

Predicting non-response in patient-reported outcome measures: results from the Swiss quality assurance programme in cardiac inpatient rehabilitation

STEFANIE KÖHN¹, ANNA SCHLUMBOHM¹, MANUELA MARQUARDT¹, ANKE SCHEEL-SAILER², STEPHAN TOBLER³, JAN VONTOBEL⁴ and LUISE MENZI⁵

¹Abteilung Rehabilitationsforschung, Institute of Medical Sociology and Rehabilitation Science, Charité—Universitätsmedizin Berlin, Charitéplatz 1, Berlin 10117, Germany

²Swiss Paraplegic Centre, Guido-A.-Zäch-Strasse 1, Nottwil CH-6207, Switzerland

³Kliniken Valens, Taminaplatz 1, Valens CH-7317, Switzerland

⁴Hochgebirgsklinik Davos AG, Herman-Burchard-Strasse 1, Davos CH-7265, Switzerland

⁵Swiss National Association for Quality Development in Hospitals and Clinics - ANQ, Weltpoststrasse 5, Bern CH-3015, Switzerland

Address reprint requests to: Stefanie Köhn, Institut für Medizinische Soziologie und Rehabilitationswissenschaften, Abteilung Rehabilitationsforschung, Charité—Universitätsmedizin Berlin, Charitéplatz 1, Berlin 10117, Germany. Tel: +49 30 450 517 156; Fax: +49 30 450 517 932; E-mail: stefanie.koehn@charite.de

Abstract

Background: Quality assurance programmes measure and compare certain health outcomes to ensure high-quality care in the health-care sector. The outcome of health-related quality of life is typically measured by patient-reported outcome measures (PROMs). However, certain patient groups are less likely to respond to PROMs than others. This non-response bias can potentially distort results in quality assurance programmes.

Objective: Our study aims to identify relevant predictors of non-response during assessment using the PROM MacNew Heart Disease questionnaire in cardiac rehabilitation.

Methods: This is a cross-sectional study based on data from the Swiss external quality assurance programme. All patients aged 18 years or older who underwent inpatient cardiac rehabilitation in 16 Swiss rehabilitation clinics between 2016 and 2019 were included. Patients' socio-demographic and basic medical data were analysed descriptively by comparing two groups: non-responders and responders. We used a random intercept logistic regression model to estimate the associations of patient characteristics and clinic differences with non-response.

Results: Of 24 572 patients, there were 33.3% non-responders and 66.7% responders. The mean age was 70 years, and 31.0% were women. The regression model showed that being female was associated with non-response [odds ratio (OR) 1.22; 95% confidence interval (CI) 1.14–1.30], as well as having no supplementary health insurance (OR 1.49; 95% CI 1.39–1.59). Each additional year of age increased the chance of non-response by an OR of 1.02 (95% CI 1.02–1.02). Not being a first language speaker of German, French or Italian increased the chance of non-response by an OR of 6.94 (95% CI 6.03–7.99). Patients admitted directly from acute care had a higher chance of non-response (OR 1.23; 95% CI 1.10–1.38), as well as patients being discharged back into acute care after rehabilitation (OR 3.89; 95% CI 3.00–5.04). Each point on the cumulative illness rating scale total score increased the chance of non-response by an OR of 1.05 (95% CI 1.04–1.05). Certain diagnoses also influenced the chance of non-response. Even after adjustment for known confounders, response rates differed substantially between the 16 clinics.

Conclusion: We have found significant non-response bias among certain patient groups, as well as across different treatment facilities. Measures to improve response rates among patients with known barriers to participation, as well as among different treatment facilities, need to be considered, particularly when PROMs are being used for comparison of providers in quality assurance programmes or outcome evaluation.

Key words: non-response, patient-reported outcome measures, cardiac rehabilitation, quality assurance, health care, Switzerland

Introduction

In recent years, patient-reported outcome measures (PROMs) have received increased attention as a tool to measure health-care outcomes. PROMs focus on patients' perspectives and are the only way to address personal satisfaction and health-related quality of life (HRQOL). The use of PROMs is associated with many useful applications in clinical care (i.e. shared decision-making and individualized treatment

adjustments) and quality improvement [1–3]. In external quality assurance programmes, PROMs are used to compare treatment results by different health-care providers; in health-care systems, they are used as indicators for public reporting and performance-based contracting [4, 5]. Several PROM initiatives increasingly promote the standardization and institutionalization of the patients' perspective in the health systems (e.g. European Organisation for Research and

Treatment of Cancer, International Consortium for Health Outcome Measurement, OECD Patient-Reported Indicator Surveys and Patient-Reported Outcomes Measurement Information System). This focus on the patients' perspective when measuring health was closely related to the implementation of the International Classification of Functioning, Disability and Health (ICF) in 2001 [6] to promote the biopsychosocial model of health. Although PROMs have great potential to anchor the patient perspective more firmly in medical care and quality assurance, their usability has limitations due to high rates of non-response [7, 8]. Non-response is not necessarily problematic for the representativeness of data; however, it becomes problematic when non-response is systematic and certain groups are under-represented (non-response bias) [9–11]. For national health reporting or comparisons of health providers, it becomes particularly relevant if these subgroups, for example, are in poorer condition or achieve better treatment results than the responders. Systematic non-response biases such as these can lead to over- or under-estimation of health status or treatment outcomes [12, 13].

The rehabilitation sector has a long tradition of collecting PROMs due to its early adoption of the biopsychosocial model offered by the ICF. External quality assurance programmes in rehabilitation in Germany and Switzerland have used PROMs as outcome quality indicators for many years already, such as the MacNew Heart Disease questionnaire (MacNew Heart), to measure HRQOL of patients with heart disease [14, 15].

In this paper, we analysed MacNew Heart response rates of patients undergoing cardiac rehabilitation in Switzerland. The aim was to explore whether a non-response bias exists and whether this non-response bias is due to specific patient characteristics, features of the clinic or both.

Methods

Study design and setting

In 2009, the Swiss National Association for Quality Development in Hospitals and Clinics (ANQ) was founded as a collaboration of all major partners in the Swiss health-care system (Swiss Hospital Association H+, insurance associations, all 26 Swiss cantons and the Principality of Liechtenstein) (www.anq.ch). The ANQ is responsible to conduct external quality assurance for the Swiss inpatient health sector and, amongst others, for rehabilitation clinics. The rehabilitation sector is further divided into several functional departments, for which the ANQ developed detailed measurement schemes. In order to measure HRQOL, the ANQ decided to use PROMs within a number of departments, alongside functional assessments reported by clinicians and performance tests. For quality assurance within cardiac rehabilitation, the chosen PROM is the MacNew Heart.

For this cross-sectional study, we analysed data from all patients within the functional department of cardiac rehabilitation from 2016 to 2019. The cardiac rehabilitation facilities admit patients for inpatient treatment after being in acute care due to cardiac events or surgery, as well as patients with severe chronic cardiac diseases for conservative management. Patients are supported by an interdisciplinary team and receive intensive, multimodal rehabilitation treatment that lasts on average for 3 weeks. After admission and before discharge, the patients were requested to complete the MacNew

Heart. The questionnaire was administered as a paper-pencil licensed version in German, French and Italian, which are official languages of Switzerland.

Additionally, clinics submit socio-demographic information, basic medical data (minimum medical dataset of the Swiss federal statistical office [16]) and a measurement for comorbidity based on the cumulative illness rating scale (CIRS) total score [17] for each patient.

Eligibility criteria

All cardiac rehabilitation patients in Switzerland aged 18 years or older, who were admitted into a rehabilitation clinic between 2016 and 2019 and for whom rehabilitation treatment has been completed, were eligible for this study. Cases where treatment was discontinued or shorter than 7 days were categorized as dropouts and not eligible for analysis. Reasons for discontinuation of treatment were death of a patient, transfer to an acute care clinic for more than 24 hours or a request for discontinuation of treatment by the patient.

Furthermore, patients were only included if socio-demographic, basic medical and CIRS data were available. Additionally, the MacNew Heart had to be sufficiently completed by the patient, or refusal or inability of the patient to complete the questionnaire had to be recorded by the clinic personnel (recorded non-response).

Outcome: recorded non-response of the MacNew Heart

The MacNew Heart consists of 27 items; each item can have values between 1 ('severe restriction') and 7 ('no restriction at all'). The total score is calculated from the mean value of all completed items (a minimum of 50% of the items must be filled in).

If a patient is unable to complete the questionnaire, clinics must record this as a 'non-response'. All patients recorded as non-responders at admission or discharge were categorized as non-responders, whereas all patients who filled in the questionnaire at admission *and* discharge were categorized as responders.

Exposure and covariates

Covariates were collected from socio-demographic and basic medical data, which clinics have to submit to the Swiss federal statistical office for each patient.

We used the following socio-demographic variables as exposure and covariates: age (in years), gender (female or male) and supplementary health insurance (yes or no). Residents in Switzerland may purchase supplementary health insurance on top of their compulsory health insurance, which guarantees a single or double room in hospital and treatment by a head physician. We coded language proficiency of German, French or Italian [L2 (second language) vs L1 (first language) speaker] based on the nationality of each patient. We categorized patients as L1 speakers if their nationality corresponded with a country that has one of these three languages as official language.

Furthermore, we used the following variables to capture comorbidities and health status of the patients: CIRS total score (in points), length of rehabilitation (in days), patient

location before admission (home vs acute care), patient location after discharge (home vs acute care) and main diagnosis of each patient (categorized into eight groups of diagnoses). Finally, we adjusted for the year in which the rehabilitation treatment was concluded (2016–2019).

Reporting of our study followed the Strengthening the Reporting of Observational Studies in Epidemiology checklist [18] and the Good Epidemiological Practice guidelines [19].

Statistical analysis

Statistical analyses were conducted using R 4.1.2 [20], lme4 [21] and performance [22].

We used descriptive statistics to determine patients' characteristics in the full sample, as well as stratified by non-responders versus responders. Additionally, to get an impression of how much the response rates vary by each clinic, we showed the percentages of non-response for each clinic in each year.

The associations of patients' socio-demographic and health characteristics with being categorized as non-responder were tested using a random intercept logistic regression model. We estimated odds ratios (ORs) as well as corresponding 95% confidence intervals (CIs) and *P* values. In the regression model, we used all covariates and stratified by clinic.

Results

Participants

Between 2016 and 2019, data of 29 086 cardiac rehabilitation patients were collected. We excluded 1998 dropouts, as well as 321 patients with incomplete MacNew Heart and 1477 patients with no records for the MacNew Heart. Another 718 patients were excluded because their socio-demographic and basic medical data were missing or incomplete. Our analysis included 24 572 cardiac rehabilitation patients (Figure 1).

Clinics

All 16 cardiac rehabilitation clinics that treated patients between 2016 and 2019 were included in the analyses. Some clinics submitted data for all 4 years, while others only submitted data for a number of years (Figure 2).

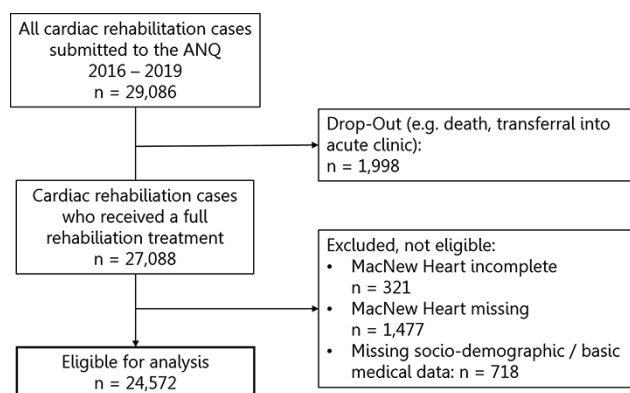


Figure 1 Flow of participants.

Descriptive data

Descriptive statistics are shown for the full sample ($n = 24\,572$), as well as stratified by response status [$n_{\text{non-responders}} = 8172$ (33.3%); $n_{\text{responders}} = 16\,400$ (66.7%)] in Table 1. Non-responders were more often female, less often L1 speakers of Italian, French and German and had less often supplementary health insurance. Non-responders also scored higher on the comorbidity scale (CIRS total score) and were more likely to be discharged back into acute care.

Non-response rates vary largely between the 16 clinics: some clinics have considerably lower non-response rates while others have non-response rates well above the average (Figure 2). There is also variation within each clinic over the years.

Inference statistics

The intraclass correlation coefficient (ICC) in the regression model, the marginal and the conditional R^2 showed that about half of the total explained variance (0.10 of 0.196) can be explained by differences at clinic level (Level 2) (Table 2). The model showed an overall R^2 of 0.196 and thus a high goodness of fit since an R^2 above 0.2 is considered excellent [23].

All socio-demographic predictors were relevant and statistically significant. Female patients had a higher chance of non-response, as well as patients without supplementary health insurance. The chance of non-response was further increased by age. Finally, not being native speaker of German, French or Italian increased the chance of non-response.

A number of medical exposures also showed significant associations: Those admitted directly from acute care had a higher chance of non-response, as well as those being discharged back into acute care instead of being discharged home. Each additional point of CIRS total score increased the chance of non-response. A number of different diagnoses also increased or decreased the chance of non-response, compared to the diagnosis 'Chronic ischaemic heart disease'. Every additional day of rehabilitation treatment increased the chance of non-response slightly. Compared to the year 2016, the chance of non-response remained similar in 2017, decreased in 2018 and increased in 2019.

Discussion

Statement of principal findings

On individual level, our results show that there were significant differences between patients who completed the MacNew Heart (responders) and those who did not (non-responders). The most striking finding was that patients who spoke German, French or Italian (language of the questionnaire) as second language had a higher chance of non-response. Women, older patients and patients without supplementary health insurance also had a higher chance of non-response. Being able to afford supplementary health insurance could indicate higher income and could be interpreted as a proxy for higher socio-economic status.

Furthermore, indicators that directly or indirectly measure disease severity were shown to be important. Patients who received longer rehabilitation treatment and those with a higher comorbidity score had an increased chance of non-response. Patients admitted to a rehabilitation clinic from an

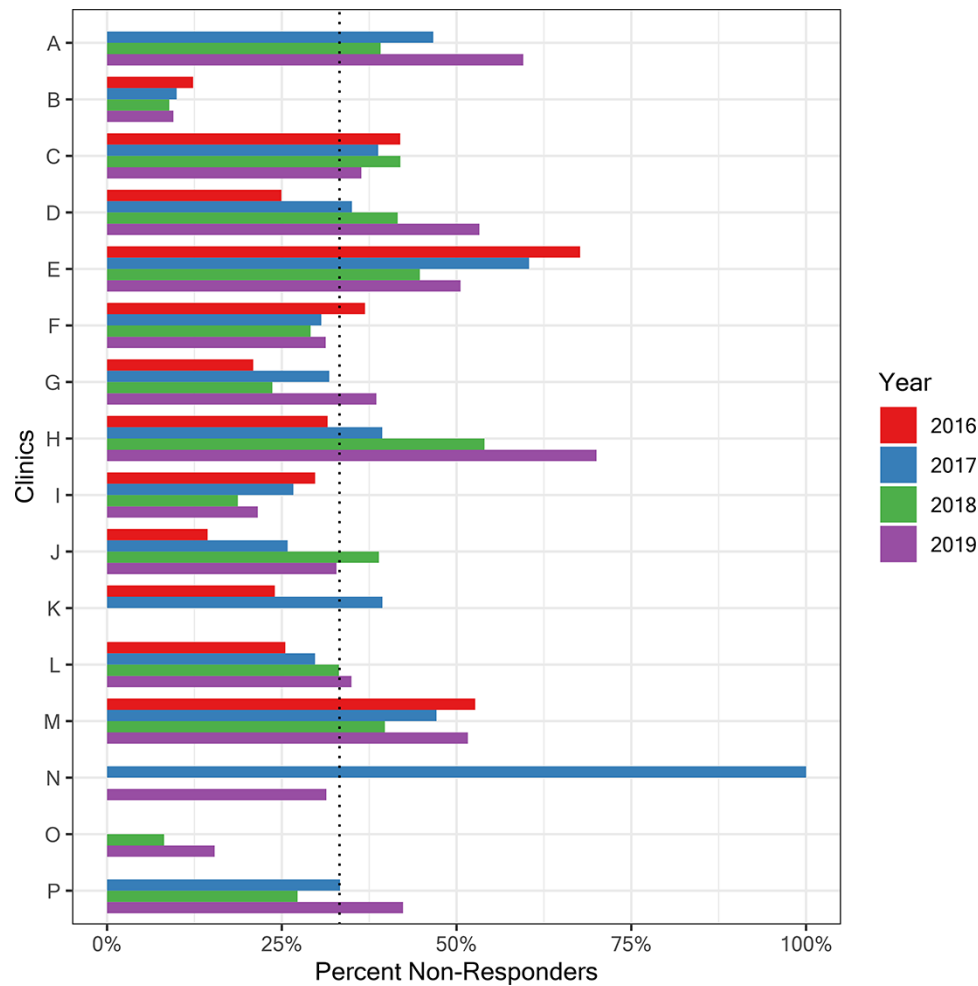


Figure 2 Patient non-response at clinics in percent; dotted line depicts the overall rate of non-response.

acute care hospital and those discharged back into acute care after rehabilitation were also less likely to respond.

On clinical level, our results show that there were substantial differences in response rates between each facility. In fact, they accounted for about half of the explained variance in the full model.

Strengths and limitations

Several limitations of this study must be considered: The available confounding factors do not cover all potential reasons for non-response such as marital status, education and complete information on socio-economic status. Another limitation is potential misclassification of language status. The only information available was the patient's nationality and not their first language. Patients who are Swiss-born or naturalized Swiss citizens may not have the linguistic resources to answer the questionnaire but, according to this metric, would be considered first language speakers. Moreover, a patient may be fluent in German, French or Italian despite holding a passport from a country where these are not official languages.

Furthermore, a lack of information on clinical conditions (number of beds and staffing levels) means that it was difficult to explain possible differences between clinics. Additionally,

the number of Level 2 units (16 clinics) is rather small. Including Level 2 predictors would probably not lead to a better model fit. We thus decided to estimate random intercept logistic regression models despite the small number of Level 2 units since the ICC is rather high [28], and the assumption of independence of errors would thus be violated when using logistic regression models without a second level.

The main strength of our analysis is the dataset on which this study was based. As the Swiss quality assurance programme for inpatient rehabilitation is a comprehensive survey, the data give a thorough overview of all patients in Swiss cardiac inpatient rehabilitation. Socio-demographic and basic medical data were (mostly) available, even for non-responders. The data furthermore included information about where each patient was treated. We can thus show how much of the explained variance is due to individual-level factors (such as age, gender or language status) and how much variation occurs *between* the clinics.

Interpretation within the context of the wider literature

Our results are consistent with existing research that found language barriers to be important predictors of non-response when using PROMs. Burrus *et al.* [10] found being a non-native speaker led to higher non-response rates in vocational

Table 1 Sample characteristics: full sample and stratified by response status MacNew Heart questionnaire

	Full sample (n = 24 572)			Non-responders MacNew Heart (n = 8172)			Responders MacNew Heart (n = 16 400)		
	N	%	Median (IQR)	N	%	Median (IQR)	N	%	Median (IQR)
Age (years)	24 572		70 (45)	8172		71 (44)	16 400		69 (45)
Gender									
Female	7623	31.0		2771	33.9		4852	29.6	
Male	16 949	68.9		5401	66.1		11 548	70.4	
German, Italian or French L1 speaker									
No	1144	4.7		797	9.8		347	2.1	
Yes	23 428	95.3		7375	90.2		16 053	97.9	
Supplementary health insurance									
Yes	17 667	71.9		6489	79.4		11 178	68.2	
No	6905	28.1		1683	20.6		5222	31.8	
CIRS (total score)	24 572		16 (24)	8172		17 (25)	16 400		15 (24)
Length of stay (days)	24 572		20 (23)	8172		20 (25)	16 400		20 (21)
Pre-rehabilitation location									
Acute care	22 677	92.3		7535	92.2		15 142	92.3	
Home	1895	7.7		637	7.8		1258	7.7	
Post-rehabilitation location									
Acute care	297	1.2		193	2.4		104	0.6	
Home	24 275	98.9		7979	97.8		16 296	99.4	
Main diagnosis									
Chronic ischaemic heart disease	7414	30.2		2504	30.6		4910	29.9	
Other ischaemic heart diseases	4963	20.2		1629	19.9		3334	20.3	
Non-rheumatic mitral valve disorders	1827	7.4		485	5.9		1342	8.2	
Non-rheumatic aortic valve disorders	4816	19.6		1489	18.2		3327	20.3	
Other forms of heart disease	2727	11.1		1023	12.5		1704	10.4	
Diseases of arteries	1483	6.0		499	6.1		984	6.0	
Other disorders of the circulatory system	1030	4.2		383	4.7		647	3.9	
Other diseases	312	1.3		160	2.0		152	0.9	
Year									
2016	6033	24.6		1845	22.6		4188	25.5	
2017	6109	24.9		1989	24.3		4120	25.1	
2018	6169	25.1		2006	24.5		4163	25.4	
2019	6261	25.5		2332	28.5		3929	24.0	

IQR = interquartile range

rehabilitation patients in Switzerland. This was also found for patients in an orthopaedic outpatient setting in Australia [24], for a cohort with total joint arthroplasty in the USA [8] and for patients who underwent elective surgery in England [11]. Several national health surveys found higher non-response rates among ethnic minorities, immigrants and their descendants [7, 25].

Our finding that patients without supplementary insurance had a higher chance for non-response is in line with a study in which American patients who had Medicare or Medicaid were more often non-responders than patients with commercial medical insurance. Schamber *et al.* [8] interpreted this as a proxy for lower socio-economic status. Additionally, several studies also identified that more deprived people were less likely to respond [7, 11, 25, 26].

The association of non-response with poorer health status and higher levels of comorbidity was also reported in several surveys. A higher need for personal assistance during completion of the PROM was reported for patients with health-related disabilities [8, 10, 11, 26].

Contrary to our findings, several studies found that non-responders are more often male [11, 25, 26], or no association between gender and response rate was reported [7, 8]. Our result that elder patients have higher odds for non-response is in line with the existing research [7, 8, 24]; others found the opposite [11, 26]. These discrepancies

are possibly due to the differing sub-populations that were evaluated.

Based on our results, we cannot conclude whether non-responders have in fact lower outcomes pertaining to HRQOL after rehabilitation treatment, as information on this was not available for non-responders. Only a few studies could show that non-response was systematically associated with a worse overall health condition [7] or lower results in outcome measurements [27, 28]. The observed differences in response rates between clinics (the provider effect) are also in line with the existing research. Two studies [12, 13] have found large differences between single providers in their analysis of pre-operative PROM data from the National Health Service in England. Both studies were unable to report on specific clinic conditions that could potentially be responsible for differing response rates. Further evidence can also be found in an experimental study by Ho *et al.* [26], where eight cohorts of patients were asked to complete a PROM under differing conditions, resulting in non-response rates between 13% and 81% among the eight groups.

Implications for policy, practice and research

Non-response bias may seem like a predominantly methodological problem. However, it has strong implications for the question of how much confidence we should have in

Table 2 Predicting non-response to the MacNew Heart questionnaire—random intercept logistic regression model

Predictors	Random intercept logistic regression model		
	OR	95% CI	<i>P</i>
Intercept	0.02	0.02–0.04	<0.001
Gender: female (ref: male)	1.22	1.14–1.30	<0.001
Age (years)	1.02	1.02–1.02	<0.001
German, Italian or French L1 speaker: yes (ref: no)	6.95	6.04–8.00	<0.001
Supplementary health insurance: yes (ref: no)	1.49	1.39–1.59	<0.001
Length of stay (days)	1.01	1.00–1.01	0.008
Pre-rehabilitation location: acute care (ref: home)	1.23	1.10–1.38	<0.001
Post-rehabilitation location: acute care (ref: home)	4.14	3.12–5.50	<0.001
CIRS (total score in points)	1.05	1.04–1.05	<0.001
Main diagnosis (ref: Chronic ischaemic heart disease)			
Other ischaemic heart diseases	1.01	0.93–1.10	0.858
Non-rheumatic mitral valve disorders	0.85	0.75–0.96	0.010
Non-rheumatic aortic valve disorders	0.97	0.89–1.05	0.436
Other forms of heart disease	1.25	1.13–1.38	<0.001
Diseases of arteries	1.15	1.01–1.31	0.031
Other disorders of the circulatory system	1.18	1.02–1.37	0.031
Other diseases	2.09	1.62–2.70	<0.001
Year (ref: 2016)			
2017	0.98	0.90–1.06	0.562
2018	0.90	0.83–0.98	0.018
2019	1.09	1.00–1.19	0.040
Random effects			
σ^2	3.29		
τ_{00}	0.38 _{clinic}		
<i>N</i> groups	16 _{clinic}		
<i>N</i> observations	24 572		
Model indices			
ICC	0.10		
Marginal <i>R</i> ² /conditional <i>R</i> ²	0.103/0.196		
AIC	27 927.981		
BIC	28 090.169		
RMSE	0.438		

Notes: Italic *P* values are <0.05; ICC = intraclass correlation coefficient; AIC = Akaike information criterion; BIC = Bayesian information criterion; RMSE = root mean squared error.

the correctness of our results [9]. High non-response rates in specific clinics lead to small case numbers and could lead to smaller (and non-significant) differences. Additionally, the samples may not represent the true population since certain groups, such as people who are not L1 speakers, less affluent patients and more severely ill patients, are under-represented.

Results of PROMs and clinic comparisons based on these results must be interpreted carefully, with attention being paid to non-response rates and non-response bias.

Methodologically, these issues can be approached with multiple imputation techniques [13, 29]. However, the question is whether imputation on the outcome variable is desirable in a quality assurance context. In fact, complete data should be one aspect of good quality. The indicator would then be the response rate. When action is being taken to increase response rates of PROMs, subgroups with higher chance of non-response should be addressed specifically [7], as non-response bias can remain even with high response rates [9].

The importance of systematic integration of PROMs into organizational procedures and treatment processes should be emphasized in quality assurance programmes. Results should also be discussed with the patient, which can lead to shared decision-making, more adherence and a strengthened self-efficacy of the patients [2, 3, 30]. This may result in the medical team being more motivated to pay attention to completed questionnaires.

Completing a PROM questionnaire should not burden patients unduly. Offering shorter or proxy versions of the questionnaire for particularly ill patients [31] or questionnaires in different languages for non-native speakers could be an effective strategy for improving response rates for these groups. The easiest way to facilitate this is to offer online questionnaires. They provide instant feedback, and the results can directly be discussed with the patient [32].

Conclusions

We have found significant non-response bias among certain patient groups, as well as across different treatment facilities. Several patient characteristics were associated with non-response, and response rates differed considerably across clinics (provider effect). Measures to improve response rates among patients with known barriers to participation, as well as among different treatment facilities, need to be considered. For quality assurance programmes and outcome evaluations of medical treatments, non-response bias should be a serious concern, particularly when PROMs are being used; their results should be interpreted with caution. However, even though PROMs tend to have high non-response rates, they are necessary for evaluating treatment quality, as several important health outcomes, such as HRQOL, can only be measured from the patient's perspective.

Data availability statement

Data are available upon request from the Head of Rehabilitation at ANQ, Luise Menzi (luise.menzi@anq.ch).

Contributorship

All authors made substantial contributions to data collection, study conception and interpretation of results. Data analysis was performed by A.S., S.K. and M.M. The first draft was written by S.K., A.S., M.M. and A.S.-S. (Introduction), and the manuscript was revised critically for important intellectual content by A.S.-S., S.T., J.V. and L.M. Necessary revision to the first draft was taken over by S.K., A.S. and M.M. All authors read and approved the final version of the manuscript.

Acknowledgements

We would like to acknowledge and thank all other members of the quality committee at the ANQ: Dr sc med. Gavin Brupbacher, Annette Egger, MPH, Dr med. Stefan Goetz, Angelina Hofstetter, Barbara Lüscher, MHA, Dr med. Pierre-André Rapin, Dr med. Isabelle Rittmeyer, Dr med. Thomas Sigrist and Dr med. Christian Sturzenegger.

We would especially like to thank Anika Zembic for important editorial notes and proofreading of the final version of this article.

Funding

This work was funded by the ‘Swiss National Association for Quality Development in Hospitals and Clinics - ANQ’.

Ethics and other permissions

This study does not require the approval of an ethics committee since the data are part of the Swiss national quality assurance programme.

References

1. Rotenstein LS, Huckman RS, Wagle NW. Making patients and doctors happier—the potential of patient-reported outcomes. *N Engl J Med* 2017;377:1309–12.
2. Valderas J, Kotzeva A, Espallargues M *et al*. The impact of measuring patient-reported outcomes in clinical practice: a systematic review of the literature. *Qual Life Research* 2008;17:179–93.
3. Noonan VK, Lyddiatt A, Ware P *et al*. Montreal Accord on Patient-Reported Outcomes (PROs) use series—Paper 3: patient-reported outcomes can facilitate shared decision-making and guide self-management. *J Clin Epidemiol* 2017;89:125–35.
4. 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. 2017.
5. Black N. Patient reported outcome measures could help transform healthcare. *BMJ* 2013;346: f167.
6. WHO. *ICF: International Classification of Functioning, Disability and Health*. Genf: World Health Organization, 2001.
7. Coste J, Quinquis L, Audureau E *et al*. Non response, incomplete and inconsistent responses to self-administered health-related quality of life measures in the general population: patterns, determinants and impact on the validity of estimates—a population-based study in France using the MOS SF-36. *Health Qual Life Outcomes* 2013;11:1–15.
8. Schamber EM, Takemoto SK, Chenok KE *et al*. Barriers to completion of patient reported outcome measures. *J Arthroplasty* 2013;28:1449–53.
9. Halbesleben JR, Whitman MV. Evaluating survey quality in health services research: a decision framework for assessing nonresponse bias. *Health Serv Res* 2013;48:913–30.
10. Burrus C, Ballabeni P, Deriaz O *et al*. Predictors of nonresponse in a questionnaire-based outcome study of vocational rehabilitation patients. *Arch Phys Med Rehabil* 2009;90:1499–505.
11. Hutchings A, Neuburger J, Grosse Frie K *et al*. Factors associated with non-response in routine use of patient reported outcome measures after elective surgery in England. *Health Qual Life Outcomes* 2012;10:1–8.
12. Hutchings A, Neuburger J, van der Meulen J *et al*. Estimating recruitment rates for routine use of patient reported outcome measures and the impact on provider comparisons. *BMC Health Serv Res* 2014;14:1–8.
13. Gomes M, Gutacker N, Bojke C *et al*. Addressing missing data in patient-reported outcome measures (PROMs): implications for the use of PROMs for comparing provider performance. *Health Econ* 2016;25:515–28.
14. Höfer S, Lim L, Guyatt G *et al*. The MacNew Heart Disease health-related quality of life instrument: a summary. *Health Qual Life Outcomes* 2004;2:3.
15. Thompson DR, Yu C-M. Quality of life in patients with coronary heart disease-I: assessment tools. *Health Qual Life Outcomes* 2003;1:1–5.
16. Bundesamt für Statistik. In: *Variablen der Medizinischen Statistik, Spezifikationen*. Bern, 2019.
17. Linn BS, Linn MW, Gurel L. Cumulative illness rating scale. *J Am Geriatr Soc* 1968;16:622–6.
18. vonElm E, Altman DG, Egger M *et al*. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *PLoS Med* 2007;4:e297.
19. Hoffmann W, Latza U, Baumeister SE *et al*. Guidelines and recommendations for ensuring Good Epidemiological Practice (GEP): a guideline developed by the German Society for Epidemiology. *Eur J Epidemiol* 2019;34:301–17.
20. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2021.
21. Bates D, Mächler M, Bolker BM *et al*. Fitting linear mixed-effects models using lme4. arXiv:1406.5823, 2014.
22. Lüdtke D, Ben-Shachar MS, Patil I *et al*. Performance: an R package for assessment, comparison and testing of statistical models. *J Open Source Softw* 2021;6: 3139.
23. McFadden D. *Quantitative Methods for Analysing Travel Behaviour of Individuals: Some Recent Developments*. No 474 Cowles Foundation Discussion Papers. Cowles Foundation for Research in Economics, Yale University, 1977.
24. Ho A, Purdie C, Tirosh O *et al*. Improving the response rate of patient-reported outcome measures in an Australian tertiary metropolitan hospital. *Patient Relat Outcome Meas* 2019;10:217–26.
25. Ahlmark N, Algren MH, Holmberg T *et al*. Survey nonresponse among ethnic minorities in a national health survey—a mixed-method study of participation, barriers, and potentials. *Ethn Health* 2015;20:611–32.
26. Kwong E, Neuburger J, Petersen SE *et al*. Using patient-reported outcome measures for primary percutaneous coronary intervention. *Open Heart* 2019;6:e000920.
27. Kwon SK, Kang YG, Chang CB *et al*. Interpretations of the clinical outcomes of the nonresponders to mail surveys in patients after total knee arthroplasty. *J Arthroplasty* 2010;25:133–7.
28. Hutchings A, Frie KG, Neuburger J *et al*. Late response to patient-reported outcome questionnaires after surgery was associated with worse outcome. *J Clin Epidemiol* 2013;66:218–25.
29. Rombach I, Gray AM, Jenkinson C *et al*. Multiple imputation for patient reported outcome measures in randomised controlled trials: advantages and disadvantages of imputing at the item, subscale or composite score level. *BMC Med Res Methodol* 2018;18:1–16.
30. Rowland C, Walsh L, Harrop R *et al*. What do UK orthopedic surgery patients think about PROMs? Evaluating the evaluation and explaining missing data. *Qual Health Res* 2019;29:2057–69.
31. Nguyen H, Butow P, Dhillion H *et al*. A review of the barriers to using patient-reported outcomes (PROs) and patient-reported outcome measures (PROMs) in routine cancer care. *J Med Radiat Sci* 2021;68:186–95.
32. Richter JG, Becker A, Koch T *et al*. Self-assessments of patients via Tablet PC in routine patient care: comparison with standardised paper questionnaires. *Ann Rheum Dis* 2008;67:1739–41.