

BMJ Open Review of response rates over time in registry-based studies using patient-reported outcome measures

Katherine Wang ^{1,2}, Cathrine N Eftang,³ Rune Bruhn Jakobsen,^{4,5} Asbjørn Årøen^{5,6}

To cite: Wang K, Eftang CN, Jakobsen RB, *et al.* Review of response rates over time in registry-based studies using patient-reported outcome measures. *BMJ Open* 2020;**10**:e030808. doi:10.1136/bmjopen-2019-030808

► Prepublication history for this paper is available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2019-030808>).

Received 02 April 2019
Revised 22 May 2020
Accepted 04 June 2020



© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹University of Oslo Faculty of Medicine, Oslo, Norway

²Oslo Sports Trauma Research Center, Oslo, Norway

³Department of Pathology, Akershus University Hospital, Lorenskog, Norway

⁴Department of Health Management and Health Economics, Institute of Health and Society, University of Oslo Faculty of Medicine, Oslo, Norway

⁵Department of Orthopaedic Surgery, Akershus University Hospital, Lorenskog, Norway

⁶Institute of Clinical Medicine, University of Oslo Faculty of Medicine, Oslo, Norway

Correspondence to

Katherine Wang;
wkath8@gmail.com

ABSTRACT

Objectives Gain an overview of expected response rates (RRs) to patient-reported outcome measures (PROMs) in clinical quality registry-based studies and long-term cohorts in order to better evaluate the validity of registries and registry-based studies. Examine the trends of RRs over time and how they vary with study type, questionnaire format, and the use of reminders.

Design Literature review with systematic search.

Data sources PubMed, MEDLINE, EMBASE, kvalitetsregistre.no, kvalitetsregister.se and sundhed.dk.

Eligibility criteria Articles in all areas of medical research using registry-based data or cohort design with at least two follow-up time points collecting PROMs and reporting RRs. Annual reports of registries including PROMs that report RRs for at least two time points.

Primary outcome measure RRs to PROMs.

Results A total of 10 articles, 12 registry reports and 6 registry articles were included in the review. The overall RR at baseline was 75%±22.1 but decreased over time. Cohort studies had a markedly better RR (baseline 97%±4.7) compared with registry-based data at all time points (baseline 72%±21.8). For questionnaire formats, paper had the highest RR at 86%±19.4, a mix of electronic and paper had the second highest at 71%±15.1 and the electronic-only format had a substantially lower RR at 42%±8.7. Sending one reminder (82%±16.5) or more than one reminder (76%±20.9) to non-responders resulted in a higher RR than sending no reminders (39%±6.7).

Conclusions The large variation and downward trend of RRs to PROMs in cohort and registry-based studies are of concern and should be assessed and addressed when using registry data in both research and clinical practice.

INTRODUCTION

Clinical quality registries (CQRs) have played an important role in monitoring and benchmarking patient outcomes, and have thus improved treatment for patients.¹⁻⁴ A CQR systematically collects data about specific patient groups for a predetermined objective.⁴ Many health authorities have also increased focus on incorporating patient-reported outcome measures (PROMs) when developing patient-centred healthcare systems.⁵ PROMs help to quantitatively assess

Strengths and limitations of this study

- The first review with a systematic search for published response rates of patient-reported outcome measures in registries and prospective cohorts.
- Inclusion of non-peer-reviewed sources such as annual reports from national clinical quality registries.
- Lack of standardisation in this area of research makes it difficult to generate a comprehensive search string to identify all relevant articles. Searches were conducted in multiple databases, however, not all relevant articles were identified in these searches and were instead retrieved from publication lists of registries.
- Including data from both cohorts and registries makes the data heterogeneous, and therefore, difficult to analyse.

a patient's health condition and quality of life and are included in many national and international CQRs to follow up patients over time. Data from properly designed and well-executed registries can be analysed to provide a realistic view of clinical practice and patient outcomes, and compare effectiveness of treatments.^{2 4 6} Applying PROM-data in clinical practice brings many benefits for both patients and physicians, such as more objective and quantifiable measure of symptoms and complaints, more in-depth understanding of the patient's perspective and better informed decision making through discussions about health-related quality of life.⁶⁻⁸ Furthermore, with increased government spending on healthcare in many countries, CQRs can improve clinical practice at a relatively low cost and have a significant net-positive return on investment.^{8 9}

However, for registries to provide an accurate representation of reality, it is important that there is both a high reporting completeness and a high response rate (RR) to PROM questionnaires, distributed evenly across the patient demographic. Reporting completeness refers to the percentage of all patients

meeting inclusion criteria that are reported to, and included in, the registry. RR is the percentage of patients who fill in and submit PROMs at baseline and subsequent follow-ups. High reporting completeness and RR provide a large data sample and increased statistical power, and are therefore essential for society to benefit from registries and for registries to maintain a positive cost–benefit ratio.⁶ However, there is no standard definition of RR and no uniform formula by which RRs are calculated. The American Association for Public Opinion Research reports six different methods for calculating RR.¹⁰ One of the main differences among methods for calculating RR is whether deceased patients or patients who did not receive the questionnaire from the sample are excluded.¹¹

Many registries aim for a reporting completeness of over 90% and an RR of at least 80% but this is difficult to achieve. RRs to PROMs tend to decrease over time, making conclusions less representative in the long term. Previous studies have demonstrated that age, gender and other factors such as socioeconomic status affect RRs.^{12–14} Furthermore, studies have indicated that patients with higher health-related quality of life scores and satisfaction may be over-represented in PROM studies, even in studies with relatively high RRs.^{15 16} These potential selection and attrition biases can affect the resulting estimate of the treatment effect, especially in registries that collect data over multiple time points.^{5 17} There have been several studies on how to increase RR to questionnaires and a number of methods have been identified,¹⁸ yet many registries do not employ these methods.

With an increasing number of registries being established and invested in, it is important to understand what kind of RR to PROMs is realistic and achievable, how RRs change over time and what can be done to improve them. The nature of registry data and PROM follow-ups make them especially useful for comparing methods of treatment where randomised controlled trials (RCTs) are not feasible and long-term prognosis is unknown. Registries are often not limited by the strict inclusion or exclusion criteria used in clinical studies; these criteria may result in selection bias and thereby reduce external validity.¹⁹ Registry manuscripts are increasingly important for clarifying crucial questions debated among healthcare workers, especially in the field of orthopaedic surgery where pain and health-related quality of life are often the deciding factors for surgical treatment.²⁰

One important issue, therefore, is to define the level of evidence these registry-based studies reveal and identify what a realistic RR may be in manuscripts using registry data to test a hypothesis. RR is not the only indicator of the quality of a registry, but its easily measurable and comparable qualities make it suitable as a primary outcome. There have been numerous studies demonstrating the beneficial effects of CQRs but none, to our knowledge, that review RRs to PROMs in registries. This article aims to review the available literature and gain an overview of the expected RR over time in a registry-based

study, including PROMs, in order to better evaluate the validity of future registries and registry-based studies.

METHODS

We obtained the data for this study through two systematic searches of peer-reviewed papers, as well as from annual reports from CQRs. The search protocol was adapted from Øglund *et al.*²¹ Search 1 was performed on 11 April 2017 in PubMed (www.ncbi.nlm.nih.gov), and search 2 was performed on 12 December 2017 in MEDLINE and EMBASE through Ovid (www.ovid.com). Search terms can be found in [table 1](#).

Inclusion criteria

We included journal articles for studies conducted in humans after 1 January 1990 published in English and Scandinavian languages. Articles in all areas of medical research that used PROMs and reported RR for at least two time points were included in order to understand change in RR over time. We were mainly interested in articles using registry-based data, but also included prospective cohort studies with a follow-up of at least 1 year, as we believe long-term cohorts have a similar logistical administration as registries. For registry-based articles, we accepted any follow-up time. We excluded RCTs as the nature and logistics of an RCT do not reflect the logistics of a registry. We also excluded articles where children under 16 years were respondents, articles with patients unable to answer the PROMs themselves or where the PROMs were filled in at control appointments as these factors can affect the RR. In the case of studies with multiple articles for follow-up time points, all related articles reporting RR were included.

Registry search

As we did not identify a large number of journal articles meeting inclusion criteria in the systematic searches, we supplemented them with annual reports by CQRs. We searched through all CQRs in Scandinavia listed on each country's national registry websites kvalitetsregistre.no, kvalitetsregister.se and sundhed.dk, as well as all national knee ligament or joint registries known to the authors. We also searched any registries mentioned in articles from the previous systematic searches, as well as the list of published articles from each registry to identify articles with RRs that were not identified in our previous searches. If the RRs to PROMs were described in the annual report but an article reported a different RR, both were included.

Review

The articles were assessed by two independent researchers to ensure inclusion of all relevant articles. We first performed a title review, followed by an abstract review, and lastly, a full-text review. In search 1, these phases were performed in EndNote by both primary researchers. The results were assessed after completion and an agreement

Table 1 Terms used in the systematic searches

Search 1 in PubMed	Search 2 in EMBASE and Ovid MEDLINE
1. ("register" OR "registry" OR "prospective cohort" OR "prospective cohort study")	1. exp Registries/or exp Register/
2. ("PROMs" OR "subjective outcome" OR "patient outcome" OR "clinical symptoms")	2. (register* or registr*).tw.
3. ("compliance" OR "response rate")	3. prospective cohort*.tw.
4. 1 and 2 and 3	4. 1 or 2 or 3
	5. exp Patient Reported Outcome Measures/ or exp Patient Outcome Assessment/or exp "Outcome Assessment (Health Care)"/ or exp Patient Reported Outcome/ or exp Outcome Assessment/ or exp "Outcome and Process Assessment (Health Care)"/
	6. patient-reported outcome.tw.
	7. PROM.tw.
	8. exp self report/
	9. self report*.tw.
	10. exp Questionnaires/ or exp Questionnaire/
	11. exp Survey/and Questionnaires/
	12. subjective outcome*.tw.
	13. patient outcome*.tw.
	14. clinical symptom*.tw.
	15. response rate.tw.
	16. compliance.tw.
	17. 15 or 16
	18. randomized controlled stud*.mp.
	19. randomized controlled trial*.mp.
	20. rct.mp
	21. 18 or 19 or 20
	22. (retrospective or qualitative stud*).mp.
	23. 21 or 22
	24. 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14
	25. 4 and 17 and 24
	26. 25 not 23

was made in case of any differences. In search 2, these phases were conducted in Covidence (www.covidence.org). The annual reports from CQRs were assessed by only one of the researchers. Due to the nature of the annual reports as describing the activity of a registry, we considered it unnecessary for two researchers to assess them. Risk of bias was not assessed as this was deemed not relevant for the main outcome measure.

Data extraction

Relevant data from the included articles were extracted by the same two researchers to ensure accuracy. Data from annual reports were extracted by one researcher as they were presented more clearly. Data extracted from the articles and reports include year of publication, lead author, format of questionnaires, number of reminders sent, and RRs and their time points. We also collected data on which questionnaires were sent at each follow-up and if this changed between follow-ups, but these data were not readily available or specified clearly in all of articles and was therefore not included in this paper. Methods used to calculate RR were not explicitly stated in most of the articles and were therefore not collected. When subgroups within articles or reports had separate RRs reported, these were treated as separate subgroups instead of averaging the RR. For example, some studies reported separate RRs for different types of PROMs or

for different patient groups. Where indicated, the RR was calculated by a researcher using the available numbers. The main outcome measure was RRs to PROMs. If the relevant data were not described in the article or report, an email was sent to one of the authors or the registry with a request for the missing information.

Statistics

The results are presented with descriptive statics (average \pm SD) and figures using IBM SPSS Statistics V.25. Microsoft Excel was used to visualise the data. No statistical tests were performed to compare the values due to the heterogeneity in the reporting of RRs. No meta-analysis was conducted.

Patient and Public Involvement

Patients or the public were not involved in the designing, conducting, reporting or dissemination of our research.

RESULTS

A total of 5379 articles and 219 registries were identified, and 10 articles from the systematic searches, 12 registry reports and 6 registry articles were included in the final review (figures 1 and 2). Five of the 10 (50%) articles identified in the systematic searches were registry based, and 21 of the 28 (75%) articles and reports included

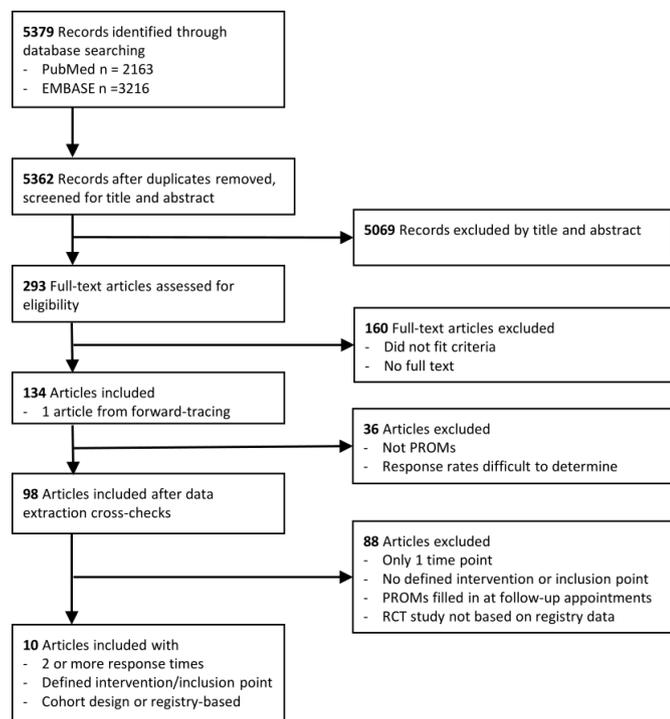


Figure 1 Flow chart of article selection from systematic searchers. PROMs, patient-reported outcome measures; RCT, randomised controlled trial.

were in the field of orthopaedic surgery. The included articles and annual reports are outlined in table 2. We also looked at the effect of reminders and the format of the questionnaires, with more than one reminder and paper-only format being the most common. Additionally, four studies changed the number of PROMs sent at subsequent follow-ups. With regard to RR calculation, eight articles did not report how RR was calculated, while seven excluded deceased patients or those lost to follow-up and only one study included them. We did not find any mention of how RR was calculated in any of the registry annual reports.

Overall RR

The overall RR at baseline starts at $75\% \pm 22.1$ but decreases over time. This average includes all RRs defined by the

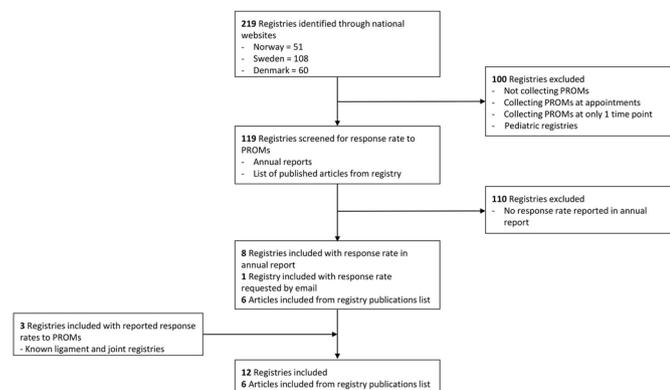


Figure 2 Flow chart of registry annual reports and article selection. PROMs, patient-reported outcome measures.

author to be baseline RRs and RRs to PROMs collected preintervention. The average RRs at different times are shown in table 3. They demonstrate a wide variation in RRs over time and the trend is visualised in figure 3.

Subgroup RR

Table 3 also shows average RRs with SD for all subgroups. When RR in cohort studies was compared with that in registry-based data, the scatter plot indicates a higher average RR for cohorts (figure 4), and the average for cohort studies at baseline of $97\% \pm 4.7$ is far above both the goal RR of 80% and the registry average of $72\% \pm 21.8$. One reminder led to a higher RR at $82\% \pm 16.5$ and a slower decrease compared with no reminders ($39\% \pm 6.7$) or more than one reminder ($76\% \pm 20.9$). Questionnaires on paper had an RR of $86\% \pm 19.4$ at baseline and showed better results than electronic at $42\% \pm 8.7$, or a mix of paper and electronic formats at $71\% \pm 15.1$.

Change in RR per year

Figure 5 shows the amount of change in RR per year. This was calculated by subtracting the final RR from the baseline or first reported RR and dividing the difference by the total length of follow-up time. The data points appear to approach zero with time, indicating that RR decreases less the longer the follow-up time.

DISCUSSION

To our knowledge, this is the first review with systematic searches examining RRs in registry-based articles and cohorts including at least two time points, to gain a better understanding of long-term trends in RRs to PROMs. There is an ongoing debate as to which level of evidence registry-based studies belong in, as there are inherent issues with registries such as a non-homogeneous definition of outcomes and lack of control of confounding factors.²² We believe the current study will be helpful for reviewers assessing registry manuscripts to determine what RR can be expected. Hopefully, this will contribute to an improvement in quality in this field of research.

Considering the goal that many registries have of reaching and maintaining an 80% RR, the average baseline RR for all included articles and reports of 75% is close but with a large SD. As expected, the results depicted a downward trend in RRs over time, but with an unexpected slight increase after the 5 year follow-up time. This was reflected in the 2–5 years RR being the lowest ($50\% \pm 16.1$), and a subsequent increase in the 5–10-years period, with the average reaching $61\% \pm 23.0$. Most likely, this is due to the large number of studies with only up to 2 years of follow-up, resulting in data that is negatively skewed before the 5 year time point. There are seven studies where RR increased more than 1% at subsequent follow-ups. Of these, two studies only sent questionnaires to responders. Another four are orthopaedic studies, where negative symptoms often increase at later stages post-intervention, which could have motivated patients to

Table 2 Included articles and annual registry reports with extracted data

ID	Lead author (year)	Format	No of reminders	RR (%) at follow-up points					Total follow-up time (years)	Change in RR (%/year)*
				#0	#1	#2	#3	#4		
Database searches—prospective cohort studies										
1	Ahmed† ²⁷	Paper	1	90	85	82			2	-4.0
2	Ashley† ²⁸	Mix	>1		85	70	66		1.25	-25.3
3	Blanchard ²⁹	ND	ND	99	85	82	81	50	4	-11.0
4				99	70	57	44	45	4	-13.5
5	Gjeilo ³⁰	Paper	1	100	89	89	82		5	-3.6
6	Pieterse† ³¹	Paper	>1		86	81			2	-5.0
Database searches—registry-based studies										
7	Helsten§,† ³²	Mix	>1		62	71			1	9.1
8	Imam† ¹⁶	Paper	>1	99	91	89	83	79	2	-10.1
9	Olsson ³³	Paper	1	88	82				8	-0.7
10	Porchet† ³⁴	Paper	>1	87	98	91			1	4.0
11	Simony ³⁵	Paper	1	100	84				1	-15.8
Registry articles										
12	Bengtsson ³⁶	Paper	1		93	87			5	-1.2
13	Elkan¶,† ²⁴	Paper	1	78	73	62			2	-8.0
14				100	98	99				-0.6
15	Gjertsen ³⁷	Paper	0		54	49			1	-7.5
16	Kvist ³⁸	Mix	>1	66	50	48	48		5	-3.6
17				61	49	47	46			-3.0
18				72	40	47	38			-6.8
19				68	38	46	38			-6.0
20				61	45	45	42			-3.8
21				58	43	44	42			-3.2
22	Lind† ³⁹	Web	0	46	31				1	-15.0
23	Rahr-Wagner† ⁴⁰	Web	0	33	27				1	-6.5
Registry reports										
24	SweSpin†,§ ⁴¹	Mix	ND		71	57	44		5	-6.7
25	Danish† ⁴²	Web	0	37	31				1	-5.4
26	Danish Hip Arthroscopy Reg ⁴²	Web	>1	52	56	48	34		3	-6.0
27	DaneSpin†,** ⁴³	Paper	1		83	73	59		5	-5.9
28					48	37	24			-6.0
29	Swedish Fracture Reg †,§ ⁴²	Paper	1	55	41				1	-13.7
30	Swedish Hip Arthroplasty Reg†,¶ ⁴²	Mix	1	83	88				1	4.4
31	Swedish Knee Ligament Reg† ²⁵	Mix	1	67	50	47	45	39	10	-2.7
32				62	48	47	44	38		-2.3
33	Norwegian Hip Fracture Reg† ⁴⁴	Paper	0		57	57	55		3	-0.6
34	UK National Ligament Reg ⁴⁵	Mix	ND	58	35	30			2	-14.0
35	UK National Joint Registry†,¶,§ ⁴⁶	Paper	>1	42	44				0.5	4.3
36				100	75	61	84		3	-5.1
37				100	76				0.5	-48.2
38	New Zealand ACL-Reg†,** ²³	Mix	>1	100	86	80	86		2	-7.0
39				100	85	77	83		2	-8.5

Continued



Table 2 Continued

ID	Lead author (year)	Format	No of reminders	RR (%) at follow-up points					Total follow-up time (years)	Change in RR (%/year)*
				#0	#1	#2	#3	#4		
40	Norwegian Knee Ligament Reg†,**	Paper	1		62	58	59		10	-0.4

Follow-up point #0 is the reported baseline time point or time of intervention as specified in the article or report. The subsequent time point intervals (#2, #3...) vary for each article or report and is not the follow-up year. Articles or reports with multiple rows represent separate subgroups with individual RRs reported within the article.

*This is calculated by taking the difference between the last and first RRs and dividing it by the number of follow-up years between the last and first reported RR to show the average change in RR per year.

†Author/registry emailed for reminders and/or format information. ID 1, Rens Kempeneers, Feb 2019 (personal communication); ID 6, Moniek ter Kuile, Feb 2019 (personal communication); ID 10, Anne Mannion, Feb 2019 (personal communication); ID 22, 23, 25, Martin Lind, Feb 2019 (personal communication); ID 27, 28, Karen H. Hansen, Feb 2019 (personal communication); ID 29, Karin Pettersson, Feb 2019 (personal communication); ID 30, Sandra Olausson, Feb 2019 (personal communication); ID 31, 32, Anna Pappas, Feb 2019 (personal communication); ID 33, Kari A. Vågstøl, Feb 2019 (personal communication); ID 35, 36, 37, Beth at NJR Helpdesk, Feb 2019 (personal communication); ID 38, 39, Charlotte Smith, Feb 2019 (personal communication); ID 40, Anne M. Fenstad, Feb 2019 (personal communication).

‡Registry emailed about RRs. ID 27, 28, Karen H. Hansen, Feb 2019 (personal communication); ID 38, 39, Charlotte Smith, Feb 2019 (personal communication); ID 40, Anne M. Fenstad, Feb 2019 (personal communication).

§Some or all RRs calculated by author using numbers available in article or report.

¶Follow-up questionnaires only sent to previous responders.

**Change in number of PROMs or questions sent at subsequent follow-ups.

ND, not determined; PROMs, patient-reported outcome measures; RR, response rate.

fill in the PROMs to report their symptoms. The last study by Porchet *et al* appears to have only sent questionnaires to responders at subsequent follow-ups and contacted some patients to fill in the questionnaire over the phone.

A good CQR should have a high reporting completeness and therefore track all patients that meet inclusion criteria, while a long-term cohort study follows only a selection of patients meeting inclusion criteria over a specific time period. The higher RR in cohort studies compared with that in registry data raises the question of whether long-term PROM follow-ups of all patients included in registries are worth the monetary investment as compared with performing single long-term cohort studies with PROMs. The latter can be nested within registries by predetermining specific cohorts (eg, a selection of patients annually or a full yearly cohort every 5 years) and investing more in improving long-term RRs to PROMs of these cohorts. This is similar to the model used by the New Zealand Joint Registry for their hip and knee arthroplasty patients, where a random selection of patients is

sent the PROMs in order to achieve a 20% RR, which the registry deemed sufficient to provide powerful statistical analyses.²³ The average RR for cohorts is well over 80% at both the baseline and the 1-year follow-up, and just below 80% at the 2 years follow-up. However, many registries have only recently started collecting PROMs and few have had multiple 10-year follow-up collections, so it may be too early to suggest nested cohorts. Furthermore, Elkan *et al* showed that RR did not affect PROM results after lumbar discectomy.²⁴ This indicates that a goal of 80% may not be necessary, but further studies are needed to determine an acceptable limit. The limit may also vary depending on the field, as previous studies have demonstrated significant differences between responders and non-responders both in demographic and in outcomes in various fields.^{12 15}

Another result worthy of further investigation is the higher RR to paper-based questionnaires compared with electronic or mixed formats. With advancements in technology resulting in almost everyone in western

Table 3 Average response rates with SD for time periods

Group	Baseline	≤1 year	>1–2 years	>2–5 years	>5–10 years
All	75±22.1	67±20.9	61±18.4	50±16.1	61±23.0
Cohort	97±4.7	82±7.8	74±11.5	61±18.9	–
Registry	72±21.8	64±21.3	57±18.5	47±13.9	61±23.0
Paper	86±19.4	75±18.6	71±19.4	60±22.1	76±15.3
Web	42±8.7	36±13.4	48	34	–
Mix	71±15.1	62±18.5	53±16.0	43±3.4	39±0.7
Reminders: 0	39±6.7	41±13.3	–	55	–
Reminders: 1	82±16.5	74±20.3	64±20.4	52±19.6	61±23.0
Reminders: >1	76±20.9	68±18.8	58±17.8	47±15.7	–

If there is no SD, the average consists of only one data point.

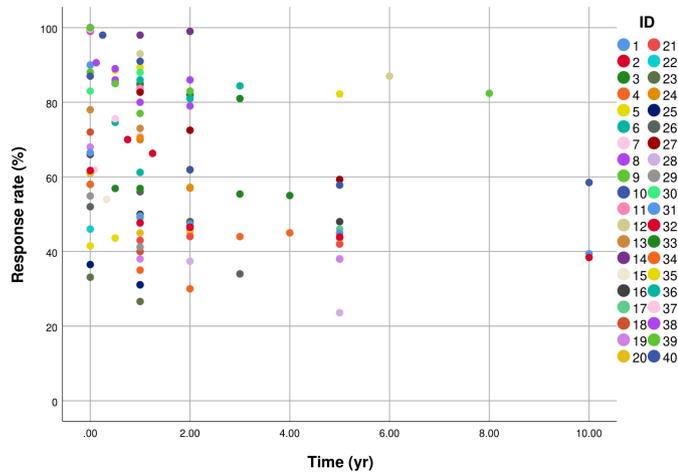


Figure 3 All response rates over time.

populations having easily accessible internet, more and more registries are moving towards web-based PROMs. An online questionnaire requires less time and money to administrate, making it an attractive alternative to collecting paper surveys. However, this approach has been met with problems such as difficulties registering emails or reminders and notifications ending up as spam in already crowded inboxes. Some registries are looking into other alternatives such as social media and mobile applications to better reach patients.²⁵ A recent study found that a mix of electronic and paper surveys achieved a higher RR than only electronic surveys, raising the baseline RR to 100% and the subsequent follow-ups at 3, 6 or 12 months to 83%, from ca. 55%.²⁶ However, this came at an increased cost, which raises the question of whether the additional paper survey costs are justifiable,²⁶ and further highlights the need for identifying an RR that validates PROM data.

With regard to reminders, our results reflect conclusions from previous studies that at least one reminder improves RR.¹⁸ However, the lack of studies with a follow-up longer than 3 years with more than one reminder means we cannot suggest with any certainty how many reminders to send for longer follow-ups. We also did not examine the effect of the number of follow-ups on RRs and this

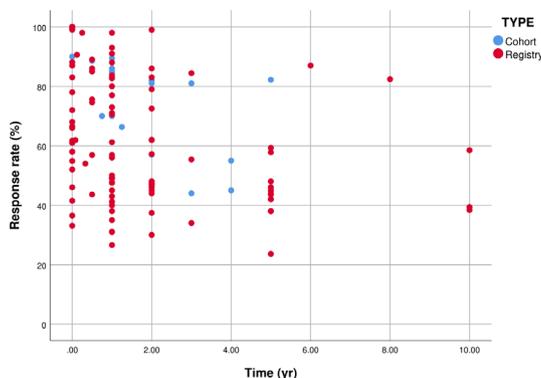


Figure 4 Comparing response rate over time for cohort studies versus registry-based data.

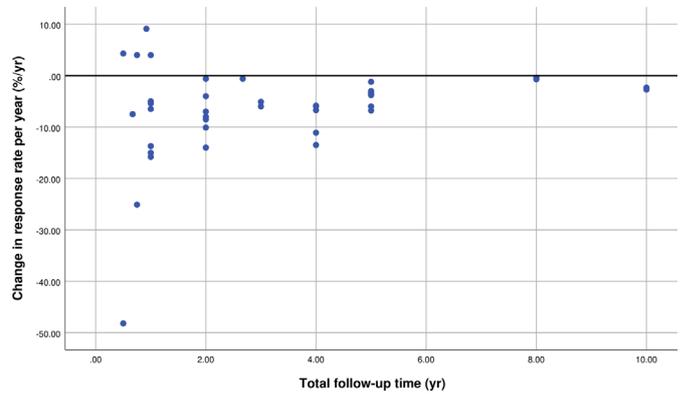


Figure 5 Average change in response rate per year over total follow-up time for each article and report.

could also affect the number of reminders that should be sent. If the follow-ups are more frequent, more than one reminder may negatively affect the RR, while more than one reminder may be necessary to maintain or improve the RR for more spread-out follow-ups.

Figure 5 shows that the change in RR decreases over time and appears to approach zero. This suggests that if a registry starts with a low overall RR that does not decrease significantly at subsequent follow-ups and is consistent within a certain subgroup, those results can potentially still be representative for that subgroup. Thus, the results would still provide important information regarding outcomes and treatment decisions for that specific subgroup, but would not support any conclusions for the general patient population in the registry.

Limitations

There are a number of limitations to this review. RRs were not uniformly calculated and reported across all sources. Fifteen of the studies and registries had insufficient information regarding processes for sending out questionnaires and reminding patients, and eight articles did not report how RRs were calculated. Four studies and registries only sent follow-up questionnaires to previous responders and three changed the number of PROMs at subsequent follow-ups. The heterogeneity of the RR calculation methods makes direct statistical comparison impossible, however, we chose to visualise the data points to better understand the trends in RR over time. In order to improve this research area and create better grounds for future research, we made some general comparisons such as calculating averages despite weak statistical grounds.

Although we wanted to focus on the RRs to PROMs in registries, a large portion of registries did not report the RRs even though they are a key factor in the validity of the data and results in their annual reports. The registries we examined were mainly Scandinavian with the inclusion of only a few other registries from the UK and New Zealand to help supplement the data collected from articles. Further studies are needed to systematically evaluate trends in RRs across Europe, USA and other countries. We did not examine how the number of questions

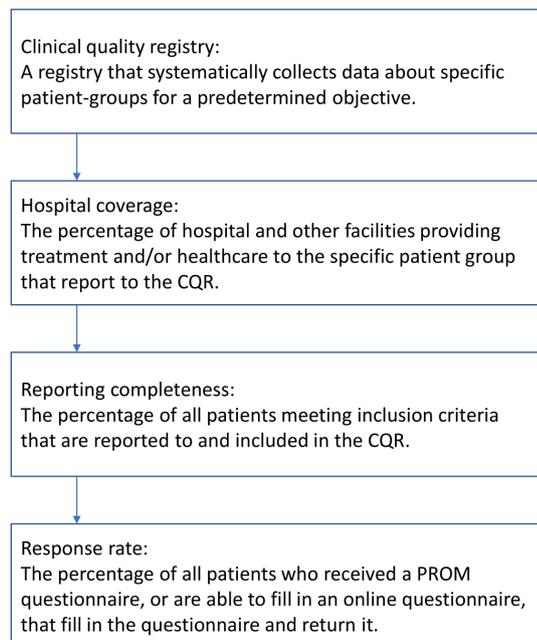


Figure 6 Suggested standardised terminology for registries and articles reporting rrs to PROMs. PROMs, patient-reported outcome measures.

or length of questionnaires affected the RRs as longer questionnaires are known to decrease RR.¹⁸ Lastly, we did not consider the effect of the number of follow-ups on RRs as this may also affect how willing a patient is to respond in the long term. This is especially important for some orthopaedic interventions that may give promising results in the short term but lead to decreased quality of life in the long term. We recognise that these limitations restrict the generalisability of the results, but the intention of this review was mainly to gain an overview of this area of research and make recommendations based on the limitations we found to improve the foundation for future research.

Standard terminology and recommendations for registries

Based on the data collected, the literature we have reviewed, and the authors' experience with registries, we have created a set of recommendations to improve standards for future research in this area. We also noticed a lack of consistent terminology across articles and registries, and therefore, have created a suggested standard terminology as depicted in figure 6.

- ▶ If a study involves PROMs, RRs should always be reported. If there are grounds for not reporting RR, they should be explained.
- ▶ All CQRs should clearly report RRs to PROMs according to the recommendations specified in this article.
- ▶ All studies involving PROMs should specify the number of questionnaires, which (standardised) questionnaires are included and if they have been modified or if only parts or specific questions are used.
- ▶ The method for calculating RR should always be stated. We also recommend including a flow chart of

patient inclusion and response similar to a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow chart for systematic reviews. Examples can be found in the studies by Gjeilo *et al* and Olsson *et al*.

– When possible and applicable, RR should always be reported as the percentage of the total number of patients who are able to fill in a questionnaire, who fill in the questionnaire and return it. This total number should not include patients who have passed away or patients who are not reachable (eg, the wrong address), or be affected by whether questionnaires were filled incompletely and/or correctly. These factors are not included so that the RR accurately represents the percentage of all patients who are able to fill in the questionnaire that actively constitute the study sample. However, if the study protocol or relevant statistical analysis does not allow for this formula, we maintain that the method of calculating RR should still be clearly stated.

- ▶ We recommend sending follow-up questionnaires to all patients able to fill in a questionnaire, and not only to the patients who responded to previous questionnaires. This may help decrease non-response bias, and previous non-responders can still contribute to future follow-up time points.
- ▶ We recommend all registries send at least one reminder as that has been shown to improve RR.
- ▶ Although paper format produced a higher RR in this study, we recognise this may not be sustainable as more registries move to electronic questionnaires. Instead, we recommend baseline PROMs to be filled in pre-intervention at the hospital as this has been shown to increase post-intervention participation.¹⁴

CONCLUSIONS

The large variation and downward trend of RRs to PROMs over time in cohorts, registry-based studies and registries are of concern. There are actions CQRs can take in order to improve their RRs, such as sending at least one reminder or implementing a mixed electronic and paper approach. However, further research is warranted to clarify an acceptable RR in order for PROM data in registry-based studies to be used to inform treatment decisions.

Contributors CNE, RBJ and AA devised the project and the main conceptual ideas. CNE and KW performed the searches and data extraction. KW analysed the data and wrote the manuscript with support from CNE and consultation with RBJ and AA.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository. Extra data can be accessed via the Dryad data repository at <http://datadryad.org/> with the doi: 10.5061/dryad.hhmgqkpcp.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iD

Katherine Wang <http://orcid.org/0000-0003-2209-4549>

REFERENCES

- Engebretsen L, Forssblad M, Lind M. Why registries analysing cruciate ligament surgery are important. *Br J Sports Med* 2015;49:636–8.
- Liu FX, Rutherford P, Smoyer-Tomic K, et al. A global overview of renal registries: a systematic review. *BMC Nephrol* 2015;16:31.
- Stey AM, Russell MM, Ko CY, et al. Clinical registries and quality measurement in surgery: a systematic review. *Surgery* 2015;157:381–95.
- Agency for Healthcare Research and Quality (US). *Registries for evaluating patient outcomes: a user's guide*. Rockville (MD): Agency for Healthcare Research and Quality (US), 2014.
- 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;50:1060–70.
- Wilcox N, McNeil JJ. Clinical quality registries have the potential to drive improvements in the appropriateness of care. *Med J Aust* 2016;205:S27–9.
- Trotter JP. *Patient registries: a new gold standard for "real world" research*, 2002.
- ACSQHC. *Economic evaluation of clinical quality registries: final report*. Sydney: The Australian Commission on Safety and Quality in Health Care, 2016.
- Larsson S, Lawyer P, Garellick G, et al. Use of 13 disease registries in 5 countries demonstrates the potential to use outcome data to improve health care's value. *Health Aff* 2012;31:220–7.
- The America Association for Public Opinion Research. Standard definitions: final dispositions of case codes and outcome rates for surveys, 2016. Available: [https://www.aapor.org/Standards-Ethics/Standard-Definitions-\(1\).aspx](https://www.aapor.org/Standards-Ethics/Standard-Definitions-(1).aspx)
- Fincham JE. Response rates and responsiveness for surveys, standards, and the Journal. *Am J Pharm Educ* 2008;72:43.
- Reinholdsson J, Kraus-Schmitz J, Forssblad M, et al. A non-response analysis of 2-year data in the Swedish knee ligament register. *Knee Surg Sports Traumatol Arthrosc* 2017;25:2481–7.
- Juto H, Gärtner Nilsson M, Möller M, et al. Evaluating non-responders of a survey in the Swedish fracture register: no indication of different functional result. *BMC Musculoskelet Disord* 2017;18:278.
- Patel J, Lee JH, Li Z, et al. Predictors of low patient-reported outcomes response rates in the California joint replacement registry. *J Arthroplasty* 2015;30:2071–5.
- de Rooij BH, Ezendam NPM, Mols F, et al. Cancer survivors not participating in observational patient-reported outcome studies have a lower survival compared to participants: the population-based profiles registry. *Qual Life Res* 2018;27:3313–24.
- Imam MA, Barke S, Stafford GH, et al. Loss to follow-up after total hip replacement: a source of bias in patient reported outcome measures and registry datasets? *Hip Int* 2014;24:465–72.
- Hamrin Senorski E, Svantesson E, Engebretsen L, et al. 15 years of the Scandinavian knee ligament registries: lessons, limitations and likely prospects. *Br J Sports Med* 2019;53:1259–60.
- Edwards PJ, Roberts I, Clarke MJ, et al. Methods to increase response to postal and electronic questionnaires. *Cochrane Database Syst Rev* 2009;3:Mr000008.
- Juliusson G, Lazarevic V, Hörstedt A-S, et al. Acute myeloid leukemia in the real world: why population-based registries are needed. *Blood* 2012;119:3890–9.
- Rolfson O, Rothwell A, Sedrakyan A, et al. Use of patient-reported outcomes in the context of different levels of data. *J Bone Joint Surg Am* 2011;93 Suppl 3:66–71.
- Øglund GP, Hildebrand M, Ekelund U. Are birth weight, early growth, and motor development determinants of physical activity in children and youth? A systematic review and meta-analysis. *Pediatr Exerc Sci* 2015;27:441–53.
- Grassi A, Zaffagnini S. Meta-analysis of the risk of infections after anterior cruciate ligament reconstruction by graft type: letter to editor. *Am J Sports Med* 2018;46:NP20–1.
- New Zealand Orthopaedic Association. New Zealand joint registry 19 year report - Jan 1999 to Dec 2017, 2018. Available: <https://nzoa.org.nz/nzoea-joint-registry>
- Elkan P, Lagerbäck T, Möller H, et al. Response rate does not affect patient-reported outcome after lumbar discectomy. *Eur Spine J* 2018;27:1538–46.
- Swedish Knee Ligament Registry. Annual report 2017 [Årsrapport 2017], 2017. Available: <https://aclregister.nu/documents/>
- Pronk Y, Pilot P, Brinkman JM, et al. Response rate and costs for automated patient-reported outcomes collection alone compared to combined automated and manual collection. *J Patient Rep Outcomes* 2019;3:31.
- Ahmed Ali U, Issa Y, van Goor H, et al. Dutch chronic pancreatitis registry (care): design and rationale of a nationwide prospective evaluation and follow-up. *Pancreatol* 2015;15:46–52.
- Ashley L, Jones H, Thomas J, et al. Integrating patient reported outcomes with clinical cancer registry data: a feasibility study of the electronic patient-reported outcomes from cancer survivors (ePOCS) system. *J Med Internet Res* 2013;15:e230.
- Blanchard P, Davis JW, Frank SJ, et al. Quality of life after brachytherapy or bilateral nerve-sparing robot-assisted radical prostatectomy for prostate cancer: a prospective cohort. *BJU Int* 2018;121:540–548.
- Gjeilo KH, Wahba A, Klepstad P, et al. Survival and quality of life in an elderly cardiac surgery population: 5-year follow-up. *Eur J Cardiothorac Surg* 2013;44:e182–8.
- Pieterse QD, Kenter GG, Maas CP, et al. Self-reported sexual, bowel and bladder function in cervical cancer patients following different treatment modalities: longitudinal prospective cohort study. *Int J Gynecol Cancer* 2013;23:1717–25.
- Helsten DL, Ben Abdallah A, Avidan MS, et al. Methodologic considerations for collecting patient-reported outcomes from unselected surgical patients. *Anesthesiology* 2016;125:495–504.
- Olsson A, Sandblom G, Fränneby U, et al. Impact of postoperative complications on the risk for chronic groin pain after open inguinal hernia repair. *Surgery* 2017;161:509–16.
- Porchet F, Bartanusz V, Kleinstueck FS, et al. Microdiscectomy compared with standard discectomy: an old problem revisited with new outcome measures within the framework of a spine surgical registry. *Eur Spine J* 2009;18:360–6.
- Simony A, Hansen KH, Ernst C, et al. [Implementation of the Danish national database Danespine for spinal surgery]. *Ugeskr Laeger* 2014;176:V01130019.
- Qvarnström M, Kahan T, Kieler H, et al. Persistence to antihypertensive drug classes: a cohort study using the Swedish primary care cardiovascular database (SPCCD). *Medicine* 2016;95:e4908.
- Gjertsen J-E, Baste V, Fevang JM, et al. Quality of life following hip fractures: results from the Norwegian hip fracture register. *BMC Musculoskelet Disord* 2016;17:265.
- Kvist J, Kartus J, Karlsson J, et al. Results from the Swedish national anterior cruciate ligament register. *Arthroscopy* 2014;30:803–10.
- Lind M, Menhert F, Pedersen AB. The first results from the Danish ACL reconstruction registry: epidemiologic and 2 year follow-up results from 5,818 knee ligament reconstructions. *Knee Surg Sports Traumatol Arthrosc* 2009;17:117–24.
- Rahr-Wagner L, Thillemann TM, Lind MC, et al. Validation of 14,500 operated knees registered in the Danish knee ligament reconstruction register: registration completeness and validity of key variables. *Clin Epidemiol* 2013;5:219–28.
- SweSpine. Annual report [Årsrapport], 2018. Available: <http://www.swespine.se/Aarsrapporter.htm>
- Swedish Hip Arthroplasty Registry. Annual report [Årsrapport], 2017. Available: <https://shpr.registercentrum.se/shar-in-english/annual-reports/p/rkeyyeElz>
- DaneSpine. Spine surgery annual report [Ryggekirurgi Årsrapport], 2017. Available: <http://drksdanespine.dk/wm420129>
- National Hip Fracture Registry. Report [Rapport], 2018. Available: <http://nrlweb.ihelse.net/>
- UK National Ligament Registry. The fourth annual report, 2018. Available: <https://www.uknrl.co.uk/media.php>
- UK National Joint Registry. 12th annual report, 2015. Available: <http://www.njrcentre.org.uk/njrcentre/NJR-12th-Annual-Report>