ScienceDirect

Contents lists available at sciencedirect.com
Journal homepage: www.elsevier.com/locate/jval

Patient-Reported Outcomes

Using Patient-Reported Outcome Measures to Evaluate Care for Patients With Inflammatory Chronic Rheumatic Disease

Daniela Rodrigues, MSc, Andrew Street, PhD, Maria José Santos, MD, PhD, Ana Maria Rodrigues, MD, PhD, João Marques-Gomes, PhD, Helena Canhão, MD, PhD

ABSTRACT

Objectives: Few countries integrate patient-reported outcome measures (PROMs) in routine performance assessment and those that do focus on elective surgery. This study addresses the challenges of using PROMs to evaluate care in chronic conditions. We set out a modeling strategy to assess the extent to which changes over time in self-reported health status by patients with inflammatory chronic rheumatic disease are related to their biological drug therapy and rheumatology center primarily responsible for their care.

Methods: Using data from the Portuguese Register of Rheumatic Diseases, we assess health status using the Health Assessment Questionnaire-Disability Index for rheumatic patients receiving biological drugs between 2000 and 2017. We specify a fixed-effects model using the least squares dummy variables estimator.

Results: Patients receiving infliximab or rituximab report lower health status than those on etanercept (the most common therapy) and patients in 4 of the 26 rheumatology centers report higher health status than those at other centers.

Conclusions: PROMs can be used for those with chronic conditions to provide the patient's perspective about the impact on their health status of the choice of drug therapy and care provider. Care for chronic patients might be improved if healthcare organizations monitor PROMs and engage in performance assessment initiatives on a routine basis.

Keywords: comparative effectiveness of treatments, patient-reported outcome measures, performance of healthcare organizations, rheumatology.

VALUE HEALTH. 2022; 25(11):1885–1893

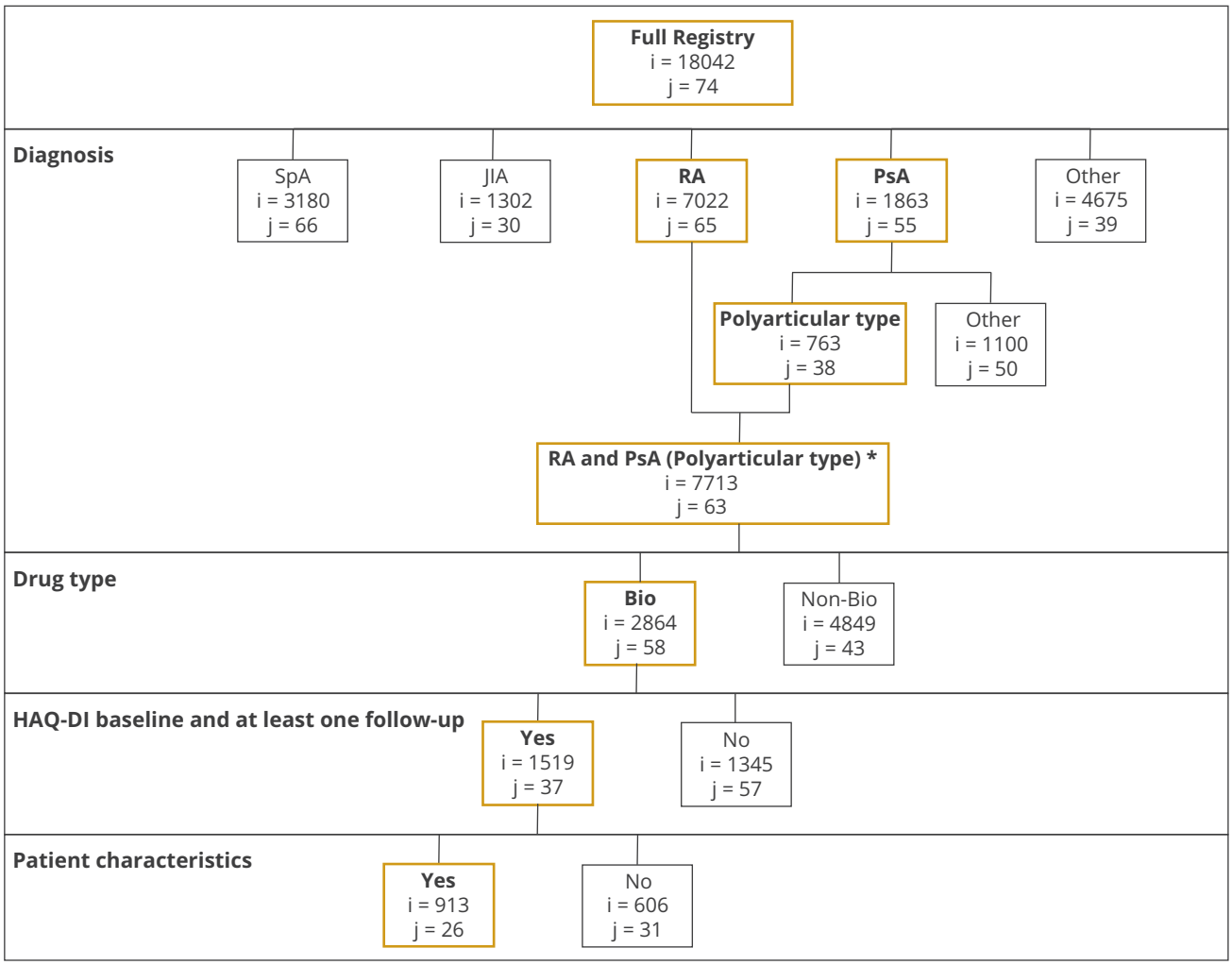
Introduction

Routine comparison of the performance of healthcare providers and treatments on the basis of how they improve patients' health is a prerequisite for developing a learning health system in which the latest evidence is used to guide patients and providers in their decision making. Nevertheless, estimating the effect of providers and treatments on patients' health is a challenge.[1,2] A key issue is how to measure "health" accurately. Traditional measures focused on clinical processes and patient outcomes such as survival. Although survival is the top priority for any patient, there are other aspects of health that are of great importance such as symptoms, functional status, and health-related quality of life (HRQoL). These health dimensions can be captured using patient-reported outcome measures (PROMs), questionnaires completed by patients.[3,4]

PROMs have been used not only to compare providers and treatments but also to support public reporting and inform value-based payment models.[5-8] The Dartmouth-Hitchcock Spine Center, Swedish Rheumatology Quality Register hosted by Karolinska University Hospital, and Group Health pioneered the use of PROMs in clinical practice and healthcare management.[9] Sweden has a long history of national registries, many of which collect, analyze, and report PROMs at the national level,[10,11] and has supported the development of several value-based payment models.[12] Another important initiative has been the English National Health Service's PROMs program in which patients who have a hip or knee replacement or hernia repair complete questionnaires about their HRQoL before and after the intervention.[11,13] Comparative data on change in HRQoL are published regularly and hospitals face financial penalties for poor response rates and outcomes.[14] In the United States, there has also been a shift toward value-based pricing,[15,16] which runs alongside the practice of public reporting of outcome data.[17] For example, PROMs have been incorporated into the Medicare Merit-Based Incentive Payment System and other payment models,[18] with an initial focus on joint replacement[19] and oncology.[20] Other countries such as Australia,[21-23] The Netherlands,[24] and Portugal[25] have started to promote the routine collection and use of PROMs at the national level. There is a growing interest worldwide in international

1098-3015/Copyright © 2022, International Society for Pharmacoeconomics and Outcomes Research, Inc. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Figure 1. Analytical sample.



* 72 patients participating in clinical trials were excluded at this stage

i indicates patients with inflammatory chronic rheumatic disease; j, rheumatology centers; Bio, biological drug; HAQ-DI, Health Assessment Questionnaire-Disability Index; JIA, juvenile idiopathic arthritis; Non-Bio, nonbiological drug; PsA, psoriatic arthritis; RA, rheumatoid arthritis; SpA, spondyloarthritis.

benchmarking using standard sets of PROMs for different health conditions.[26-28]

There is now substantial experience of collecting and reporting PROMs from those undergoing elective surgery. Attention is now turning to those with chronic conditions,[29,30] but applying PROMs to those with chronic diseases requires addressing specific challenges. First, elective surgery typically involves a single intervention, easily identifiable for the purposes of evaluation, and for which it is possible to establish baseline health status before the intervention. Instead, treatment for chronic conditions often involves a succession of interventions and pharmacotherapy, with baseline health status reset each time there is a change in therapy. This makes it challenging to assess the impact of each intervention on health status. Second, people with chronic conditions may well have a long treatment history, and this may affect their responses to any current intervention. Therefore, it is necessary to take account of treatment history in the evaluation. Third, for elective surgery, follow-up (FU) data are often collected at 1 or 2 time points, but for those with chronic conditions, there is no definitive time (other than death) at which treatment can be said to be completed. Evaluation necessitates use of long-term FU data.

We address these challenges in our analysis of care delivered to chronic disease patients with rheumatoid arthritis (RA) and polyarticular type of psoriatic arthritis (PsA) ("rheumatoid like") in Portugal. Patients receive treatment in specific rheumatology centers and are invited to complete the Health Assessment Questionnaire-Disability Index (HAQ-DI) before each biological drug therapy has been initiated and afterward during ongoing FU visits to the center. We use these questionnaires to assess whether patient's health status is related to their biological drug therapy and to the rheumatology center responsible for their care.

Methods

Data

Information was retrieved from the Portuguese Register of Rheumatic Diseases, collated by the Portuguese Society of Rheumatology.[25] A article-based registry was set up in 2000 with the goal of capturing demographic and clinical information about all patients in Portugal with RA receiving biological drug therapies. In 2003, patients with spondyloarthritis, PsA, and juvenile idiopathic

arthritis were also included. An electronic version of the registry, Reuma.pt, became active in 2008 as a desktop application, expanded to include patients receiving either biological or nonbiological drugs. In 2012, it was replaced by an online version (www.reuma.pt). Reuma.pt has been approved by the National Data Protection Authority and Ethics Committees from rheumatology centers and patients consented to be registered. As of May 2018, Reuma.pt included 74 rheumatology centers in public and private hospitals and clinics and 18 042 rheumatic patients with RA (39%), spondyloarthritis (18%), PsA (10%), juvenile idiopathic arthritis (7%), or any of the other less frequent conditions (26%).

Figure 1 shows which patient records were extracted from the registry to comprise the analytical sample for this study.

Our analytical sample was restricted to patients with RA or the polyarticular type of PsA taking biological drug therapies because these are 2 very similar groups of patients. In contrast to more traditional drugs such as conventional synthetic disease-modifying antirheumatic drugs (csDMARDs), biological drugs are genetically engineered agents that target specific cells, cellular interactions, and cytokines related to the inflammation process.[31] Biological drugs are a relatively new type of therapy for patients with inflammatory chronic rheumatic diseases, often prescribed when csDMARDs do not show a significant effect on patient outcomes. The medical decision to prescribe a biological drug is based not only on clinical guidelines[32,33] but also on clinical experience and professional judgment. This means that rheumatology centers may differ in their prescribing patterns and, hence, in the outcomes that their patients experience.

Each biological drug has a recommended dosage and administration route. If intravenous injection, patients receive it at the center; if subcutaneous injection, patients collect it at the center and then administer it themselves at home. Patients undertaking biological therapies were generally monitored every 3 to 6 months. As part of the monitoring process, patients were asked to complete the HAQ-DI, among other outcome measures. HAQ-DI was chosen because it captures patients' functional status, which is an important health dimension for these patients, and it has been shown to discriminate well across levels of RA severity.[34] Reuma.pt uses the Portuguese version of the short HAQ,[35] which was created and validated based on original proposals.[36,37] The short HAQ comprises a disability index, HAQ-DI, that can take any values from 0 (no difficulty) to 3 (very severe disability). To simplify the interpretation of the results, we inverted the HAQ-DI scale so that higher values indicate better outcomes. Patients were also asked to complete the Disease Activity Score with 28-joint counts (erythrocyte sedimentation rate) (DAS-28 [ESR]), which we also inverted for purposes of analysis. See Appendix Table A1 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2022.05.012 for more information about the HAQ-DI and DAS-28 (ESR).

Only patients with HAQ-DI data at both baseline (before each therapy is initiated) and at least one FU time were included in the study. In addition, we included only those patients for whom there was full information at baseline about the history of csDMARDs, the DAS-28 (ESR), age, duration of the disease, sex, smoking status, educational attainment, and job status. The analytical sample comprises 913 patients with RA or polyarticular type of PsA receiving from up to 11 different biological drugs in 26 different centers between 2000 and 2017. To assess the characteristics of those with missing HAQ-DI data, we compare the patients in our main sample who completed the HAQ-DI at both the baseline and at least one FU period with each of the following 3 samples: (1) those with only baseline HAQ-DI data, (2) those with only FU HAQ-DI data, and (3) those with no baseline or FU HAQ-DI data. Only patients with all characteristics were considered in each sample. Pearson's chi-squared test was performed for categorical variables; Student's t test was used for continuous variables. The results show that the main and 3 subsamples are very similar (see Appendix Tables A2-A4 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2022.05.012), thereby supporting the assumption that missing HAQ-DI data are completely at random.

Main Analysis

We specify a fixed-effects model using the least squares dummy variables estimator to assess whether patient-reported health status, as measured by HAQ-DI, varies according to the biological drug they are taking and the rheumatology center responsible for their care. Recognizing that patients are observed on multiple occasions, standard errors are clustered at patient level. The model takes the following form:

$$y_{ijt} = \sum_{k=1}^{K} \gamma_k V_{kijt} + \sum_{m=1}^{M} \delta_m W_{mijt} + \beta_0 y^o_{ijt} + \beta_1 z_{ijt} + \sum_{n=1}^{N} \lambda_n X_{nijt} + \tau_t + u_j + \varepsilon_{ijt} \quad (1)$$

where y_{ijt} indicates the patient's HAQ-DI response at the relevant FU time t (t = 3, 6, 12, 18 months), with i indexing patients and j indexing centers.

V_{kijt} is a set of k categorical variables that indicate each biological drug (k = etanercept, adalimumab, tocilizumab, infliximab, rituximab, golimumab, and a residual category, labeled "other," that comprises all other biological drugs, namely, certolizumab, abatacept, ustekinumab, secukinumab, and anakinra), with etanercept, the most common biological drug in the sample, acting as the reference. These categorical variables capture the relationship between reported health status and each biological drug. A positive coefficient $\widehat{\gamma}_k$ for biological drug k indicates that patients on that drug report better HAQ-DI scores than those on etanercept, $\widehat{\gamma}_E$. We also calculate the percent change in magnitude for each drug relative to the reference category. We do this by first estimating the marginal effect (me) of each drug on HAQ-DI, which is the mean patient's predicted outcome for each drug adjusted to the case mix of the entire estimation sample. This was determined by using *margins* command in Stata/BE 17.[38] The percent change is calculated as: $\frac{me_k - me_E}{me_E} \times 100$, for each biological drug k.

If patients have a long and complex history of either csDMARDs or biological drug therapies, we expect them to report worse health status. We control for this possibility by the set of m variables denoted W_{mijt}. These variables include a count of previous biological drugs and a vector of dummy variables for each sequence of previous biological drugs that contains 15 or more observations (for more information, see Appendix Table A5 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2022.05.012), a count of previous csDMARDs, and a vector of dummy variables for each sequence of previous csDMARDs that contains 15 or more observations (for more information, see Appendix Table A6 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2022.05.012).

It is likely that those reporting better health and disease status when starting the drug will report better health status at each FU time. Hence, we account for the patient's HAQ-DI, y^o_{ijt}, and DAS-28 (ESR), z_{ijt}, before starting the biological drug therapy.

Other patient characteristics, namely, their diagnosis, age, duration of the disease, sex, smoking status, educational attainment, and job status, are captured by the set of n variables comprising X_{nijt}.

We have also included a set of dummy variables, τ_t, indicating the FU time at which the HAQ-DI was completed with t = 6, 12, and

Table 1. Descriptive statistics for the main sample.

Sample characteristics	Mean (SD)/n
Total number of centers	26
Total number of patients	913
Number of patients per center	35 (60)
Total number of baseline HAQ-DI measures	1193
Total number of observations	3033

Variable name	Mean (SD)/n (%)
HAQ-DI (inverted scale) at follow-up time	1.898 (0.659)
Biological drug therapy	
Etanercept	374 (31)
Adalimumab	220 (18)
Tocilizumab	197 (17)
Infliximab	123 (10)
Rituximab	119 (10)
Golimumab	113 (9)
Other	47 (4)
Number of previous biological drug therapy*	1.636 (0.968)
Top 8 sequences of previous biological drug therapies*	
None	709 (59)
Etanercept	109 (9)
Infliximab	93 (8)
Adalimumab	53 (4)
Golimumab	23 (2)
Etanercept, adalimumab	17 (1)
Infliximab, etanercept	17 (1)
Tocilizumab	15 (1)
Number of previous csDMARDs*	2.331 (1.381)
Top 13 sequences of previous csDMARDs*	
Methotrexate	345 (29)
Sulfasalazine, methotrexate	76 (6)
Methotrexate, sulfasalazine	67 (6)
Hydroxychloroquine, methotrexate	64 (5)
Methotrexate, leflunomide	61 (5)
None	41 (3)
Hydroxychloroquine, methotrexate, sulfasalazine	32 (3)
Methotrexate, leflunomide, methotrexate	31 (3)
Methotrexate, sulfasalazine, methotrexate	28 (2)
Sulfasalazine, hydroxychloroquine, methotrexate	28 (2)
Methotrexate, hydroxychloroquine, methotrexate	19 (2)
Methotrexate, sulfasalazine, hydroxychloroquine	18 (2)
Methotrexate, hydroxychloroquine	17 (1)
Baseline HAQ-DI (inverted scale)*	1.522 (0.626)
Baseline DAS-28 (ESR) (inverted scale)*	3.384 (1.254)
Diagnosis	
Rheumatoid arthritis	793 (87)
Polyarticular type of PsA	120 (13)
Age (years)*	53 (12)
Duration of the disease (years)*	12 (9)
Sex	
Female	752 (82)
Male	161 (18)
Smoking status	
Never smoker	663 (73)
Current smoker	127 (14)
Ex-smoker	123 (13)
Educational attainment	
Low (< secondary)	608 (67)
Medium (secondary/postsecondary)	126 (14)
High (bachelor/master/doctorate)	131 (14)
No education	48 (5)
Job status	
Working (full/part-time)	408 (45)
Retired	383 (42)
Unemployed	59 (6)
Sick leave	63 (7)
Follow-up time (months)	
3	792 (26)
6	838 (28)
12	778 (26)
18	625 (21)

Continued in the next column

Table 1. Continued

csDMARDs indicates conventional synthetic disease-modifying antirheumatic drugs; DAS-28 (ESR), Disease Activity Score with 28-joint counts (erythrocyte sedimentation rate); HAQ-DI, Health Assessment Questionnaire-Disability Index; PsA, psoriatic arthritis.
*At the beginning of every biological drug therapy.

18 months and 3 months acting as the reference category. The categorization of the FU time (3FU, 6FU, 12FU, and 18FU months) is calculated as the difference in months between the dates of the baseline and FU appointments for each therapy. Naturally, there is variation in how much time elapsed between baseline and FU visits, as shown in Appendix Figure A1 in the Supplemental Materials found at https://doi.org/10.1016/j.jval.2022.05.012. Patients are allocated to a FU time if their visit falls within the following windows: $2mths < 3\text{FU} < 5mths$; $5mths < 6\text{FU} < 9mths$; $9mths < 12\text{FU} < 15mths$; and $15mths < 18\text{FU} < 21mths$, where *mths* indicates months after the baseline visit.

The rheumatology center-specific effect is captured by u_j, which we interpret as a measure of performance. A positive coefficient for center c, (u_c), indicates that patients from that center report better health status than the reference center (u_r), which was arbitrarily defined as center 1. To assess the difference in magnitude among all centers, we estimate the marginal effect of each center on HAQ-DI over and above the influence of all variables included in the model. We use funnel plots to compare center performance.[39,40] The adjusted HAQ-DI is plotted against the number of patients for each center, with 95% (2 SDs) and 99.7% (3 SDs) control limits indicating those centers in which patients report significantly higher or lower HAQ-DI scores than the average. Centers located above (below) the 95% and 99.7% control limits are deemed to have a positive (negative) "alert" and "alarm" status, respectively, their patients reporting much better (worse) health status than expected.

Finally, ε_{ij} is a random error term with the usual classical properties, $\varepsilon_{ij} \sim \text{N}(0, \sigma_\varepsilon^2)$. In terms of model performance, we report the adjusted R^2 as a measure of goodness of fit and we examine the histogram and Q-Q plot of the residuals from the regression model to assess whether these were normally distributed.

Sensitivity Analysis

We conducted 3 sensitivity analyses to assess the robustness of the treatment and center effects to the choice of analytical sample. First, it is possible that patients with RA and those with polyarticular type of PsA should not be pooled together in the same analysis if there are distinct unobserved characteristics. Our first

sensitivity analysis included only patients with a diagnosis of RA. Second, some centers report data for few patients, so our second sensitivity analysis excluded centers with HAQ-DI responses from fewer than 15 patients. Third, patients do not always visit their rheumatology center at each FU time and those that do might not always complete the HAQ-DI. To examine this, we performed a complete case analysis including only those patients with HAQ-DI at every FU time.

To test robustness of results to these sensitivity analyses, we examined the coefficients of the biological drug therapies and calculated the Pearson's correlation coefficient of the estimated center effects across the 3 different subsamples: (1) only patients with RA, (2) only patients within centers with 15 or more patients, and (3) only patients with HAQ-DI at every FU time: 3, 6, 12, and 18 months. We also report descriptive statistics for these subsamples to compare them with the main sample. For more information, see Appendix Tables A7 to A9 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2022.05.012.

Results

Descriptive Statistics

Descriptive statistics of the main sample are presented in Table 1.

The main sample encompasses 26 centers, 913 patients, and 3033 FU observations. Although 482 patients were on just a single biological drug therapy throughout the study period, 431 switched drugs, with some doing so up to 7 times. A baseline HAQ-DI was completed each time a new biological therapy was initiated, yielding a total of 1193 baseline HAQ-DI measures, with corresponding FU measures, although these were not always collected for each of the 4 FU times.

Generally, patients reported improved functional status after therapy, with the (inverted) average HAQ-DI score at each FU time being higher than the baseline score.

The 6 biological drugs most commonly received by patients were etanercept (31%), adalimumab (18%), tocilizumab (17%), infliximab (10%), rituximab (10%), and golimumab (9%). On average, patients had 1.6 biological therapies. Before starting any biological drug therapy, the average patient had received 2.3 csDMARDs, with methotrexate being the most common.

The sample mostly comprised RA patients (87%), with 13% of patients having polyarticular type of PsA. At baseline, the average patient's age was 53 years and they had the disease for 12 years. Most patients were female (82%), were never smokers (73%), had less than secondary education (67%), and were either working (full or part-time) (45%) or retired (42%).

Regression Analysis

The coefficients and 95% confidence intervals (CIs) for the biological drug therapy and rheumatology center in the main sample are shown in Figure 2. The parameter estimates and robust standard errors for all variables in all samples across the main and sensitivity analyses are reported in Appendix Table A10 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2022.05.012, along with the adjusted R^2 statistics all of which are between 0.45 and 0.49. Both the histogram and Q-Q plot of the residuals from the regression models resemble a normal distribution (see Appendix Fig. A2 in the Supplemental Materials found at https://doi.org/10.1016/j.jval.2022.05.012). For the main sample (column 1 in Appendix Table A10 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2022.05.012), these multivariate estimates are also compared with those derived from univariate analysis of each variable in Appendix Table A11 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2022.05.012. The changes to the magnitude and significance of the univariate estimates demonstrate the importance of controlling for other factors when assessing the relationship between 2 variables.

Comparative effectiveness of biological drugs

Figure 2 shows how the (inverted) HAQ-DI scores reported at each FU time relate to the biological drug therapy, relative to etanercept, the reference category. Compared with those on etanercept, patients reported lower HAQ-DI scores if on infliximab (10% lower, $\widehat{\gamma}_I = -0.185$, $P < .01$) or rituximab (7% lower, $\widehat{\gamma}_R = -0.142$, $P < .01$). These significant negative effects remained if analyzing only those with RA (column 2) or in centers with ≥15 patients (column 3) but were not significant if restricting the sample to those with complete FU measures (column 4).

As Appendix Table A10 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2022.05.012 confirms, at $P < .001$ level, there were only 3 other significant influences on HAQ-DI. Higher FU HAQ-DI scores were reported by those with higher HAQ-DI baseline scores, for younger patients, for men, and at the 18-month FU. These effects were also evident for each subsample. In general, none of the variables capturing previous use of csDMARDs or biological drugs individually proved significant influences on HAQ-DI at $P < .001$, although some were at lower levels of significance. Nevertheless, taken together, the variables measuring previous use of biological drugs were jointly significant (F test $P < .001$), although this was not the case for the csDMARD variables (F test $P < .055$).

Relative performance of rheumatology centers

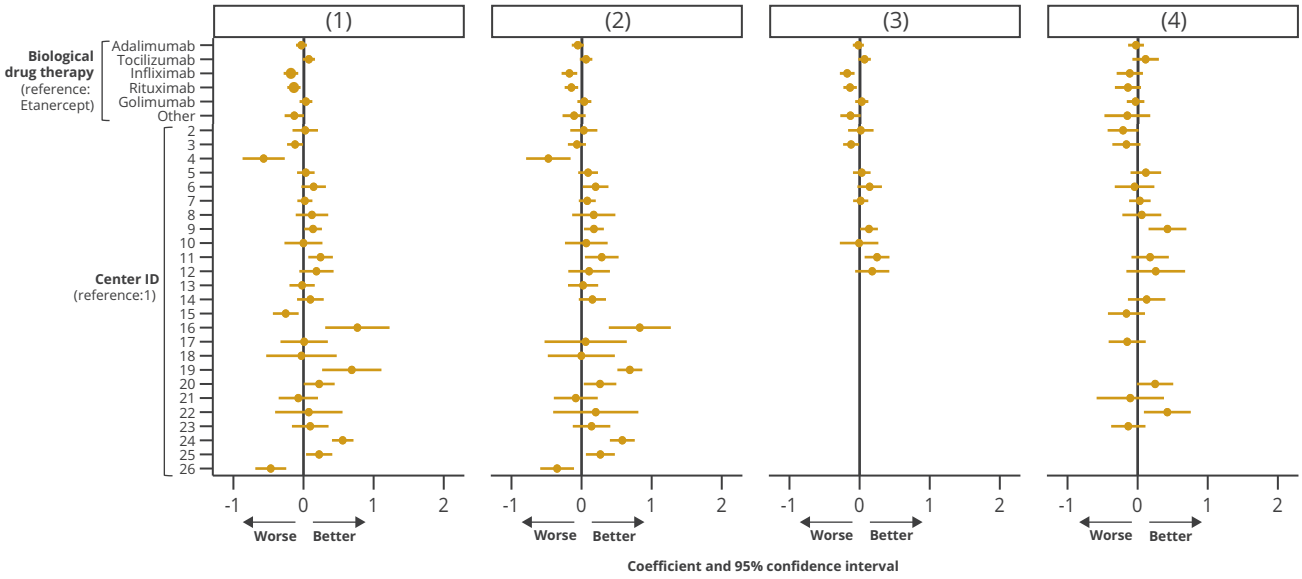
The fixed-effects estimates of the performance of each center with respect to the reference category (center 1) are shown in Figure 2, with the funnel plot presented in Figure 3.

A total of 4 centers (9, 11, 16, 19) lie above the 99.7% upper control limit, indicating that their patients report significantly higher HAQ-DI scores than those at other centers. The opposite is evident for 2 centers (3, 4) that lie below the 99.7% lower control limit. The Pearson's correlation coefficients indicate that the relative performance of centers is robust to focusing solely on RA patients $r_{12} = 0.990$ (95% CI 0.978-0.996) and to restricting analysis to only those centers with ≥15 patients $r_{13} = 1.000$ (95% CI 0.998-1.000), but is less robust to restricting the sample to patients with complete FU measures $r_{14} = 0.617$ (95% CI 0.193-0.846).

Discussion

This study exploited rich longitudinal registry data to assess whether patients' self-reported health status, measured using HAQ-DI for those with inflammatory chronic rheumatic disease, was related to their biological drug therapy and to the rheumatology center responsible for their care. The evaluation necessitated addressing 3 key challenges associated with assessing changes in health status of those with chronic conditions. First, patients regularly switch therapies, so determining the effect of each therapy requires measuring health status every time a new therapy is about to commence. This was possible because the registry data contain baseline measures each time a patient changes therapy. Second, treatment history may have an effect on how patients respond to their current therapy. The registry data included the previous use of csDMARDs and biological drugs, allowing us to take this complex history into account by

Figure 2. Forests plots with coefficients and 95% confidence intervals for the biological drug therapy and rheumatology center in (1) the main sample and each sample in the sensitivity analysis, (2) only patients with rheumatoid arthritis, (3) only patients within centers of 15 or more patients, and (4) only patients with HAQ-DI at every FU time, 3, 6, 12, and 18 months. Please refer to Appendix Table A10 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2022.05.012 to assess parameter estimates and robust standard errors for all variables in all samples across the main and sensitivity analyses.



HAQ-DI indicates Health Assessment Questionnaire-Disability Index; ID, identifier.

constructing indicator variables for sequences with at least 15 observations. An alternative could have been to disregard the ordering of csDMARDs and biological drugs when considering treatment history. Third, recognizing that treatment is an ongoing process, patients were asked to report their health status during FU visits held 3, 6, 12, and 18 months after commencing each biological therapy.

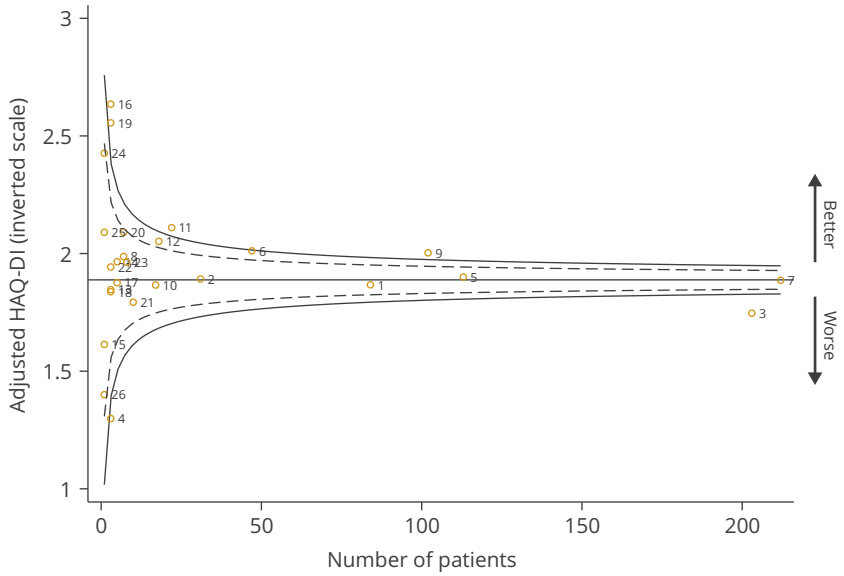
Some of our results contrast with findings from previous studies that have analyzed these registry data, but which focused on the DAS-28 as the primary measure of outcome.[41,42] Canhao et al[41] found no significant difference in DAS-28 among etanercept, adalimumab, and infliximab, whereas we found that patients on infliximab and rituximab report lower HAQ-DI scores than those on etanercept. Romao et al[42] reported tocilizumab to be associated with better DAS-28 scores than etanercept, whereas we found no significant difference when analyzing the HAQ-DI. The contrasting insights when using the HAQ-DI suggest that it should be considered alongside other outcome measures, in particular if we are to value the opinion of patients about the care they receive.[6,43]

The analyses identify 4 rheumatology centers in which patients report significantly higher HAQ-DI scores than elsewhere, this not being due the characteristics of the patients accounted for in the regression models. These "positive outliers" may repay more detailed investigation via qualitative research to identify why their patients report better health status and, if this suggests good practice, to share lessons across the sector.[44,45] For example, it may be that these centers are more effective than others at prescribing the right biological drug therapy to their patients, organizing care, and managing the whole treatment process, including acting within a reasonable time by switching the patient to another biological drug therapy if showing a poor response to current therapy.

Limitations

This study has limitations. First, rather than bespoke primary data collection, we relied on secondary data for the analysis. Secondary data are available at much lower cost, but may be less accurate, although the comprehensiveness, detail, and accuracy of registry data tend to exceed that of administrative data, and registry data have been recommended for use in evaluating the organization and delivery of healthcare.[46–48] Second, in the main analysis, we assumed missing HAQ-DI data across FU times to be completely at random. If that is not the case, results may be subject to selection bias in the form of nonrandom participation. To address this issue, we conducted an analysis only of those patients with complete FU records. Of course, perhaps these patients have complete records because they are different, requiring closer monitoring of their condition. Reassuringly, most results were not sensitive to focusing on this subgroup, but the negative effects of infliximab and rituximab, relative to etanercept, became insignificant. Third, the number of patients from the rheumatology centers ranged from some with >100 patients to some with fewer than 15 patients. Nevertheless, results were robust to excluding those centers contributing fewer patients to the study. Fourth, patients are not randomly allocated to either therapies or centers, compromising our ability to isolate the effects of therapies and centers on outcomes. To address this, we accounted for a large number of patient characteristics in the analysis, notably treatment history, but this comes at the risk of model overfitting. This concern informed our decision to include dummy variables for each sequence of previous biological drugs or of previous csDMARDs only if there were 15 or more observations. Finally, although the registry provides extremely rich data, the study may yet endure omitted variable bias. For instance, the patient's post-treatment health status may also be influenced by the care

Figure 3. Funnel plot of the adjusted score of the HAQ-DI for the main sample. Hollow circles and associated numbers represent the center IDs; dashed lines, 95% control limits; solid lines, 99.7% control limits.



HAQ-DI indicates Health Assessment Questionnaire-Disability Index; ID, identifier.

provided by primary and social care services, over and above that provided by the rheumatology center. Such information is not captured in the registry but might be identified through further qualitative analysis of the positive outliers that our study has identified.

Policy Implications

It is important to consider the patient's perspective in the assessments of the effectiveness of therapies and performance of healthcare organizations. These assessments are more challenging when assessing chronic rather than acute conditions, but this study benefited from using rich registry data that allow these challenges to be addressed. This type of exercise would give more insights if executed on a routine basis, perhaps as part of an audit of healthcare organizations. For example, the analysis could be run every year and findings published in an annual report, perhaps publicly available, to promote accountability and motivate quality improvement.[49] This is more likely to be successful if undertaken as a collaborative exercise among analysts, healthcare professionals, and policy makers and by making use of existing data.[50] Research using registry data to evaluate organizational performance remains in its infancy,[46] but our study provides evidence that national clinical databases can be used for this evaluative purpose. Policy makers might encourage participation in the evaluative exercise by rewarding organizations that monitor PROMs on a routine basis,[14] and in the first instance, promoting accurate and comprehensive data collection might be more effective at generating improvement than rewarding or penalizing those providers performing better or worse than average.[51,52]

Conclusions

We assessed the associations among the HAQ-DI, biologic drugs, and the rheumatology center in which care was delivered to patients with inflammatory chronic rheumatic disease. PROMs can be used for those with chronic conditions to provide the patient's perspective about the impact on their health status of the choice of drug therapy and care provider. Care for chronic patients might be improved if healthcare organizations monitor PROMs and engage in performance assessment initiatives on a routine basis.

Supplemental Materials

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.jval.2022.05.012.

Article and Author Information

Accepted for Publication: May 12, 2022

Published Online: June 23, 2022

doi: https://doi.org/10.1016/j.jval.2022.05.012

Author Affiliations: NIHR Imperial Patient Safety Translational Research Center, Institute of Global Health Innovation, Department of Surgery & Cancer, Imperial College London, London, England, UK (D. Rodrigues); Department of Health Policy, London School of Economics and Political Science, London, England, UK (Street); Rheumatology Department, Hospital Garcia de Orta, Almada, Portugal (Santos); JE Fonseca Lab, Instituto de Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa, Lisboa, Portugal (Santos); EpiDoC Unit, NOVA Medical School, Universidade NOVA de Lisboa, Lisboa, Portugal (A.M. Rodrigues, Canhão); CHRC, Comprehensive Health Research Center, NOVA Medical School, Universidade NOVA de Lisboa, Lisboa, Portugal (A.M. Rodrigues, Canhão); Rheumatology Unit, University Central Hospital Lisbon (CHULC), Lisboa, Portugal (A.M. Rodrigues); ReumaPt, Sociedade Portuguesa de Reumatologia, Lisboa, Portugal (Marques-Gomes); Nova School of Business and Economics, Carcavelos, Portugal (Marques-Gomes); NOVA Medical School, Universidade NOVA de Lisboa, Lisboa, Portugal (Marques-Gomes).

Correspondence: Daniela Rodrigues, MSc, NIHR Imperial Patient Safety Translational Research Centre, Institute of Global Health Innovation, Department of Surgery & Cancer, Imperial College London, 10th Floor, Queen Elizabeth the Queen Mother Wing (QEQM), St Mary's Campus,

London W2 1NY, England, United Kingdom. Email: d.rodrigues@imperial.ac.uk

Author Contributions: *Concept and design:* D. Rodrigues, Street, Santos, A.M. Rodrigues, Marques-Gomes, Canhão
Acquisition of data: Santos, A.M. Rodrigues, Canhão
Analysis and interpretation of data: D. Rodrigues, Street
Drafting of the manuscript: D. Rodrigues, Street
Critical revision of the paper for important intellectual content: D. Rodrigues, Street, Santos, A.M. Rodrigues, Marques-Gomes, Canhão
Statistical analysis: D. Rodrigues, Street
Administrative, technical, or logistic support: Marques-Gomes, Canhão
Supervision: Street, Marques-Gomes, Canhão

Conflict of Interest Disclosures: The authors reported no conflicts of interest.

Funding/Support: The authors received no financial support for this research.

Acknowledgment: The authors thank the Portuguese Society of Rheumatology for providing access to the data used in the current study and the journal's editor and reviewers for their insightful and constructive comments on previous versions of the article.

Ethical Approval: This study has been unanimously approved by the Ethics Research Committee NMS|FCM-UNL.

REFERENCES

1. Raine R, Fitzpatrick R, Barratt H, et al. *Challenges, Solutions and Future Directions in the Evaluation of Service Innovations in Health Care and Public Health*. Southampton, United Kingdom: NIHR Journals Library; 2016.
2. Sutton M, Garfield-Birkbeck S, Martin G, et al. *Economic Analysis of Service and Delivery Interventions in Health Care*. Southampton, United Kingdom: NIHR Journals Library; 2018.
3. Devlin N, Appleby J. *Getting the Most Out of PROMs - Putting Health Outcomes at the Heart of NHS Decision-Making*. London, United Kingdom: The King's Fund; 2010.
4. Basch E, Spertus J, Dudley RA, et al. Methods for developing patient-reported outcome-based performance measures (PRO-PMs). *Value Health*. 2015;18(4):493–504.
5. Ahmed S, Berzon RA, Revicki DA, et al. The use of patient-reported outcomes (PRO) within comparative effectiveness research: implications for clinical practice and health care policy. *Med Care*. 2012;50(12):1060–1070.
6. Black N. Patient reported outcome measures could help transform healthcare. *BMJ*. 2013;346:f167.
7. Smith PC, Street AD. On the uses of routine patient-reported health outcome data. *Health Econ*. 2013;22(2):119–131.
8. Van Der Wees PJ, Nijhuis-Van Der Sanden MW, Ayanian JZ, Black N, Westert GP, Schneider EC. Integrating the use of patient-reported outcomes for both clinical practice and performance measurement: views of experts from 3 countries. *Milbank Q*. 2014;92(4):754–775.
9. Nelson EC, Hvitfeldt H, Reid R, et al. *Using Patient-Reported Information to Improve Health Outcomes and Health Care Value: Case Studies From Dartmouth, Karolinska and Group Health*. Lebanon, NH: The Dartmouth Institute for Health Policy and Clinical Practice, Center for Population Health; 2012.
10. Emilsson L, Lindahl B, Koster M, Lambe M, Ludvigsson JF. Review of 103 Swedish healthcare quality registries. *J Intern Med*. 2015;277(1):94–136.
11. Prodinger B, Taylor P. Improving quality of care through patient-reported outcome measures (PROMs): expert interviews using the NHS PROMs Programme and the Swedish quality registers for knee and hip arthroplasty as examples. *BMC Health Serv Res*. 2018;18(1):87.
12. Wohlin J. SVEUS-National collaboration for value-based reimbursement and monitoring of healthcare in Sweden. OECD Expert group meeting on payment systems. https://www.oecd.org/els/health-systems/Item7a_SVEUS-National-collaboration-for-value-based-reimbursement-and-monitoring-of-health-care-in-Sweden_Wohlin.pdf. Accessed February 15, 2022.
13. Patient reported outcome measures (PROMs). NHS England. https://www.england.nhs.uk/statistics/statistical-work-areas/proms/. Accessed December 22, 2020; 2009.
14. Gutacker N, Street A, Gomes M, Bojke C. Should English healthcare providers be penalised for failing to collect patient-reported outcome measures? A retrospective analysis. *J R Soc Med*. 2015;108(8):304–316.
15. Porter ME, Kaplan RS. *How should we pay for health care?* Harvard Business School; 2015. https://www.hbs.edu/faculty/Pages/item.aspx?num=48340. Accessed December 22, 2020.
16. Porter ME, Kaplan RS. How to pay for health care. *Harv Bus Rev*. 2016;94(7-8):88–100.
17. Campanella P, Vukovic V, Parente P, Sulejmani A, Ricciardi W, Specchia ML. The impact of Public Reporting on clinical outcomes: a systematic review and meta-analysis. *BMC Health Serv Res*. 2016;16:296.
18. Lavallee DC, Chenok KE, Love RM, et al. Incorporating patient-reported outcomes into health care to engage patients and enhance care. *Health Aff*. 2016;35(4):575–582.
19. Comprehensive care for joint replacement model. Centers for Medicare & Medicaid Services. https://innovation.cms.gov/innovation-models/cjr. Accessed December 30, 2020.
20. Oncology care model overview and application process. Centers for Medicare & Medicaid Services. https://innovation.cms.gov/files/slides/ocmintro.pdf. Accessed December 30, 2020.
21. Burgess P, Coombs T, Clarke A, Dickson R, Pirkis J. Achievements in mental health outcome measurement in Australia: reflections on progress made by the Australian Mental Health Outcomes and Classification Network (AMHOCN). *Int J Ment Health Syst*. 2012;6(1):4.
22. Thompson C, Sansoni J, Morris D, Capell J, Williams K. *Patient-Reported Outcome Measures: an Environmental Scan of the Australian Healthcare Sector*. Sydney, NSW: Australian Commission on Safety and Quality in Health Care (ACSQHC); 2016.
23. Williams K, Sansoni J, Morris D, Grootemaat P, Thompson C. *Patient-Reported Outcome Measures: Literature Review*. Sydney, Australia: Australian Commission on Safety and Quality in Health Care (ACSQHC); 2016.
24. van Veghel D, Marteijn M, de Mol B, Measurably Better Study Group. The Netherlands) and Advisory Board. First results of a national initiative to enable quality improvement of cardiovascular care by transparently reporting on patient-relevant outcomes. *Eur J Cardiothorac Surg*. 2016;49(6):1660–1669.
25. Canhão H, Faustino A, Martins F, Fonseca JE. Rheumatic Diseases Portuguese Register Board Coordination, Portuguese Society of Rheumatology. Reuma.pt - the rheumatic diseases Portuguese register. *Acta Reumatol Port*. 2011;36(1):45–56.
26. OECD. *Recommendations to OECD Ministers of Health From the High-Level Reflection Group on the Future of Health Statistics - strengthening the international comparison of health system performance through patient-reported indicators*. OECD; 2017. https://www.oecd.org/health/health-systems/Recommendations-from-high-level-reflection-group-on-the-future-of-health-statistics.pdf. Accessed December 30, 2020.
27. ICHOM standard sets. International Consortium for Health Outcomes Measurement. https://www.ichom.org/standard-sets/. Accessed December 30, 2020.
28. Alonso J, Bartlett SJ, Rose M, et al. The case for an international patient-reported outcomes measurement information system (PROMIS(R)) initiative. *Health Qual Life Outcomes*. 2013;11:210.
29. Peters M, Crocker H, Jenkinson C, Doll H, Fitzpatrick R. The routine collection of patient-reported outcome measures (PROMs) for long-term conditions in primary care: a cohort survey. *BMJ Open*. 2014;4(2):e003968.
30. Zrubka Z, Beretzky Z, Hermann Z, et al. A comparison of European, Polish, Slovenian and British EQ-5D-3L value sets using a Hungarian sample of 18 chronic diseases. *Eur J Health Econ*. 2019;20(suppl 1):119–132.
31. Curtis JR, Singh JA. Use of biologics in rheumatoid arthritis: current and emerging paradigms of care. *Clin Ther*. 2011;33(6):679–707.
32. Furst DE, Breedveld FC, Kalden JR, et al. Updated consensus statement on biological agents, specifically tumour necrosis factor {alpha} (TNF{alpha}) blocking agents and interleukin-1 receptor antagonist (IL-1RA), for the treatment of rheumatic diseases, 2005. *Ann Rheum Dis*. 2005;64(suppl 4):iv2–iv14.
33. Fonseca JE, Bernardes M, Canhão H, et al. Portuguese guidelines for the use of biological agents in rheumatoid arthritis - October 2011 update. *Acta Reumatol Port*. 2011;36(4), 385-358.
34. Marra CA, Woolcott JC, Kopec JA, et al. A comparison of generic, indirect utility measures (the HUI2, HUI3, SF-6D, and the EQ-5D) and disease-specific instruments (the RAQoL and the HAQ) in rheumatoid arthritis. *Soc Sci Med*. 2005;60(7):1571–1582.
35. Santos RA, Reis P, Rebelo L, Dias FC, Rosa CM, Queiroz MV. Health assessment questionnaire (versão curta): adaptação para Língua Portuguesa e estudo da sua aplicabilidade. *Acta Reumatol Port*. 1996;XXI(76):15–20.
36. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. *Arthritis Rheum*. 1980;23(2):137–145.
37. Fries JF, Spitz PW, Young DY. The dimensions of health outcomes: the health assessment questionnaire, disability and pain scales. *J Rheumatol*. 1982;9(5):789–793.
38. Williams R. Using the margins command to estimate and interpret adjusted predictions and marginal effects. *STATA J*. 2012;12(2):308–331.
39. Spiegelhalter DJ. Funnel plots for comparing institutional performance. *Stat Med*. 2005;24(8):1185–1202.
40. Spiegelhalter D. Funnel plots for institutional comparison. *Qual Saf Health Care*. 2002;11(4):390–391.
41. Canhao H, Rodrigues AM, Mourao AF, et al. Comparative effectiveness and predictors of response to tumour necrosis factor inhibitor therapies in rheumatoid arthritis. *Rheumatology (Oxford)*. 2012;51(11):2020–2026.
42. Romao VC, Santos MJ, Polido-Pereira J, et al. Comparative effectiveness of tocilizumab and TNF inhibitors in rheumatoid arthritis patients: data from

the rheumatic diseases Portuguese Register, Reuma.pt. *Biomed Res Int*. 2015;2015:279890.
43. Ovretveit J, Zubkoff L, Nelson EC, Frampton S, Knudsen JL, Zimlichman E. Using patient-reported outcome measurement to improve patient care. *Int J Qual Health Care*. 2017;29(6):874–879.
44. Bradley EH, Curry LA, Ramanadhan S, Rowe L, Nembhard IM, Krumholz HM. Research in action: using positive deviance to improve quality of health care. *Implement Sci*. 2009;4:25.
45. O'Hara JK, Grasic K, Gutacker N, et al. Identifying positive deviants in healthcare quality and safety: a mixed methods study. *J R Soc Med*. 2018;111(8):276–291.
46. Black N, Tan S. Use of national clinical databases for informing and for evaluating health care policies. *Health Policy*. 2013;109(2):131–136.
47. Pasquali SK, He X, Jacobs JP, et al. Measuring hospital performance in congenital heart surgery: administrative versus clinical registry data. *Ann Thorac Surg*. 2015;99(3):932–938.
48. Westaby S, Archer N, Manning N, et al. Comparison of hospital episode statistics and central cardiac audit database in public reporting of congenital heart surgery mortality. *BMJ*. 2007;335(7623):759.
49. Prang KH, Maritz R, Sabanovic H, Dunt D, Kelaher M. Mechanisms and impact of public reporting on physicians and hospitals' performance: a systematic review (2000-2020). *PLoS One*. 2021;16(2):e0247297.
50. Strome TL. *Healthcare Analytics for Quality and Performance Improvement*. Hoboken, NJ: Wiley; 2013.
51. Berenson RA, Rice T. Beyond measurement and reward: methods of motivating quality improvement and accountability. *Health Serv Res*. 2015;50(suppl 2):2155–2186.
52. Pross C, Geissler A, Busse R. Measuring, reporting, and rewarding quality of care in 5 nations: 5 policy levers to enhance hospital quality accountability. *Milbank Q*. 2017;95(1):136–183.