

Risk Adjustment Considerations – PROMs 2.0

Decision Support –

What will the output of the system be used for? - What decisions will be made based on information from PROMs 2.0?

What are the estimated consequences of this decision being wrong?

What are the factors that can impact on this, how they can be measured and to what certainty?

These questions have been looked at from the perspective of different stakeholders.

Commissioners (CCG's)

Clinical Commissioning Groups have the responsibility of shaping the healthcare delivered to their local population.

Commissioning decisions –

The Royal College of GP's define Clinically-led commissioning as a continual process of analysing the needs of a community, designing pathways of care, then specifying and procuring services that will deliver and improve agreed health and social outcomes, within the resources available.

Good commissioning places patients at the heart of the process. It is about improving peoples' lives and providing high quality services that are designed around the individual. Clinical commissioning groups, local authorities and others need to work together to plan and deliver better integration of local services to ensure that communities enjoy the highest quality responsive, affordable and personalised services that are also shaped directly by the people who use them.

The cycle of commissioning can be broken down into 4 stages –

Analysis and Plan

– investigates why a care pathway important.

Design Pathways

– relies on available data on quality, effectiveness, cost, need and use of services to inform the design of pathways.

Specify and Procure

– reviews what a good care pathway might look like based on the local and national evidence.

Deliver and improve

– looks at the services provided and how to improve safety, quality, effectiveness, patient experience and value for money, as well as performing to the contract and delivering national and local quality standards.

What will the output of the system be used for? - What decisions will be made based on information from PROMs 2.0?

PROMs 2.0 has a clear role in providing data on the outcomes of healthcare episodes and on patient experience. These data can inform during the design stage and again later during the delivery and improvement stage, where the commissioned service is evaluated.

Serial measurements within one organisation will be subject to variation over time due to fluctuations in the factors mentioned above. If raw data is used, caution must be exercised to avoid concluding that health outcomes are improving or worsening when these variations are due to spurious factors, unrelated to changes in the service quality. Risk adjustment is required in this scenario to mitigate these effects and to allow more precise interpretation of efforts towards service improvement. However, the more important pitfall in relation to commissioning is where a decision is made to procure services from one provider in place of another provider, based on raw outcome data.

For example, Provider A may achieve better improvements in raw outcome scores in their population compared to Provider B due to a younger, more motivated population with fewer co-morbidities, rather than delivering better quality care. Commissioning decisions need to be based on information that gives a fair comparison of different providers.

Real data are available to illustrate this. Hospital Episode Statistics online (HESonline™) provides public access to the National PROMs results (Patient Reported Outcome Measures) for NHS providers. This is based on data collected for varicose vein surgery, inguinal hernia repairs, hip and knee replacements. Unadjusted health gains and adjusted health gains are reported and different providers can be compared.

General health gain is judged by improvement in EQ5D index from pre-op to post op. Condition specific outcome scores are also reported. For Hip replacement, the improvement in Oxford Hip Score is used (OHS). Table 1 shows the results for two geographically adjacent NHS providers. From a commissioning stand point, decisions based on raw data alone may favour one trust whilst the adjusted data would not.

Provider	Raw general health gain (Change in EQ5D index)	Adjusted general health gain (Change in EQ5D index)	Raw improvement in OHS	Adjusted improvement in OHS
Northumbria Healthcare NHS Foundation Trust	0.404	0.430	20.13	20.63
The Newcastle Upon Tyne Hospitals NHS Foundation Trust	0.426	0.415	19.57	20.15

Table 1. Raw and adjusted health gains for patients undergoing hip replacements between April 2010 and March 2011 in two geographically adjacent catchments.

In this example, the health gains as judged by the difference in EQ5D index, in patients undergoing total hip replacements in The Newcastle Upon Tyne Hospitals NHS Foundation Trust appear to be greater than in the Northumbria Healthcare NHS Foundation Trust. However, when the results are adjusted, the reverse is true. In contrast, when health gain is judged by looking at the change in OHS for the same group of patients, the risk adjustment does alter the figures but each trust remains in the same position relative to each other.

Commissioning decisions based on raw data alone may not procure the best healthcare outcomes available. As discussed above, other data on effectiveness and quality should be considered.

What are the consequences of this decision being wrong?

The consequences of commissioning a healthcare provider delivering inferior quality healthcare is that the patients in that group will suffer worse outcomes overall. This is clearly a disadvantage for the patient but also has wider implications of possible delay in return to work, reduced physical activity levels and future costs of healthcare. The average difference using the example above would be 0.015 of the EQ5D index and 0.48 points of the OHS (1%). These are subtle differences. Large differences are less likely to be masked by population variation.

Commissioners often need to choose between providers within a finite geographical area. If the population characteristics within that area are fairly uniform then unadjusted data may reliably indicate better outcomes from providers of better quality care. However, where marked variation in population characteristics exists, risk adjustment becomes more important to allow a fair comparison. The population within the catchment of Northumbria Healthcare NHS Foundation Trust may have more comorbidity, less affluence, higher smoking rates etc, compared to the population that The Newcastle Upon Tyne Hospitals NHS Foundation Trust serves. Commissioning needs to be performed with knowledge of the local population characteristics and data adjusted to account for the variations.

What are the factors that can impact on this, how they can be measured and to what certainty?

Some of the factors influencing outcome can be measured with a high degree of certainty such as age and sex. However, other important aspects such as attitude to healthcare and willingness to engage with rehabilitation or lifestyle change are far more difficult to measure. The national PROMs project has found several variables to be measurable and predictive of post-operative outcome scores. The table 2 illustrates the factors that were found to be predictive of the post-op EQ5D index scores (Q2) for patients undergoing hip replacements.

Factor	Definition	Coefficient	Standard error
Q1 score	Patient's pre-operative questionnaire score	0.285	0.011
Q1 score (squared)	Patient's pre-operative questionnaire score, squared	-0.173	0.016
Age	Patient's age	0.006	0.0009
Age (squared)	Patient's age, squared	0.00004	0.000007
Sex: Female	Takes value 1 if patient is female, 0 if not	-0.017	0.002
Ethnicity: Asian	Takes value 1 if patient is of Asian ethnicity, 0 if not	-0.059	0.021
Ethnicity: Black	Takes value 1 if patient is of Black ethnicity, 0 if not	-0.056	0.017
Ethnicity: Not given	Takes value 1 if patient's ethnicity is not given through HES, 0 if not	0.023	0.004
IMD (Deprivation)	The IMD (Deprivation) 2004 score for the area the patient lives in	-0.001	0.0001
Assisted at Q1	Takes value 1 if patient was assisted in completing pre-operative questionnaire, 0 if not	0.013	0.003
Assisted at Q2	Takes value 1 if patient was assisted in completing post-operative questionnaire, 0 if not	-0.098	0.005
Disabled at Q1	Takes value 1 if patient considers themselves to have a disability, 0 if not	-0.066	0.002
HRG Code H80	Takes value 1 if the main HRG assigned to patient's spell was H80 (v3.5), 0 if not	0.056	0.025
HRG Code H81	Takes value 1 if the main HRG assigned to patient's spell was H81 (v3.5), 0 if not	0.065	0.025
Previous Surgery: Yes	Takes value 1 if patient has had previous surgery on their hip, 0 if not	-0.037	0.006
PRC: High blood pressure	Takes value 1 if patient has high blood pressure, 0 if not	0.017	0.004
PRC: Poor circulation	Takes value 1 if patient has poor circulation, 0 if not	-0.048	0.006
PRC: Diabetes	Takes value 1 if patient has diabetes, 0 if not	0.021	0.006
PRC: Nervous system diseases	Takes value 1 if patient has nervous system diseases, 0 if not	-0.050	0.013
PRC: Cancer	Takes value 1 if patient has (had) cancer, 0 if not	0.020	0.006
PRC: Depression	Takes value 1 if patient has depression, 0 if not	-0.103	0.006
PRC: Arthritis	Takes value 1 if patient has arthritis, 0 if not	-0.015	0.003
Patient has 1 HESRC	Takes value 1 if patient has one HES-reported comorbidity, 0 if not	-0.023	0.003
Patient has 2 HESRC	Takes value 1 if patient has two HES-reported comorbidities, 0 if not	-0.053	0.007
Patient has 3 HESRC	Takes value 1 if patient has three or more HES-reported comorbidities, 0 if not	-0.060	0.017
Patient has 2 PRCs	Takes value 1 if patient has self-reported exactly two comorbidities, 0 if not	-0.027	0.005
Patient has 3 PRCs	Takes value 1 if patient has self-reported exactly three comorbidities, 0 if not	-0.047	0.008
Patient has 4 PRCs	Takes value 1 if patient has self-reported four or more comorbidities, 0 if not	-0.080	0.014
Symptom period (1-5 yrs)	Takes value 1 if patient has experienced symptoms for between 1 and 5 years, 0 if not	-0.012	0.003

Symptom period (6-10 yrs)	Takes value 1 if patient has experienced symptoms for between 6 and 10 years, 0 if not	-0.028	0.005
Symptom period (10+ yrs)	Takes value 1 if patient has experienced symptoms for more than 10 years, 0 if not	-0.027	0.005
Constant term		0.562	0.040

PRC = Patient Reported Comorbidity

HESRC = HES Reported Comorbidity

Table 2. Variables that were statistically significant at the 95% level from the estimation step and their coefficients for predicting post-operative (Q2) scores for the Hip Replacement EQ-5D Index. The predicted score is found by multiplying each coefficient by the patient's response, summing the values and adding the constant term.

(Patient Reported Outcome Measures (PROMs) in England: The case-mix adjustment methodology Department of Health. 11 April 2012)

Within this model for this population, all the above factors had a statistically significant influence on the post op EQ5D index but the major drivers were the patients' pre operative EQ5D index, disability status, depression, assistance in completing questionnaire and deprivation score. However, even with multiple variables included in the models, their ability to predict the outcome (post operative score) is limited. The coefficient of determination R^2 for the PROM models in hip and knee replacement surgery for 2010 are shown in Table 3.

Model	No of variables	No of records	R^2
Hip Replacement EQ5D Index	18	12743	0.230
Hip Replacement EQ VAS	18	12084	0.265
Oxford Hip Score	16	14019	0.241
Knee Replacement EQ5D Index	17	14041	0.252
Knee Replacement EQ VAS	18	13303	0.299
Oxford Knee Score	20	15718	0.262

Table 3. R^2 values for orthopaedic risk adjustment models.

(PROMs risk adjustment methodology guide for general surgery and orthopaedic procedures. September 2010. A report to the Department of Health: NHS National PROMs Programme.)

These results show that, at best, these models account for less than one third of the variability in outcome. It may be that there are other variable which have not been measured which could improve the predictive value of the model. In practice, particularly in biologic systems, coefficients of determination of 30-50% are the best that can be produced¹. Despite this, the example above of the data form North East England shows that the models can adjust results to the point where commissioning decisions could be influenced.

Patients

One innovative feature of PROMs 2.0 is that patients have access to their own results and can compare their scores following treatment to their preoperative scores. They may use these tools that quantify their symptoms, function and wellbeing in addition to their own perceptions to gauge their progress. The additional feedback on their progress may reassure them if the score shows improvement. A lower score may be worrying and may prompt them to attend a clinic sooner or seek advice. As with all patient information it needs to be clearly communicated to avoid confusion plus contact details and sources of other information made available.

Within this context, risk adjustment is not a key issue as the pre and post-operative scores are for the same individual and so are matched. Factors influencing an outcome score independent of the quality of treatment given could include deterioration of other medical conditions. For example, a patient undergoing a total hip replacement may not see an improvement in their post op Oxford Hip Score (OHS) if they have worsening symptoms from their other hip or knee joints leading to pain and limitation of mobility. This may need to be elicited in consultation with a healthcare practitioner. Risk adjustment would not address this type of confounding factor for an individual.

Clinicians

What will the output of the system be used for? - What decisions will be made based on information from PROMs 2.0?

For treating clinicians, the system will be used in two distinct roles. The first of these will be in reference to an individual patient. The clinician may be reviewing the progress of a patient in their absence or they may use the PROMs 2.0 system as an adjunct to a clinic visit, having that patients outcome scores visible during the consultation. Decisions to start or change a treatment will not be made using outcome scores in isolation but seeing significantly improved scores would reassure the clinician that the current treatment is appropriate. An unanticipated deterioration in a score would highlight the need to carefully assess the patient to determine the cause.

Risk stratification in this scenario would not have utility as the patient acts as his or her matched control.

The second way that a clinician can use the PROMs 2.0 system is in reviewing and comparing groups of patients in a variety of ways. Comparing the outcomes of patients with similar diagnoses having different treatments. (eg patients with ankle arthritis having fusions compared to ankle replacements.) Looking at differences in outcomes over time to assess the impact of a change in practice or learning curve (comparison of first 50 arthroscopic rotator cuff repairs to the second 50 patients.) Alternatively the clinician may want to compare their own outcomes with a colleague within the same organisation or compared to regional or national figures. The results of

this type of audit may lead to changes in treatment choice, treatment technique or training.

What are the consequences of this decision being wrong?

An incorrect decision would expose a cohort of patients to an inferior treatment. Good clinical audit should include re-audit of results after a change has been implemented². Therefore, deterioration in outcomes ought to be identified unless further confounding factors are not recognised or corrected for.

This important second use of the system is dependent on risk adjustment to reduce the variation in outcome results caused by confounding factors such as disease severity, comorbidity, deprivation, age, etc.

As shown above, risk stratification cannot account for all of the variability in outcome between groups. However, clinicians are less likely to accept that outcome differences are due to disparity in the quality of care unless steps have been taken to account for confounding factors. The factors that can influence outcome and can be measured are detailed in the section above about Commissioning.

Managers

Hospital managers understand how their input use relates to the production of services and how that relates to cost. For example running extra theatre lists should allow more operations to be performed. However, the widespread collection of PROMs data is still new, managers probably don't understand how to administer processes to produce better outcomes. The use of PROMs-based performance indicators and the likelihood of revenue being linked to that performance will create strong pressure for managers to understand not just how their hospital's performance compares with others but also the causes of those variations. Revenue may depend on outcome measures either indirectly via patient choice and Payment by Results or directly via CQUIN payments (Commissioning for Quality and Innovation).³

What will the output of the system be used for? - What decisions will be made based on information from PROMs 2.0?

With increasing emphasis being placed on quality of care, encompassing patient safety, experience and effectiveness, managers will use PROMs 2.0 to provide objective data on outcome. Examples of good practice should be encouraged, expanded and emulated. Where 'outliers' are identified, more resources may be directed to that area to improve the effectiveness of the service. Alternatively, services may be reconfigured to concentrate on treatments offering good outcomes and withdraw services performing poorly.

What are the consequences of this decision being wrong?

What are the factors that can impact on this, how they can be measured and to what certainty?

Caution needs to be exercised before such decisions are made as reconfiguration of services has major financial and workforce implications. Managers need to understand that risk stratification can help to allow fair comparison between different clinicians, departments and trusts but variation in results will remain. As mentioned above the best prediction that can be expected for a risk adjustment model of a biologic system is 30-50%¹. Thus managers need to avoid doggedly following outcome results in a league-table form but should use the identification of an outlier as a prompt to engage with and understand the challenges of that service.³

Data Analysers / Central Team

PROMs 2.0 has a 'Central Team' to oversee system development, support implementation, monitor activity, provide information governance leadership, analyse data and publish reports.

What will the output of the system be used for? - What decisions will be made based on information from PROMs 2.0?

Clinical outcome reporting has been pioneered by the specialty of cardiothoracic surgery. Standardised mortality rates for a number of index procedures have been published for several years. Mortality rates have consistently improved, despite increases in complexity of cases and patient risk factors.⁴

PROMs 2.0 will report outcomes anonymised activity and outcome data. In line with the cardiothoracic reporting strategy, if outlier clinicians are identified, these clinicians will be directly contacted to draw their attention to the results.

***What are the consequences of this decision being wrong?
What are the factors that can impact on this, how they can be measured and to what certainty?***

The PROMs 2.0 Central Team will take no direct action or decisions. However, risk stratification needs to be used to allow a fairer comparison of providers to avoid alarming individuals wrongly. Reporting methods will include Funnel-plots. This widely recognised technique to graphically compare providers is to use a funnel-plot is also used by the National Joint Registry to report revision rates of joint replacements.⁵ (see figure 1).

As mentioned above, perfect risk stratification is not a realistic prospect. Identifying an outlier statistically is not cause for action to be taken but a prompt to understand what is causing the variation.

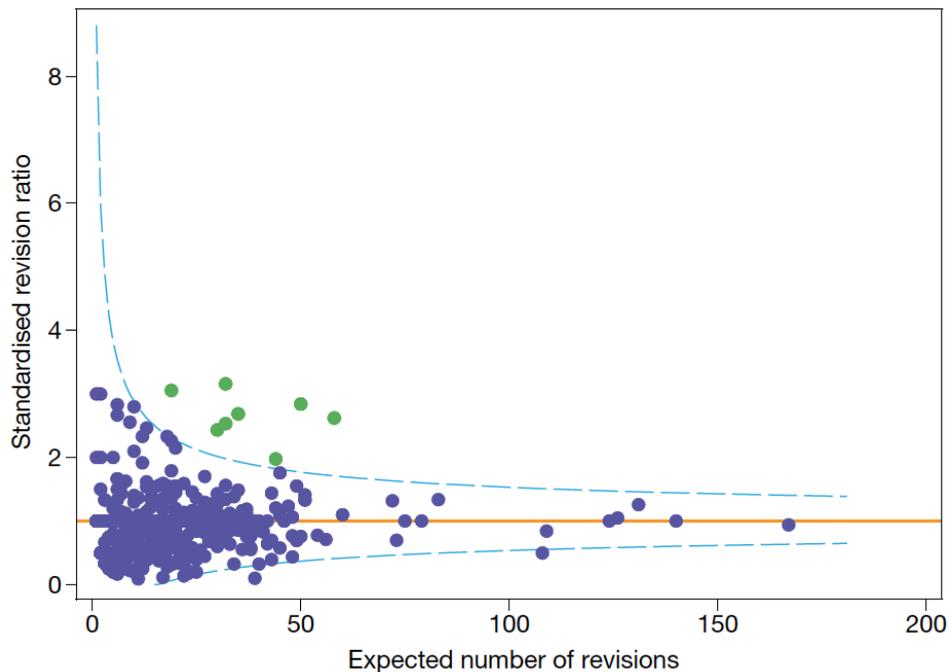


Figure 1. Funnel plot for revision rates in Total Hip Replacement. The confidence limits are set at four standard deviations. Unit results above the upper confidence limit reflect special-cause variation and are termed 'outliers'. These are indicated in green.

General Public

As with the cardiothoracic reporting, PROMs 2.0 will provide reports that are available to the general public. A systematic review of patient choice of provider found that patients do count quality of care very highly when choosing a hospital. Other highly rated factors were ease of access, cleanliness and a low rate of infection.⁶

League-table reporting should be avoided, as the limitations of any risk stratification model may be difficult to communicate.

GMC - Revalidation

¹ Risk Stratification. A practical guide for clinicians. Miller CC, Reardon MJ, Saif HJ. Cambridge. 2001. P27.

² Principles for best practice in clinical audit. National Institute For Clinical Excellence. Radcliffe Medical Press. 2002. P3.

³ Getting the most out of proms. Putting health outcomes at the heart of NHS decision-making. Devlin NJ, Appleby J. The Kings Fund. 2010

⁴ National Adult Cardiac Surgery Audit Report. Annual Report 2010-2011. National Institute for Cardiovascular Outcomes Research.

⁵ National Joint Registry for England and Wales. 9th Annual Report 2012. P176

⁶ Systematic review of the impact of patient choice of provider in the English NHS. Jones L, Mays N. Health Services Research Unit Department of Public Health and Policy London School of Hygiene and Tropical Medicine. 2009. P13.