses mixed effects models to combine estimates across groups (C). E-R = Emilia-Romagna.

the advantage of controlling for measured and unmeasured confounders within each registry without the need for explicit modeling of registry-related variables or factors but also has many limitations⁵. First, device A and device B comparisons within registries will be restricted to registries that have data on both devices, thereby impacting the precision of estimates negatively by shrinking the number of registries; second, only predefined interactions between device and risk factors (heterogeneity of treatment effect) can be studied, excluding exploration of interactions after data collection; and finally, the risk progression over time cannot be measured explicitly (e.g., risk of early versus late failure).

The ICOR presents a very flexible framework in which there is complete control on data extraction from registries. Therefore, our approach was to combine survival curves across registries by obtaining estimates of survival probabilities and their standard errors. This approach can be called the average arm effect, as it evaluates the effect of each arm and comparisons of arms are performed after arm effects are combined across registries. This approach has several advantages. First, risk progression over time can be studied, which provides a nuanced insight into the failure of devices (e.g., in a comparison between device A and device B, device A can have an increased risk of early failure compared with device B, but they both might have a comparable ten-year overall risk rate). Second, single-arm registries (having data only on one device) are not excluded in a comparison between device A and device B. Third, an investigation of interactions between risk factors can be performed in the analysis stage as long as data

are extracted for all subgroups. This is not possible in meta-analysis of hazard ratios because only predefined interactions (e.g., hazard ratios for interaction between treatment and risk factor) can be tested in the analysis stage. Fourth, time-dependent effects of risk factors can be examined in the analysis stage (e.g., including a risk factor by time interaction). Finally, such modeling can be presented in a linear mixed model framework analysis that can then be carried out using standard statistical software. However, this approach of combining survival curves does not implicitly control for registry-level confounders. This concern is addressed by modeling a fixed effect for registries in the advanced meta-analysis, which statistically controls for all registry-level characteristics influencing treatment arms within the same registry^{6,7}.

In this article, we present an overview of the data extraction process and analytic strategy employed to answer several device performance-related questions in total hip arthroplasty and total knee arthroplasty. In the process, we discuss some nuances pertinent to ICOR data that pose certain statistical challenges, and we briefly suggest strategies to be adopted to address them.

Data Extraction

Aggregate-level data were extracted from six registries each (a total of seven distinct registries) for total hip arthroplasty and total knee arthroplasty; sample sizes are given in Table I. The overall mechanism of extraction is given in Figure 1. The first step of the analytic method was to determine the inclusion

TABLE I Sample Size of Participating Registries

Registry	No. of Patients
Total hip arthroplasty	42,322 (100%)
HealthEast Joint Replacement (United States)	504 (1.19%)
Kaiser Permanente Total Joint Registry (United States)	4167 (9.85%)
Australian Orthopaedic Association National Joint Replacement Registry (Australia)	28,796 (68.04%)
Emilia-Romagna Joint Registry R.I.P.O.* (Italy)	3851 (9.1%)
Norwegian Arthroplasty Registry (Norway)	3556 (8.4%)
Catalan Arthroplasty Register (Spain)	1448 (3.42%)
Total knee arthroplasty	457,232 (100%)
Kaiser Permanente (United States)	57,599 (12.6%)
Swedish Knee Arthroplasty Register (Sweden)	84,915 (18.57%)
Australian Orthopaedic Association National Joint Replacement Registry (Australia)	247,495 (54.13%)
Emilia-Romagna Joint Registry R.I.P.O.* (Italy)	22,648 (4.95%)
Norwegian Arthroplasty Registry (Norway)	23,548 (5.15%)
Catalan Arthroplasty Register (Spain)	21,027 (4.6%)

*R.I.P.O. = Registro dell'implantologia Protesica Ortopedica.

and exclusion criteria and to harmonize demographic and device-related variables. The ICOR investigators coordinated and agreed upon the following inclusion and exclusion criteria. The inclusion criteria for total hip arthroplasty were patient age of sixty-five years and primary surgical cases performed from January 1, 2001, to December 31, 2010; the inclusion criteria for total knee arthroplasty were primary total knee replacement, an osteoarthritis diagnosis, and cases performed from January 1, 2001, to December 31, 2010. The exclusion criterion for total hip arthroplasty was articular surface replacement (ASR)⁸ implants; the exclusion criteria for total knee arthroplasty were unicompartmental, bicompartamental, constrained, and hinged devices. We also excluded the DePuy ASR Hip System (DePuy, Warsaw, Indiana) from the metal-on-metal bearing surface class, as it is an outlier implant with notoriously poor performance⁸ (leading to its subsequent recall) that would otherwise have artificially inflated the risk of revision estimates in this bearing class. SAS syntax, developed in the coordinating center, was distributed to each registry for data extraction. The event of interest was the occurrence of revision surgery, and the outcome was time to revision surgery. Hence, a survival analysis for time to revision surgery was performed within each registry, and Nelson-Aalen estimates⁹ of survival probabilities and their standard errors were obtained. The survival analysis was stratified on five attributes for total hip arthroplasty: bearing surface of the implant, size of the femoral head, method of fixation, sex, and age. Similarly, the survival analysis for total knee arthroplasty was also stratified on five attributes: mobility or stability, resurfacing, method of fixation, sex, and age. See Table II for a complete description of these attributes. The categorization of these attributes or covariates was predetermined and agreed upon by the ICOR investigators (Table II). Conceptually, we obtain one survival curve for every covariate profile within each registry (Fig. 1). An example of a covariate profile might be total hip arthroplasties of

men between the ages of forty-five years and fifty-four years with a metal-on-metal implant with a femoral head size of 28 mm and a cemented acetabular component and a femoral stem, referred to as an observation in the ensuing analysis. However, in practice, each row of the compiled data set is a survival probability estimate at each unique event time (time of revision surgery) for a particular covariate profile within a registry.

Materials and Methods

The goal of our analysis was to combine survival curves for each covariate profile within each registry in an advanced meta-analytic framework in which each covariate profile could be considered as a treatment arm and each registry could be considered as a study. The motivation for the data extraction procedure was that survival probabilities of observations (covariate profiles within registries) were dependent across the time points (occurrence of revision surgeries) and were independent between covariate profiles conditional on registries. The survival profile was then modeled as repeated measures data (over various time points) clustered within each registry in a mixed-effects model. To harmonize the time points, all unique survival times were combined or were binned into ten one-year intervals or bins based on the year in which the event (revision surgery) had taken place. Multiple events could occur in each bin or each year, and the earliest one was chosen as the representative of that time bin. This implies that estimates of survival and relative risk reflected effects early in the interval of the time specified. For instance, in the graphs of model-predicted survival provided, the first year interval (time zero to almost one year) reflected survival closer to time zero.

To work with the survival probabilities (values between 0 and 1), we subjected them to a complementary log-log ($\log(-\log(\cdot))$) transformation, a process that yields values on the entire real number line ($-\infty, \infty$). This transformation had several advantages. First, it allowed us to represent this model in a flexible linear mixed model framework. Second, the regression coefficients retained the interpretation of hazard ratio (in the log-scale). Finally, the predicted survival probabilities were restricted within (0, 1), which is not guaranteed when survival probabilities are used without any transformation. This advanced method of meta-analysis of survival curves was first proposed by Dear¹⁰ for experimental data and then was extended by Arends et al.¹¹ to random-effects models. The correlation between survival probabilities across various time points for a particular covariate profile within a registry was calculated from estimates of survival probability and its standard error¹¹ and was treated as a fixed parameter

TABLE II Risk Factors and Their Classification for Total Hip Arthroplasties and Total Knee Arthroplasties

Risk Factor	Classification
Total hip arthroplasties	
Bearing surface	Ceramic on conventional polyethylene, ceramic on highly cross-linked polyethylene, ceramic metal (e.g., Oxinium) on conventional polyethylene, ceramic metal (e.g. Oxinium) on highly cross-linked polyethylene, metal on conventional polyethylene, metal on cross-linked polyethylene, or metal on metal
Femoral head size	<28 mm, 28 mm, 32 mm, 36 mm, and >36 mm
Method of fixation	The four groups of uncemented, hybrid, cemented, and reverse hybrid were defined as: (1) uncemented, which included uncemented tibial tray uncemented acetabular component and uncemented femoral stem; (2) hybrid, which included uncemented acetabular component and cemented femoral stem; (3) cemented, which included cemented acetabular component and cemented femoral stem; and (4) reverse hybrid, which included cemented acetabular component and uncemented femoral stem
Sex	Male or female
Age	Less than forty-five years, forty-five to fifty-four years, fifty-five to sixty-four years, sixty-five to seventy-four years, and more than seventy-four years
Total knee arthroplasties	
Mobility or stability	Fixed and non-posterior stabilized, fixed and posterior stabilized, mobile and non-posterior stabilized, or mobile and posterior stabilized
Resurfacing	Patella resurfacing or no patella resurfacing
Method of fixation	The three groups of uncemented, cemented, and hybrid were defined as: (1) uncemented, which included tibial tray uncemented and femoral component uncemented; (2) hybrid, which included tibial tray cemented and femoral component uncemented; and (3) cemented, which included tibial tray cemented and femoral component cemented
Sex	Male or female
Age	Less than forty-five years, forty-five to fifty-four years, fifty-five to sixty-four years, or more than sixty-four years

in the linear mixed model. This mixed model is then iterated by updating the covariance estimates from predicted values until the estimates of coefficients stabilize^{10,11}. All covariates and a time variable were included as fixed effects, and a registry-level random intercept was also included. The inclusion of the time variable allowed the modeling of risk progression over time, and the random intercept modeled the between-registry heterogeneity through a variance parameter, thereby allowing us to extend the inference of the intercept to a general population of registries. However, the standard error of the registry-level random intercept variance estimate could be large due to a small number of registries. Thus, we also considered a model with fixed effects for registries and based our inference on selected registries only. The advantage of using registry as a fixed effect was that it controls for all registry-level characteristics (e.g., the difference in revision rates across registries due to heterogeneity in national and international revision policies). These two approaches are called the random-effects model and the fixed-effects model.

The estimated effects of covariates (regression coefficients) obtained from this model are interpreted by hazard ratios in the log-scale. A useful feature of this method is that interactions between covariates can be investigated in the analysis stage even though data were not extracted, keeping interactions in mind. Interactions with time in this model captured the differential nature of risk progression across different risk factors.

The approach¹¹ (henceforth, the Arends approach) was designed for meta-analysis of randomized controlled trials with two or very few treatment arms. The ICOR data had many treatment arms or covariate profiles, relatively few registries, a low revision rate, and a high degree of censoring. These features of the ICOR data presented considerable challenges in the implementation of the Arends approach. We provided solutions to these challenges and tailored the Arends approach for the ICOR data. We briefly describe them here for the sake of completeness.

Computational Issues

Because this method can be represented in a linear mixed model framework, its computation can be carried out with standard statistical software such as PROC MIXED of SAS 9.3 (SAS Institute, Cary, North Carolina). However, for the ICOR total hip arthroplasty data, there are as many as 1600 unique profiles across ten time points and six registries. This would result in a 96,000 × 96,000 dimensional covariance matrix, which becomes prohibitive for computation using SAS, even if the number of profiles is substantially reduced; although the covariance matrix is a fixed parameter in the linear mixed model, a bug in SAS 9.3 still tries to compute its inverse and runs out of memory, making the computation impossible. We alleviated this computational burden by considering a general linear model with a known covariance matrix by first transforming the response, fixed effects, and random effects using the square-root of the known covariance matrix.

Bias Correction

The imbalance in the distribution of sample sizes for treatments combined with small sample sizes for certain explanatory variable combinations within registry and a low event rate with a high degree of censoring could have biased the estimates of hazard ratios. Empirically, we observed that excluding observations with a high standard error (of the estimated survival probability) reduces this bias dramatically, improves coverage, and lowers the root mean squared error. For data analysis, we considered three thresholds for standard errors of ≥ 0.05 , ≥ 0.025 , and ≥ 0.0125 for the exclusion of observations, and we performed sensitivity analyses with and without observation exclusion for each threshold. We began with all observations and sequentially applied the data exclusion thresholds (highest to lowest). If estimates of hazard ratios changed substantially for a lower threshold (more observations excluded), we used the lower threshold as the exclusion criterion; otherwise, we retained the current threshold.

Confidence Intervals

Confidence interval construction for fixed-effects models is straightforward. Either the Wald method or a *t* distribution with degrees of freedom based on the number of observations minus the number of fixed effects provides reasonable coverage. However, for random-effect models with the kind of unbalanced data that we observed in the ICOR, our simulations suggested that fixed parameters with corresponding random effects have the best coverage when confidence intervals are constructed on the basis of a *t* distribution with degrees of freedom based on the number of registries minus one¹², as opposed to the Satterthwaite¹³, Kenward and Roger¹⁴, or other hybrid methods.

Discussion

We have described a decentralized, distributed health data network, the ICOR, and have presented a data extraction and data analysis strategy to synthesize evidence across seven national and international orthopaedic registries regarding the rates of revision surgery in primary total hip and total knee arthroplasties. As described, the ICOR harmonizes the definition and classification of risk factors, sends standardized SAS syntax to participating registries, obtains data summaries from each registry, and synthesizes the evidence across different registries.

Strengths

The decentralized distributed network of the ICOR overcomes many barriers (proprietary, security, operational, legal, and patient privacy) to creating a centralized database of multiple registries, yet has the advantage of strengthening estimation by borrowing information from multiple registries. The analytic method of combining survival curves is a flexible and robust approach to conduct inference. This method of analysis allows the study of risk progression over time, the exploration of interaction between risk factors, and the modeling of risk factors and time interactions to study changes in risk progression over time.

Limitations and Future Research

Although the participation of seven registries has enabled conduct of important demonstration projects, inclusion of data from additional registries would further enhance this approach. Although we found novel solutions to statistical challenges described in this article, more research on the analytic methods is warranted to understand the nuances of this methodology in the registry data framework. ■

Samprit Banerjee, PhD
Abby J. Isaacs, MS
Art Sedrakyan, MD, PhD
Weill Cornell Medical College,
402 East 67th Street,
New York, NY 10065

Guy Cafri, PhD, MStat
Elizabeth Paxton, MA
Surgical Outcomes & Analysis Department,
Kaiser Permanente,
8954 Rio San Diego Drive, Suite 406,
San Diego, CA 92108

Stephen Graves, MBBS, DPhil, FRACS, FAOrthA
Australian Orthopaedic Association National Joint
Replacement Registry,
Discipline of Public Health,
MDP DX 650 511,
University of Adelaide, Adelaide,
SA 5005, Australia

Danica Marinac-Dabic, MD, PhD
Center for Devices and Radiological Health,
Food and Drug Administration,
10903 New Hampshire Avenue,
Silver Spring, MD 20993-0002

References

1. Sedrakyan A, Paxton EW, Phillips C, Namba R, Funahashi T, Barber T, Sculco T, Padgett D, Wright T, Marinac-Dabic D. The International Consortium of Orthopaedic Registries: overview and summary. *J Bone Joint Surg Am*. 2011 Dec 21;93(Suppl 3):1-12.
2. Brown JS, Holmes JH, Shah K, Hall K, Lazarus R, Platt R. Distributed health data networks: a practical and preferred approach to multi-institutional evaluations of comparative effectiveness, safety, and quality of care. *Med Care*. 2010 Jun;48(6)(Suppl):S45-51.
3. Maro JC, Platt R, Holmes JH, Strom BL, Hennessy S, Lazarus R, Brown JS. Design of a national distributed health data network. *Ann Intern Med*. 2009 Sep 1;151(5):341-4. Epub 2009 Jul 28.
4. Sedrakyan A, Paxton EW, Marinac-Dabic D. Stages and tools for multinational collaboration: the perspective from the coordinating center of the International Consortium of Orthopaedic Registries (ICOR). *J Bone Joint Surg Am*. 2011 Dec 21;93(Suppl 3):76-80.
5. Combescore C, Courvoisier D, Haller G, Perneger T. Meta-analysis of two-arm studies: Modeling the intervention effect from survival probabilities. *Stat Methods Med Res*. 2012 Dec 24.
6. Localio AR, Berlin JA, Ten Have TR, Kimmel SE. Adjustments for center in multicenter studies: an overview. *Ann Intern Med*. 2001 Jul 17;135(2):112-23.
7. Allison PD. Fixed effects regression models (quantitative applications in the social sciences). Los Angeles: SAGE; 2009. p 123.
8. Curfman GD, Redberg RF. Medical devices—balancing regulation and innovation. *N Engl J Med*. 2011 Sep 15;365(11):975-7. Epub 2011 Aug 10.
9. Breslow N. Contribution to the discussion of the paper by DR Cox, regression models and life-tables. *J R Stat Soc B*. 1972;34(2):216-7.
10. Dear KB. Iterative generalized least squares for meta-analysis of survival data at multiple times. *Biometrics*. 1994 Dec;50(4):989-1002.
11. Arends LR, Hunink MG, Stijnen T. Meta-analysis of summary survival curve data. *Stat Med*. 2008 Sep 30;27(22):4381-96.
12. Higgins JP, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. *J R Stat Soc Ser A Stat Soc*. 2009 Jan;172(1):137-59.
13. Satterthwaite FE. An approximate distribution of estimates of variance components. *Biometrics*. 1946 Dec;2(6):110-4.
14. Kenward MG, Roger JH. Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics*. 1997 Sep;53(3):983-97.