



# Patient Reported Outcome Measures (PROMs) in England

*A Methodology for Identifying Potential Outliers*

**DH INFORMATION READER BOX**

Policy	Estates
HR / Workforce	<b>Commissioning</b>
Management	IM & T
Planning /	Finance
Performance	Social Care / Partnership Working

**Document Purpose** Consultation/Discussion

**Gateway Reference** 16340

**Title** Patient Reported Outcome Measures (PROMs) in England: A Methodology for Identifying Outliers

**Author** Department of Health

**Publication Date** April 2012

**Target Audience** PCT CEs, NHS Trust CEs, Medical Directors, Directors of Nursing

**Circulation List**

**Description** This document outlines the Department of Health's methodology for identifying potential outliers in the national PROMs programme.

**Cross Ref** N/A

**Superseded Docs** N/A

**Action Required** Comments are welcome via the PROMs mailbox. Please state "Outlier Methodology" in the subject line.

**Timing**

**Contact Details** PROMs Team  
Commissioning Development  
2<sup>nd</sup> Floor, Richmond House  
79 Whitehall  
London  
SW1A 2NS  
  
0207 210 3000

**For Recipient's Use**

# PROMs in England

## *A Methodology for Identifying Potential Outliers*

**Prepared by the PROMs team, Commissioning Development Directorate.**

© Crown copyright 2011

First published 2011

Published to DH website, in electronic PDF format only.

<http://www.dh.gov.uk/publications>

# Contents

Contents.....	4
Introduction .....	5
Summary of Key Points.....	6
Background.....	7
Identifying Potential Outliers.....	8
Further Investigation and Verification.....	16
Annex A: Draft Outlier Policy.....	17
Annex B: Example Report on Potential Outliers.....	23

# Introduction

This paper outlines the Department of Health's methodology for identifying potential outliers in the national Patient Reported Outcome Measures (PROMs) programme.

One of the stated aims of the PROMs programme is to support evaluations by clinicians, managers, regulators, commissioners and patients of the relative clinical quality and performance of providers of relevant elective procedures<sup>1</sup>.

To date, there has been no single agreed methodology for identifying organisations that are performing relatively better or worse than others using PROMs data. Various statistical methods have been applied to the data to assess relative performance<sup>2</sup> but there remains no single convention for this sort of analysis on PROMs data. This may be due to the novel nature of the data. Advanced statistical analyses are routinely applied to data with longer time series such as mortality data<sup>3 4</sup>. In contrast, there are limited examples of routine collection of PROMs data across all providers in specific clinical areas.

A starting point for developing these proposals is the guidance prepared by the National Clinical Audit Advisory Group (NCAAG) for the Department of Health<sup>5</sup>. While the PROMs programme is not a formal clinical audit as envisaged in the guidance document – data are published continuously on a monthly basis for continuous monitoring – there are common features. Specifically, the intention to produce comparisons of providers using batches of data for defined periods remains the same in both cases.

This paper sets out the main elements of our outlier policy for PROMs. A draft policy document can be found in annex A. An example report on potential outliers, based on the PROMs data series for April 2009 to December 2010 (published in May 2011), can be found at annex B.

---

<sup>1</sup> Department of Health. Guidance on the routine collection of Patient Reported Outcome Measures (PROMs) For the NHS in England 2009/10. Department of Health. December 2008.  
[http://www.dh.gov.uk/prod\\_consum\\_dh/groups/dh\\_digitalassets/@dh/@en/documents/digitalasset/dh\\_092625.pdf](http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_092625.pdf)

<sup>2</sup> Browne et al. Patient Reported Outcome Measures (PROMs) in elective surgery: Report to the Department of Health. London School of Hygiene and Tropical Medicine. December 2007.  
[http://www.lshtm.ac.uk/php/hsrp/research/proms\\_report\\_12\\_dec\\_07.pdf](http://www.lshtm.ac.uk/php/hsrp/research/proms_report_12_dec_07.pdf)

<sup>3</sup> See for example the CQC's mortality outlier programme:  
<http://www.cqc.org.uk/aboutcqc/whatwedo/respondingtoconcerns/mortalityoutliers.cfm>

<sup>4</sup> Mohammed MA, Cheng KK, Rouse A and Marshall T. "Bristol, Shipman and clinical governance: Shewart's forgotten lessons", *Lancet* 2001; 357: 463-67.

<sup>5</sup> HQIP and Department of Health. Detection and management of outliers. Department of Health. January 2011.  
[http://www.dh.gov.uk/prod\\_consum\\_dh/groups/dh\\_digitalassets/@dh/@ab/documents/digitalasset/dh\\_123888.pdf](http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@ab/documents/digitalasset/dh_123888.pdf)

## Summary of Key Points

1. Both the EQ-5D index and the condition-specific PROMs (where applicable) will be used to identify potential outliers.
2. We will use the case-mix adjusted average health gain as the primary metric of interest to identify potential outliers.
3. We will use the national mean as the benchmark to identify potential outliers.
4. We will use the funnel plot as the method to identify potential outliers.
5. We will adopt the convention of identifying “alarms” using 99.8% control limits and “alerts” using 95% control limits within the PROMs outlier policy.
6. We will apply symmetric control limits, however in cases where the volume is less than 150 some judgement must be applied if the provider is close to the funnel limit. We propose that the benefit of the doubt is given in favour of providers being “in control” in these cases and that judgements should be applied as consistently as possible.
7. There are currently no grounds for adjusting PROMs data for over-dispersion. However, the data should be tested for over-dispersion periodically. If over-dispersion is detected then it should be corrected for.
8. The publication of a list of potential outliers will be published as part of or alongside the quarterly PROMs publication. It would be the responsibility of the provider to take action to explore and improve their performance.
9. We recommend that:
  - the IC’s participation and response rates table be used by providers to assess the quality of their data. Where rates are low, providers would be expected to take action to improve them,
  - Providers consider if there are other factors which may explain their presented results, other than variation in performance,
  - Where possible, comparative information be provided to help organisations identified as potential outliers for example, how they compare with other providers on pre-operative scores or on patient characteristics. This comparative information would be developed with input from providers.

# Background

1. The routine collection of PROMs since April 2009 in England has produced a rich dataset of outcomes data for four elective interventions. The PROMs dataset has the potential to be used to identify the poorer and better performing providers and hence to facilitate improvements in quality.
2. The data series for PROMs was first published by the Health and Social Care Information Centre (IC) in September 2010 with subsequent updates every month<sup>6</sup>. From August 2011, this became a quarterly publication with data attributed to quarters and financial years rather than being cumulatively updated for the entire data collection period. The IC removed the experimental data tag 6 months later in February 2012, in line with guidance from the UK Statistical Authority.
3. Prior to August 2011, statistical outliers were not reported in the PROMs publication. However, as the results are published at provider level, with confidence limits, it was possible to draw comparisons between providers from the published data or using the IC's comparison tool<sup>7</sup>. As the identification of specific organisations as outliers depends on the methodology employed, it is important that a single standard for identifying potential outliers is agreed to avoid mixed messages for organisations being compared.

---

<sup>6</sup> The Health and Social Care Information Centre. 2010. Provisional Monthly Patient Reported Outcome Measures (PROMs) in England: April 2009 – April 2010: Pre- and post-operative data: Experimental Statistics. The Health and Social Care Information Centre.  
<http://www.hesonline.nhs.uk/Ease/servlet/ContentServer?siteID=1937&categoryID=1295>

<sup>7</sup> The IC's comparison tool is available via HES online, under monthly documents – score comparison:  
<http://www.hesonline.org.uk/Ease/ContentServer?siteID=1937&categoryID=1295>

# Identifying Potential Outliers

4. This section discusses the statistical method to be employed to identify potential outliers based on the national PROMs data.

## What is an outlier?

5. When comparing institutional performance, an outlier is defined as an institution whose performance is statistically significantly different from some pre-determined benchmark, often the national mean<sup>8</sup>, and unlikely to be due to random variation.
6. Following this definition, a potential outlying organisation on PROMs data would be one for which performance is statistically different to an agreed benchmark for performance, however defined. Positive outliers will be those organisations whose performance is statistically different and better than the benchmark. For negative outliers, vice versa.
7. Organisations will be classified as potential outliers based on a statistical analysis of the available data. Organisations will be considered genuine outliers if or when further investigation fails to identify any plausible justification for the statistical results, other than variations in performance.

## Data sources

### Choice of health status measure

8. The national PROMs programme covers four clinical procedures: Hip replacement, knee replacement, varicose vein surgery and groin hernia surgery. For each of these procedures, there are two generic PROM measures of health status; the EQ-5D index score which is derived from responses to the five EQ-5D profile questions and the EQ-VAS (Visual Analogue Scale) which provides a single summary measure of health status directly. For hip replacement, knee replacement and varicose vein surgery there are also condition specific PROMs: the Oxford Hip Score (OHS), Oxford Knee Score (OKS) and Aberdeen Varicose Vein Questionnaire (AVVQ) respectively.
9. For those conditions that have both a generic and a condition specific PROM, there is merit in using both measures to identify potential outliers because they may be assessing different aspects of health status. The condition specific measures have been designed to identify differences in

---

<sup>8</sup> National averages are commonly used when analysing mortality rates. See for example Tekis P.P., McCulloch, P., Steger, A.C., Benjamin, I.A. and Polaniecki, J. D. (2003) "Mortality Control Charts for Comparing Performance of Surgical Units: Validation Study using Hospital Mortality Data." British Medical Journal 326: 786 – 788A.



## PROMs in England: A Methodology for Identifying Potential Outliers

health status associated specifically with the relevant clinical area and will therefore be more discriminatory than generic measures when assessing outcomes for individual clinical procedures<sup>9</sup>. It is expected that condition specific PROMs will result in a greater variation in reported outcomes than generic measures and hence the identification of a greater number of potential outliers. The EQ-5D index provides a much broader evaluation of health status incorporating assessments of factors other than those specific to the condition, disease or procedure and therefore the magnitude of change may be less pronounced than on a condition-specific measure.

10. For the purpose of identifying potential outliers, both the EQ-5D index and the condition-specific PROMs will be used. Providers will be identified as an outlier via either one or both measures. We do not propose to use the EQ-VAS for the detection of outliers at this time because of lack of sensitivity to changes in health status.

### Summary Box 1

Both the EQ-5D index and the condition-specific PROMs (where applicable) will be used to identify potential outliers.

### Choice of metric

11. The IC publishes an array of health status measures as part of its monthly publication cycle for PROMs data. These include data on average pre-operative health score, average post-operative health score and average health gain at provider level. In order to assess performance on the basis of outcomes using the published data it is necessary to examine either the average health gain (mean change in pre- to post-operative health status), or the average post-operative scores. Further, when undertaking institutional comparisons it is important to adjust for differences in case-mix, in order to compare like with like. The published PROMs data contains case-mix adjusted values for post-operative score and health gain<sup>10</sup>.
12. The choice of whether to use the average health gain or average post-operative scores is largely a presentational one. The nature of the case-mix adjustment methodology is such that the measures differ by the subtraction of a constant<sup>11</sup>. We will use the case-mix adjusted average health gain as the primary measure of interest although it is equally valid to use case-mix adjusted mean post-operative scores and the conclusions drawn will be identical under both approaches.

---

<sup>9</sup> See Browne et al (2007) for a comparison of effect sizes for generic and condition-specific measures.

<sup>10</sup> "PROMs risk adjustment methodology guide for general surgery and orthopaedic procedures" Northgate Information Solutions, September 2010.  
[http://www.northgate-proms.co.uk/docs/PROMS\\_risk\\_adjustment\\_methodologies\\_SEPT\\_10.pdf](http://www.northgate-proms.co.uk/docs/PROMS_risk_adjustment_methodologies_SEPT_10.pdf)

<sup>11</sup> See the technical appendix to the draft outlier policy in Annex A

## PROMs in England: A Methodology for Identifying Potential Outliers

13. Other measures are available but currently unpublished. Specifically, a measure can be calculated as the percentage (proportion) of patients achieving a health gain of or above some specific threshold (such as a Clinically / Minimally Important Difference or MID). These measures are not currently published because further work is required to establish an appropriate threshold or benchmark for the MID. We propose that these measures are further developed before being considered for adoption within the PROMs outlier policy.

### Summary Box 2

We will use the case-mix adjusted average health gain as the primary metric of interest to identify potential outliers.

## Benchmark

14. A benchmark can be based on either external criteria, imposed from without, or internal criteria, calculated from within the dataset. A benchmark based on research or the judgement of experts that establishes a minimum level of performance is an example of an external benchmark. An internal benchmark is one derived from the data itself, for example the statistical mean of the data series over all records among all organisations.
15. The systematic collection of PROMs in England is a world first in terms of comprehensiveness and scale. Consequently, little research has been conducted into what PROMs score constitutes acceptable performance and so there are no examples of any suitable external benchmarks to be drawn upon readily, although these are likely to develop over time. The use of an internal benchmark is therefore necessary, and the arithmetic mean of all records across all organisations in the dataset is the most appropriate.

### Summary Box 3

We will use the national mean as the benchmark for identifying potential outliers.

## Methodology

### Funnel Plots

16. The Funnel Plot is an increasingly used graphical and analytical tool for institutional comparisons and the identification of outliers. It is based on traditional Shewart control charts, which typically define a range of values that fall within limits around a benchmark as being “in control”<sup>12</sup>. Those providers which are defined as “in control” are assumed to be subject to

---

<sup>12</sup> Spiegelhalter, D. J. (2005) “Funnel plots for comparing institutional performance” *Statist. Med.* 2005: **24**: 1185-1202

## PROMs in England: A Methodology for Identifying Potential Outliers

common-cause variability, whereas those that are “out of control”, i.e. fall outside of the control limits, are subject to some special cause variability.

17. Funnel plots differ from traditional control charts in that they set the control limits around the benchmark according to some measure of precision, which in the case of PROMs will be volume of procedures undertaken or volume of completed and returned PROMs questionnaires.
18. Evidence has shown that there is support among clinicians for using funnel plots to present institutional comparisons<sup>13</sup>. They are relatively easy to produce, readily interpretable and allow for additional variability in institutions with small volumes.
19. An alternative method for assessing performance against a benchmark is to use a “caterpillar” plot. Whereas funnel plots use the variation in the patient (total sample) data to derive the limits around the benchmark, caterpillar plots use confidence limits calculated for each individual organisation using the variation in the organisation’s own data. This approach is already available using the confidence limits presented in the data published by the IC or using the published spreadsheet comparison tool. A significant criticism of this approach is that the method of presenting all organisations’ data simultaneously (the “caterpillar” plot) risks the spurious ranking of institutions<sup>14 15</sup> which is avoided in the funnel plot method. Further, the funnel has the advantage of displaying the volume-outcome relationship graphically and demonstrating the increased expected variability in lower volume samples (smaller hospitals or lower rates of PROMs returns).

### Summary Box 4

We will use the funnel plot as the method for identifying potential outliers.

### Setting Control Limits

20. Commonly, a threshold of three standard deviations from the target is used to identify values that are deemed to be “out of control”<sup>16</sup>. Statistically this corresponds to testing if an indicator is different from the target at a two-sided significance level of 0.002, i.e. a control limit of 99.8%. A threshold of two standard deviations from the target is equivalent to a two-sided significance level of 0.05 and sets control limits at 95%.

---

<sup>13</sup> Allwood, D., Hildon, Z., Black, N. (2011) “Clinicians’ views of the format and context of statistical comparisons of health care performance”, Department of Health Services Research & Policy, London School of Hygiene and Tropical Medicine, unpublished.

<sup>14</sup> Spiegelhalter, D. J. (2005) “Funnel plots for comparing institutional performance” *Statist. Med.* 2005: **24**: 1185-1202

<sup>15</sup> Mohammed A. M., Deeks, J.J. (2008) “In the Context of Performance Monitoring, the Caterpillar Plot should be Mothballed in Favour of the Funnel Plot”, *Letter to the Editor, The Annals of Thoracic Surgery* 2008; **86**: 348.

<sup>16</sup> *ibid*

## PROMs in England: A Methodology for Identifying Potential Outliers

21. Use of the control limits outlined above presents some risk of Type 1 error - the risk of identifying providers as potential outliers by chance alone - i.e. a false positive. With 99.8% control limits, the chance of incorrectly identifying a negative outlier is approximately one in a thousand. With 95% control limits this risk increases to approximately 25 in a thousand.
22. Guidance on the detection and management of outliers, produced by the National Clinical Audit Advisory Group (NCAAG) for the Department of Health, recommends using 95% and 99.8% control limits as thresholds for an “alert” and an “alarm” respectively, which is consistent with the literature in this area<sup>17</sup>. We therefore propose applying this convention to the PROMs outlier policy and have applied it in our preliminary analysis (see Annex B).

### Summary Box 5

We will adopt the convention of identifying “alarms” using 99.8% control limits and “alerts” using 95% control limits within the PROMs outlier policy.

### Symmetry of Control Limits

23. The use of symmetric control limits implicitly assumes that the random variation in the provider-level data is normally distributed. However, if patient-level data is skewed this assumption may not hold. It is possible that skewness is present in PROMs data, due in part to a ‘ceiling effect’ in the scales used and also due to genuine effectiveness of treatment. This may result in patients’ PROMs scores being concentrated around the higher end of the scale.
24. In the case where the underlying data are skewed and symmetric control limits are applied, the risk of Type 1 errors (false positives) is increased. Practically, this may result in an increase in the number of providers being identified as negative outliers and a decrease in the number of providers being identified as positive outliers. However, the central limit theorem suggests that as sample size increases, distribution tends to normality. Research by Neuberger et al (2011) for the Department of Health<sup>18</sup> suggests that once volumes reach approximately 150 the problems associated with skewness diminish.
25. Taking into account the issue of skewness, Neuberger et al conclude that if there is some flexibility over how marginal cases of poor performance are handled, symmetric control limits could reasonably be used in an analysis of PROMs outliers. However, if simple classifications of performance are put in place, the calculation of asymmetric simulated control limits should be considered. Their research also finds that applying a correction formula for skewness tended to over-correct and therefore is not recommended.

---

<sup>17</sup> ibid

<sup>18</sup> Neuberger J et al. Funnel plots for comparing provider performance based on patient-reported outcome measures (PROMs). *BMJ Quality and Safety*. <http://qualitysafety.bmj.com/content/early/recent>

**Summary Box 6**

We will apply symmetric control limits, however in cases where the volume is less than 150 some judgement must be applied if the provider is close to the funnel limit. We propose that the benefit of the doubt is given in favour of providers being “in control” in these cases and that judgements should be applied as consistently as possible.

Over-dispersion

26. When comparing providers using the funnel plot method, there is a risk that the results will show a greater number of potential outliers than would be expected due to chance alone plus the presence of genuine outliers. This is known as over-dispersion. Over-dispersion typically arises when there is insufficient risk adjustment and it is more likely to occur when an indicator is based on large numbers of observations<sup>19</sup>.

27. It is possible to estimate the over-dispersion factor, O, as follows<sup>20</sup>:

$$O = 1/l * (\sum z_i^2)$$

Where l is the number of observations and z<sub>i</sub> is unadjusted z score for each observation.

28. Over-dispersion may only be assumed if O is significantly greater than 1. The table below contains the calculated over-dispersion factors (O) for the PROMs data for the period April 2009 to December 2010. It shows that over-dispersion is not present.

Intervention	Measure	Over-dispersion factor
Knee replacement	EQ-5D	0.014
	OKS	0.031
Hip replacement	EQ-5D	0.031
	OHS	0.051
Varicose vein surgery	EQ-5D	0.012
	AVVQ	0.017
Groin hernia surgery	ED-5D	0.015

<sup>19</sup> Spiegelhalter D.J. Handling over-dispersion of performance indicators. Qual Saf Health Care 2005;14: 347-51.

<sup>20</sup> *ibid*

### Summary Box 7

There are currently no grounds for adjusting PROMs data for over-dispersion. However, the data should be tested for over-dispersion periodically. If over-dispersion is detected then it should be corrected for.

### Publication

29. As part of, or alongside, the quarterly PROMs publication, we propose to publish an outlier report which would include the funnel plots and a list of potential outliers<sup>21</sup>. As now, providers would be alerted to the publication via the Department of Health's bulletin to Chief Executives, "The Week".
30. Providers that have been identified as potential outliers should first ensure that they are satisfied that the data on which the analysis is based is correct. This could include verifying the data with the IC and referring to the data quality note which will be published as part of the PROMs publication.
31. The NCAAG report to the Department of Health on the detection and management of outliers<sup>22</sup> includes recommendations on the protocols following the identification of a potential outlier. This approach is aligned with national clinical audit procedures and will be relevant to many cases of performance monitoring. In the case of PROMs, the nature of outlying performance is different. Relatively 'poor' performance as measured by PROMs indicates that patients having surgery have lower than average health gain; it does not mean that patient safety is necessarily at risk. Therefore, we do not propose having prescriptive processes in place for dealing with identified outliers.

### Summary Box 8

The publication of a list of potential outliers will be published as part of or alongside the quarterly PROMs publication. It would be the responsibility of the provider to take action to explore and improve their performance.

---

<sup>21</sup> An example of the report we propose is illustrated in annex B.

<sup>22</sup> HQIP and Department of Health. Detection and management of outliers. Department of Health. January 2011.

[http://www.dh.gov.uk/prod\\_consum\\_dh/groups/dh\\_digitalassets/@dh/@ab/documents/digitalasset/dh\\_123888.pdf](http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@ab/documents/digitalasset/dh_123888.pdf)

### Summary

1. A potential outlier organisation is one which is significantly different to the national mean case-mix adjusted health gain at either the 95% confidence level (an “alert”) or the 99.8% confidence level (an “alarm”).
2. Both the EQ-5D index and condition specific PROMs (where applicable) will be used to identify potential outliers.
3. Funnel plots are recommended as the preferred methodology for identifying potential outliers.
4. Organisations will be identified as potential outliers on the basis of statistical analysis alone. Further investigation will be required to confirm if providers are genuine outliers.

## Further Investigation and Verification

32. Once an organisation has been identified as a potential outlier, further investigation is required to determine whether this is genuinely due to variation in performance. The first step is to verify the data in order to establish if the observed results are due to a data issue.

### Data Quality

33. The IC currently publishes a data quality note with the PROMs data that considers coverage, missing values, ambiguous values, demographics and time differences and highlights any known issues. As the outlier policy only uses published data, there are no additional data quality issues specific to the outlier policy. Therefore the IC's data quality note is sufficient to cover any data quality issues affecting the outlier analysis.

34. Providers would be expected to review their data and take steps to improve participation and response rates wherever possible.

### Generalisability

35. The generalisability or representativeness of data is assessed by comparing the characteristics of cases included with those that are not, or with all eligible cases. As the PROMs data is case-mix adjusted it is assumed that the results are generalisable.

#### Summary Box 9

We recommend that:

- the IC's participation and response rates table be used by providers to assess the quality of their data. Where rates are low, providers would be expected to take action to improve them,
- Providers consider if there are other factors which may explain their presented results,
- Where possible, comparative information be provided to help organisations identified as potential outliers, for example how they compare with other providers on pre-operative scores or on patient characteristics. This comparative information would be developed with input from providers.



# Annex A: Draft Outlier Policy

## Introduction

This document sets out the Department of Health's policy for identifying potential outliers in the Patient Reported Measures (PROMs) national programme. This covers four surgical procedures: hip replacements, knee replacements, groin hernia surgery and varicose veins surgery.

PROMs data captures the health status of patients at a point in time through the completion of questionnaires by patients. Patients are asked to fill in a PROMs questionnaire before and then after a procedure. This can then be used to impute changes in health status.

It is possible to use PROMs data to make inter-provider performance comparisons. Section 1 describes the approved methodology by which valid comparisons can be made. Section 2 gives details of further investigations and verification analysis that should be done to ensure the processes are robust.

## Section 1: Methodology for detecting potential outliers

### Data

1. The national PROMs programme collects a generic measure of health status for all four procedures and a condition specific measure for three of the procedures (hip replacement, knee replacement, and varicose veins).
2. The PROMs scores provide an indicator of post-operative health status or health gain per patient, which is used to generate an average post-operative score or health gain for each provider.
3. In order to ensure that like for like comparisons of the average scores between providers are valid, a case-mix adjustment is applied the data. This adjusts for patient characteristics, e.g. age and gender, co-morbidities and pre-operative health status.<sup>23</sup>
4. Once a case-mix adjustment has been applied to the data the choice between health gain and post operative health status is largely presentational, as the data series are linear transformations of each other.

---

<sup>23</sup> "PROMs risk adjustment methodology guide for general surgery and orthopaedic procedures"  
Northgate Information Solutions, September 2010.  
[http://www.northgate-proms.co.uk/docs/PROMS\\_risk\\_adjustment\\_methodologies\\_SEPT\\_10.pdf](http://www.northgate-proms.co.uk/docs/PROMS_risk_adjustment_methodologies_SEPT_10.pdf)

## PROMs in England: A Methodology for Identifying Potential Outliers

5. The benchmark used for comparisons between provider organisations is the mean case-mix adjusted average health gain across all providers in England.

### Identifying Potential Outliers

6. The tool that we use to compare performance between providers is the funnel plot. This is a graphical tool which plots volume of procedures on the horizontal axis, against case-mix adjusted average health gain on the vertical axis. Control limits are shown on the chart by superimposing lines that form a curved “funnel” about the benchmark.
7. The funnel plot has two control limits either side of the benchmark: a 99.8% control limit and a 95% control limit<sup>24</sup>. The convention in this area is that these control limits correspond to “alarm” and “alert” categories respectively.

Category	Description
Alarm	Provider lies outside the 99.8% control limit.
Alert	Provider lies outside the 95% control limit but within the 99.8% control limit
In control	Provider lies within all control limits.

8. The methodology employs symmetrical control limits because distribution tends to normality as sample size increases (over 150 records). For provider organisations with volumes of less than 150, the use of symmetrical control limits may result in over-identification of negative outliers.
9. Provider organisations with less than 30 records are excluded from the comparisons.
10. A report of potential outlier organisations will be published four times per year alongside the quarterly PROMs data publication. Provider organisations will be notified of publication through “The Week”, the Department of Health’s bulletin to Chief Executives.
11. Providers that have been identified as potential outliers should take steps to verify the data on which the analysis is based, in order to establish if their outlying performance is genuine or due to some issue with the data. Where no data issues are present and no other reason for the outlier status can be identified, a provider is then considered to be a genuine outlier.

---

<sup>24</sup> Information on how the control limits are calculated can be found in the technical appendix.

## Section 2: Data Quality and Verification

### Data Quality

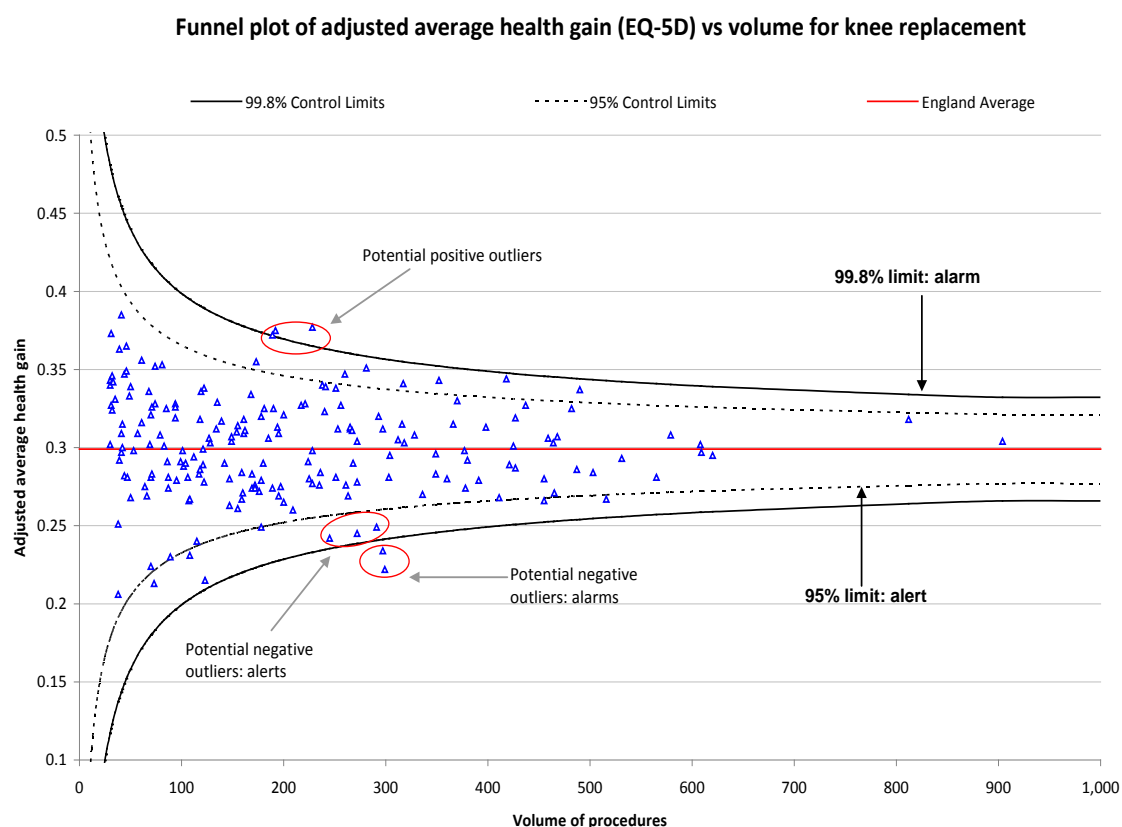
12. A Data Quality Note accompanies the publication of the PROMs data. This considers coverage, missing values, ambiguous values, demographics and time differences. It also highlights any known issues with the data. There are no data quality issues specific to the outlier policy.
13. It is the responsibility of the provider organisations to satisfy themselves that the data they provide is accurate. We recommend that providers use the IC's participation and response rates table to assess the quality of their data. Where rates are low, providers would be expected to take action to improve them. Providers should also consider if there are other factors which may confound their presented results.
14. Provider organisations have the right to request confirmation of the validity of the results in the report and of the data quality. Where possible, comparative information should be provided to help organisations identified as potential outliers, for example how they compare with other providers on pre-operative scores or on patient characteristics. This comparative information would be developed with input from providers.

## Technical appendix to the Outlier Policy

### Identifying Potential Outliers

First, a scatter plot of provider volume versus adjusted average health gain is produced. Funnel plots are constructed by plotting the upper and lower control limits at 99.8% and 95%, using all data available. The funnel limits are then super-imposed onto the scatter plot.

For all EQ-5D funnel plots and for those which plot Oxford Hip Score (OHS) and Oxford Knee Score (OKS)<sup>25</sup>, providers who lie above the upper control limits are considered positive outliers and those who lie below the lower control limits are considered negative outliers. Negative outliers can be further sub-divided into alarms and alerts; those lying outside the 99.8% control limits are considered alarms, whereas those lying outside the 95% control limit but within the 99.8% limits are considered alerts.

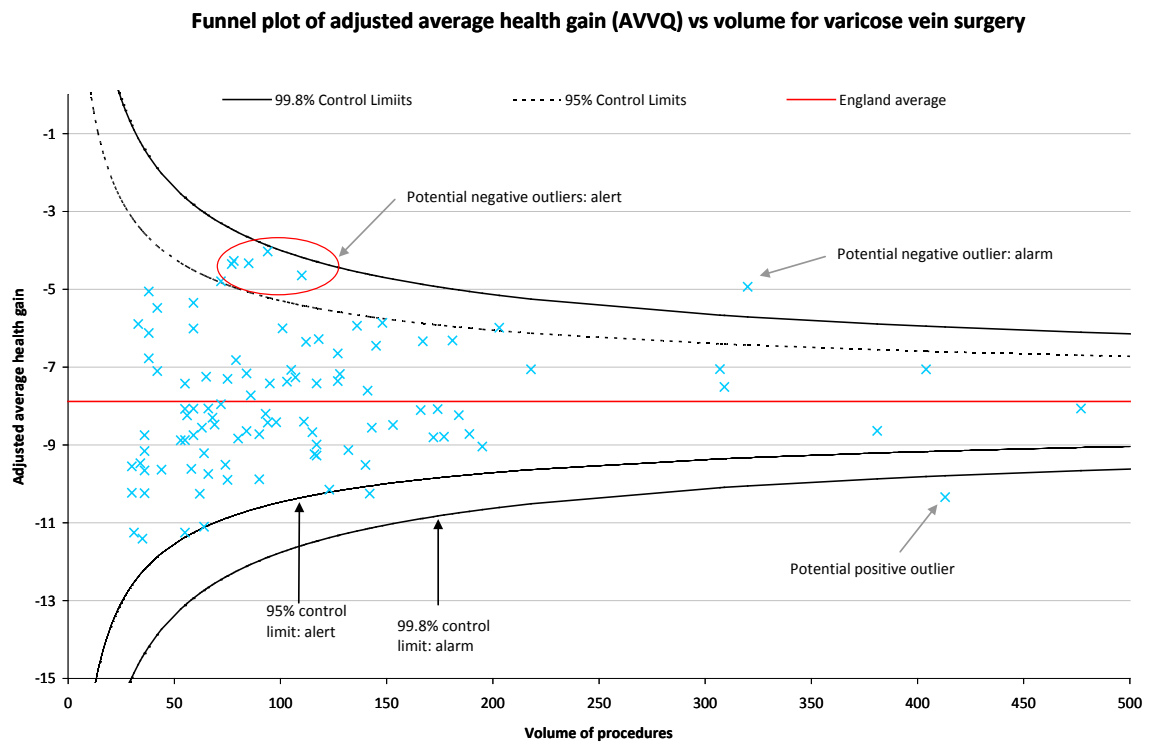


The Aberdeen Varicose Vein Questionnaire (AVVQ) condition-specific instrument uses a scale from 0-100, where 0 represents no evidence of varicose veins and 100 represents the most severe problems associated with

<sup>25</sup> For further information on the instruments and their scoring methodologies see “Provisional Monthly Patient Reported Outcome Measures (PROMs) in England: A guide to PROMs methodology. NHS Information Centre” available at:  
<http://www.hesonline.nhs.uk/Ease/servlet/ContentServer?siteID=1937&categoryID=1295>

## PROMs in England: A Methodology for Identifying Potential Outliers

varicose veins<sup>26</sup>. Therefore positive outliers are identified as those which lie below the lower control limits, whereas negative outliers are identified as those which lie above the upper control limits.



### Calculating the control limits

The standard formulae for the calculation of control limits are as follows:

$$99.8\% \text{ upper control limit } UL = \bar{X} + 3(s[X]\sqrt{(N)})$$

$$99.8\% \text{ lower control limit } LL = \bar{X} - 3(s[X]\sqrt{(N)})$$

$$95\% \text{ upper control limit } UL = \bar{X} + 2(s[X]\sqrt{(N)})$$

$$95\% \text{ lower control limit } LL = \bar{X} - 2(s[X]\sqrt{(N)})$$

Where  $\bar{X}$  represents the mean of the adjusted change scores,  $s[X]$  is the standard deviation of the change scores and  $N$  represents volume.

The 99.8% control limits correspond to testing if an indicator is different from the benchmark using a two-sided test at a significance level of 0.002. The 95% control limits correspond to a two-sided test at a significance level of 0.05.

<sup>26</sup> ibid

### Case-mix Adjustment

The case-mix adjusted health gain for provider  $j$  ( $c_j$ ) is given by the expression:

$$c_j = (r_j * h_2) - h_1$$

where  $h_2$  is the national mean post-operative score,  $h_1$  is the national mean pre-operative score and  $(r_j * h_2)$  is the case-mix adjusted post-operative score for provider  $j$ .  $r_j$  is a parameter calculated as the mean ratio of actual to predicted post-operative scores averaged across all patients for provider  $j$ .

### Calculating the standard deviation

If it is not known, it is possible to derive the sample (national) standard deviation for the change score from the published data as follows:

The published confidence limits around the national case-mix adjusted health gain ( $c_j$ ) are derived from the confidence limits around the parameter  $r_j$ :

$$r_j \pm z \cdot s(r_j) / \text{SQRT}(n_j)$$

where  $z$  is the significance level (approx 3 at 0.2% significance and 2 at 5% significance levels),  $s(r_j)$  is the standard deviation of  $r_j$  and  $n_j$  is the number of patients (observations) in provider  $j$ 's data.

From the published confidence limits we can calculate  $s(r_j)$ . Denoting the published upper confidence limit for  $c_j$  as  $u_j$ :

$$s(r_j) = (\text{SQRT}(n_j) * (u_j - c_j)) / (z * h_2)$$

$$\text{Var}(c_j) = h_2^2 * \text{Var}(r_j)$$

so

$$s(c_j) = (\text{SQRT}(n_j) * (u_j - c_j)) / z$$

# Annex B: Example Report on Potential Outliers

## Introduction

This publication reports on those providers that have been identified as potential outliers, based on statistical analysis of the national PROMs data. Further investigation is required to confirm providers' status as an outlier.

The May 2011 PROMs publication contains the ninth post-operative outcomes data release and covers the period April 2009 to December 2010. During that period there were 420,000<sup>27</sup> eligible hospital episodes. Of those, 290,000 pre-operative questionnaires were returned, a response rate of 68%. Of the 230,000 post-operative questionnaires sent out 170,000 have been returned, a return rate of 75%.

To identify potential outliers, we use funnel plot analysis to compare providers' average health gain to the national average. Further details on the methodology used can be found in the technical appendix to the outlier policy.

The current national PROMs programme collects data on knee replacement, hip replacement, varicose vein surgery and groin hernia surgery. This report presents the results of our outlier analysis for each of these individually.

## Notes on interpretation

In each of the funnel plot presentations, the dashed lines are the 95% control limits and the solid lines are the 99.8% control limits. Each provider is represented as an individual point on the scatter plot. We consider those lying outside the 95% control limits as potential "alerts" and those lying outside the 99.8% limits to be potential "alarms".

---

<sup>27</sup> Figures have been rounded to the nearest ten thousand. Exact figures can be found on the NHS Information Centre's HES Online website:  
<http://www.hesonline.org.uk/Ease/servlet/ContentServer?siteID=1937&categoryID=1488>

## KNEE REPLACEMENT

### Generic PROM

We have identified 11 providers as potential negative outliers based on mean adjusted average health gain, as measured by the EQ-5D index. Of these, 9 are considered to be in the “alert” category and 2 are considered to be in the “alarm” category. This is illustrated in figure 1.

Figure 1: Adjusted average health gain (EQ-5D) for knee replacements

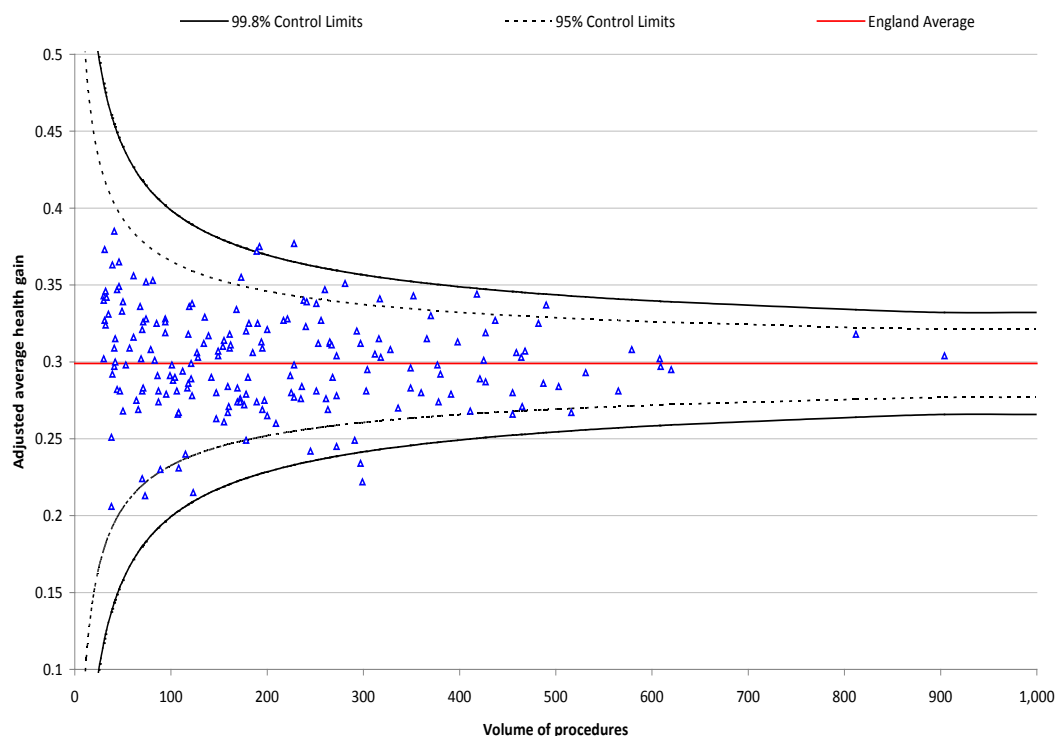


Table 1: Summary of provider status, EQ-5D for knee replacements

EQ-5D					
Control Limits	Upper @ 99.8%	Upper @ 95%	In Control	Lower @ 95%	Lower @ 99.8%
Number of outliers	3	7	171	9	2

### Condition specific PROM

We have identified 23 providers as potential negative outliers based on mean adjusted average health gain, as measured by the Oxford Knee Score (OKS). Of these, 13 are considered to be in the “alert” category and 10 are considered to be in the “alarm” category. This is illustrated in figure 2.



## PROMs in England: A Methodology for Identifying Potential Outliers

Figure 2: Adjusted average health gain (OKS) for knee replacements

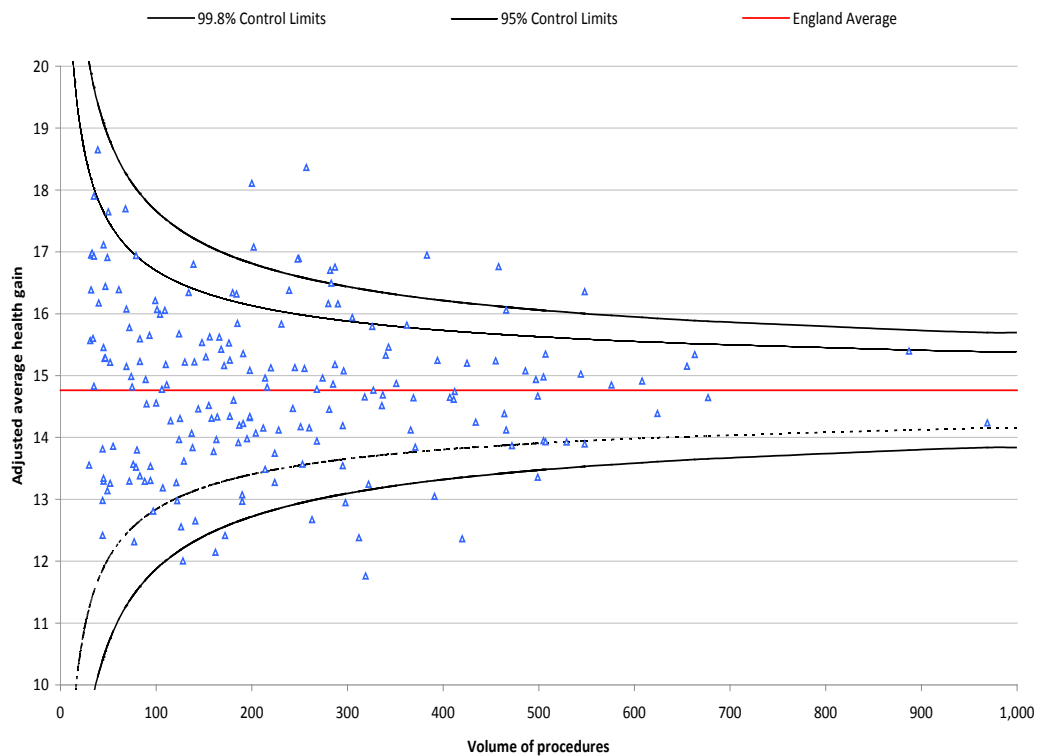


Table 2: Summary of provider status, OKS

OKS					
<b>Control Limits</b>	<b>Upper @ 99.8%</b>	<b>Upper @ 95%</b>	<b>In Control</b>	<b>Lower @ 95%</b>	<b>Lower @ 99.8%</b>
<b>Number of outliers</b>	<b>11</b>	<b>13</b>	<b>149</b>	<b>13</b>	<b>10</b>

### Potential Negative Outliers

Table 3 lists the providers who we have identified as potential negative outliers. It provides detail on which metric this is based on and which category they fall into.

Table 3: List of potential negative outliers in knee replacements

<b>PROVIDER</b>	<b>EQ-5D</b>	<b>OKS</b>
PROVIDER A	Alarm	Alarm
PROVIDER B	-	Alert
PROVIDER C	Alert	-
PROVIDER D	Alert	Alert

## Hip Replacement

### Generic PROM

We have identified 24 providers as potential negative outliers based on mean adjusted average health gain, as measured by the EQ-5D index. Of these, 16 are considered to be in the “alert” category and 8 are considered to be in the “alarm” category. This is illustrated in figure 3.

Figure 3: Adjusted average health gain (EQ-5D) for hip replacements

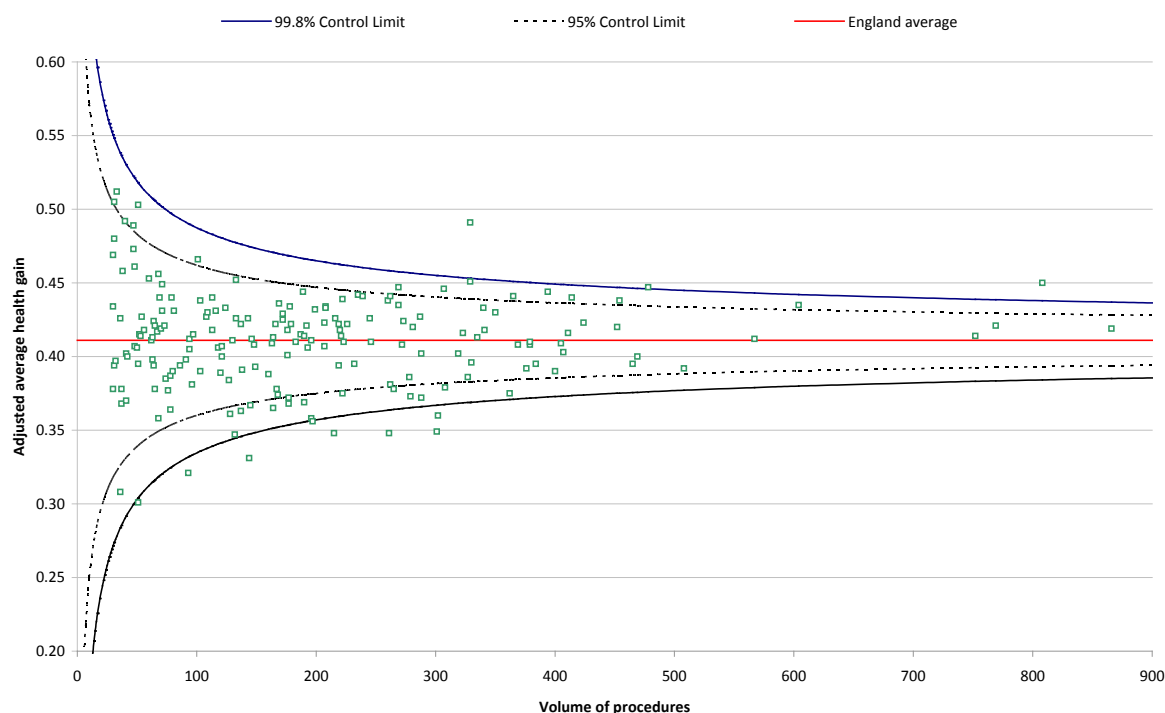


Table 4: Summary of provider status, EQ-5D for hip replacements

EQ-5D					
Control Limits	Upper @ 99.8%	Upper @ 95%	In control	Lower @ 95%	Lower @ 99.8%
Number of outliers	3	14	145	16	8

### Condition specific PROM

We have identified 30 providers as potential negative outliers based on mean adjusted average health gain, as measured by the EQ-5D index. Of these, 14 are considered to be in the “alert” category and 16 are considered to be in the “alarm” category. This is illustrated in figure 4.

# PROMs in England: A Methodology for Identifying Potential Outliers

Figure 4: Adjusted average health gain (OHS) for hip replacements

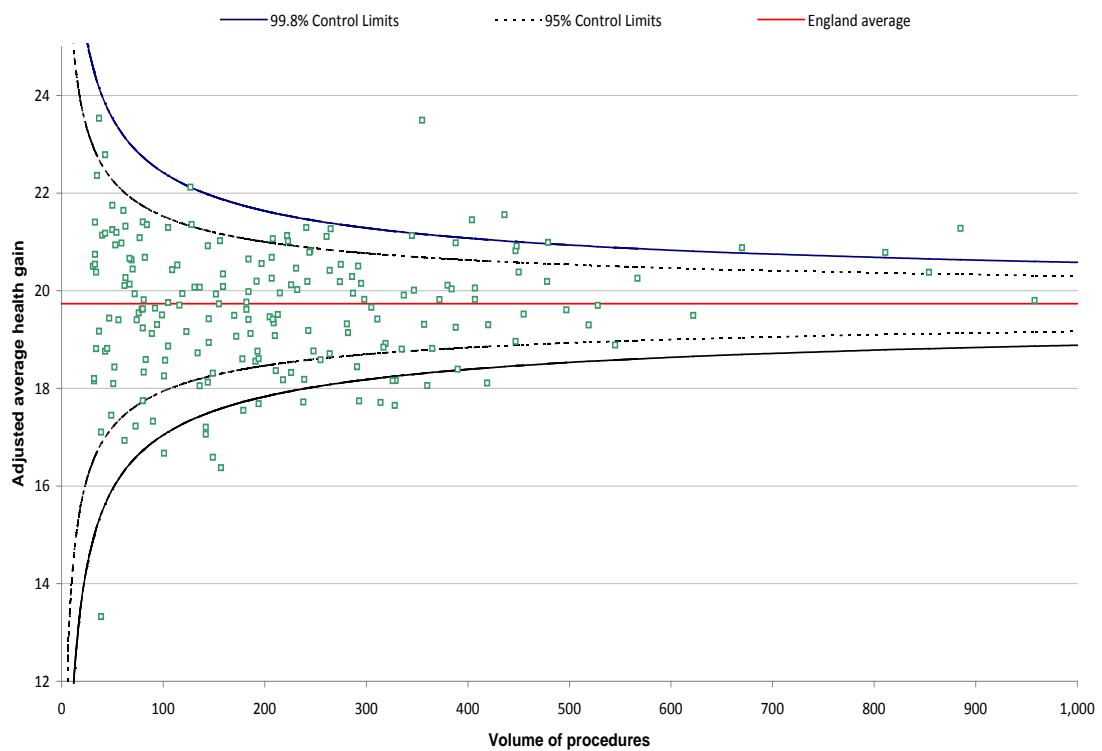


Table 5: Summary of provider status, OHS

OHS					
Control Limits	Upper @ 99.8%	Upper @ 95%	In Control	Lower @ 95%	Lower @ 99.8%
Number of outliers	8	14	143	14	16

## Potential Negative Outliers

Table 6 lists the providers who we have identified as potential negative outliers. It provides detail on which metric this is based on and which category they fall into.

Table 6: List of potential negative outliers in knee replacements

PROVIDER	EQ-5D	OKS
PROVIDER A	Alarm	Alarm
PROVIDER B	Alarm	Alarm
PROVIDER C	Alert	-
PROVIDER D	Alert	Alert

## Varicose Vein Surgery

### Generic PROM

We have identified 5 providers as potential negative outliers based on mean adjusted average health gain, as measured by the EQ-5D index. Of these, 3 are considered to be in the “alert” category and 2 are considered to be in the “alarm” category. This is illustrated in figure 5.

Figure 5: Adjusted average health gain (EQ-5D) for varicose vein surgery

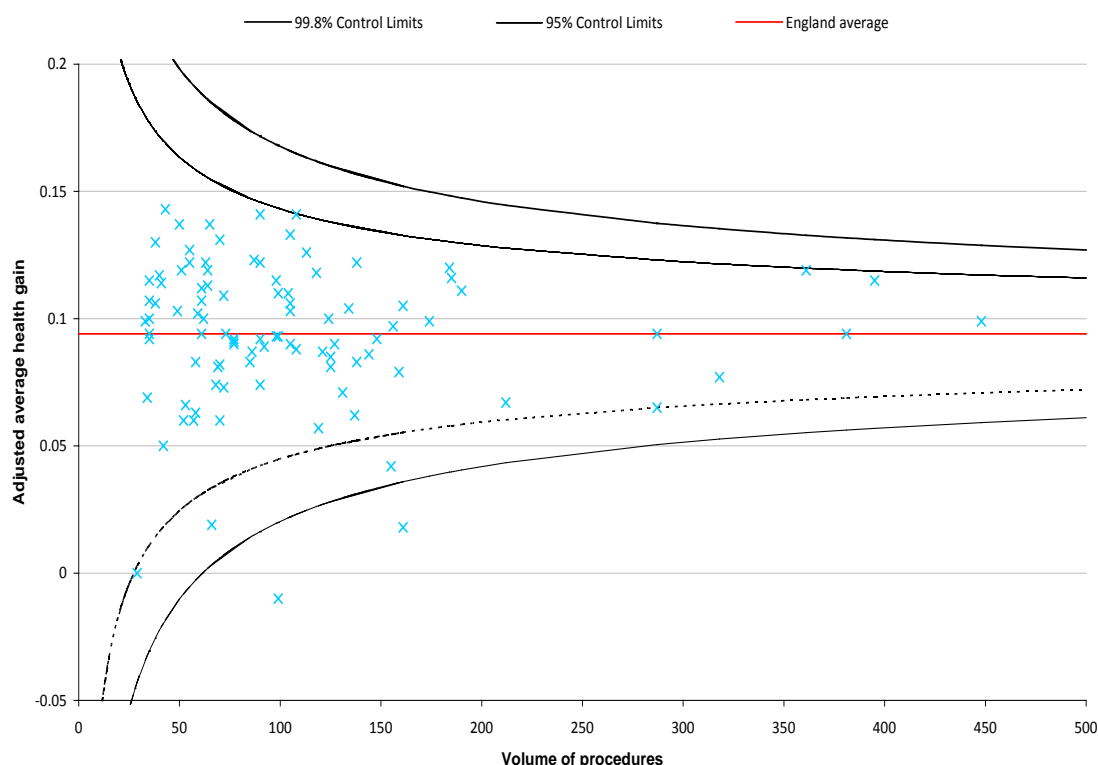


Table 7: Summary of provider status, EQ-5D for varicose vein surgery

EQ-5D					
Control Limits	Upper @ 99.8%	Upper @ 95%	In Control	Lower @ 95%	Lower @ 99.8%
Number of outliers	0	0	92	3	2

### Condition specific PROM

We have identified 8 providers as potential negative outliers based on mean adjusted average health gain, as measured by the EQ-5D index. Of these, 7 are considered to be in the “alert” category and 1 is considered to be in the “alarm” category. This is illustrated in figure 6.

## PROMs in England: A Methodology for Identifying Potential Outliers

Figure 6: Adjusted average health gain (AVVQ) for varicose vein surgery

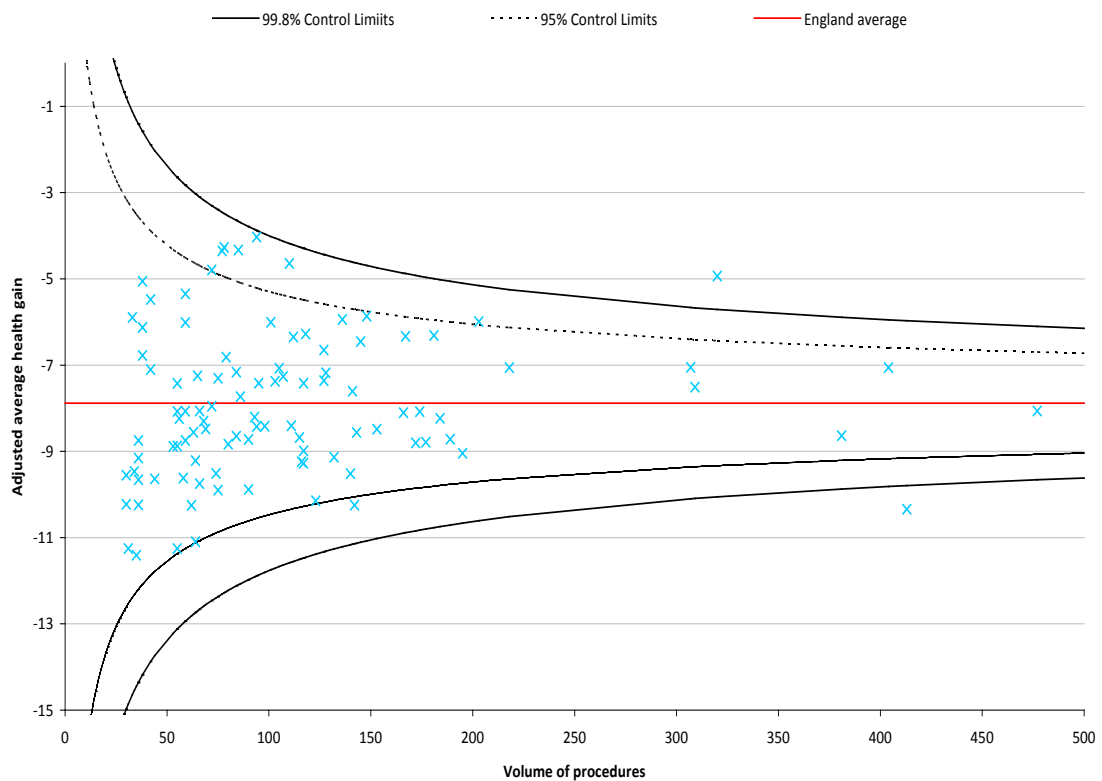


Table 8: Summary of provider status, AVVQ

AVVQ					
Control Limits	Lower @ 99.8%	Lower @ 95%	In Control	Upper @ 95%	Upper @ 99.8%
Number of outliers	1	1	89	7	1

### Potential Negative Outliers

Table 9 lists the providers who we have identified as potential negative outliers. It provides detail on which metric this is based on and which category they fall into.

Table 9: List of potential negative outliers in varicose vein surgery

PROVIDER	EQ-5D	OKS
PROVIDER A	Alarm	Alarm
PROVIDER B	Alarm	Alarm
PROVIDER C	Alert	-
PROVIDER D	Alert	Alert

## Groin Hernia Surgery

We have identified 8 providers as potential negative outliers based on mean adjusted average health gain, as measured by the EQ-5D index. All of these are considered to be in the “alert” category; none are considered to be in the “alarm” category. This is illustrated in figure 7.

Figure 7: Adjusted average health gain (EQ-5D) for groin hernia surgery

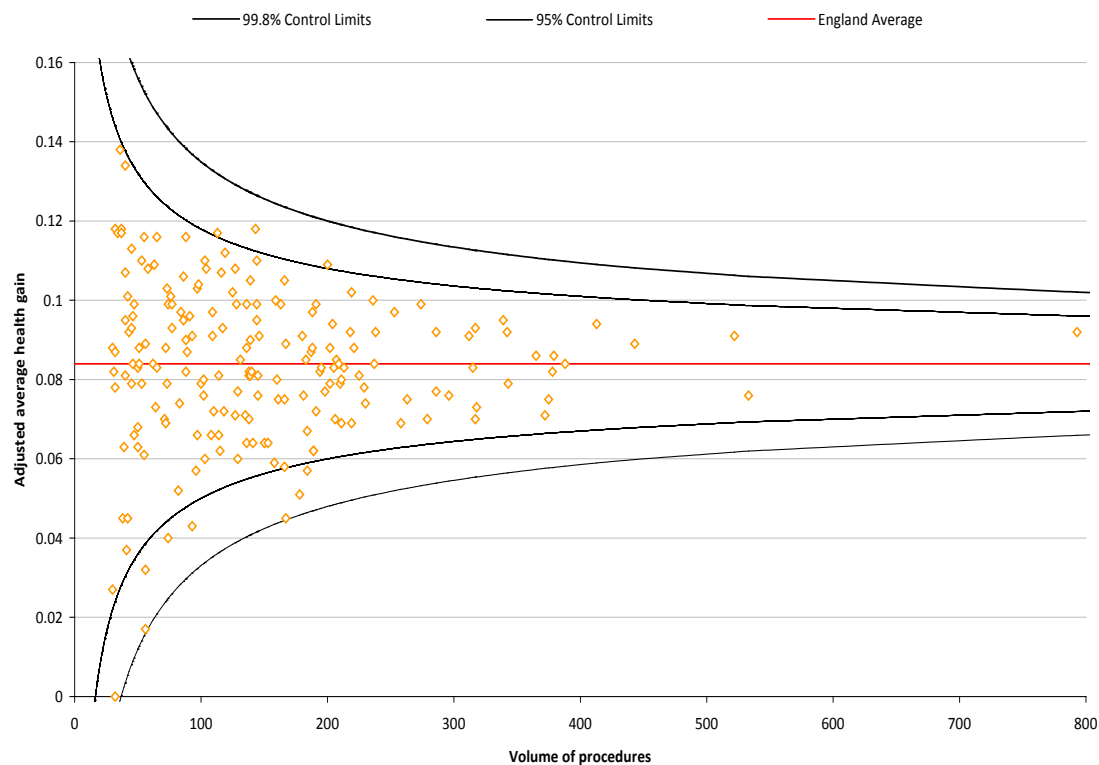


Table 10: Summary of provider status, EQ-5D for groin hernia surgery

EQ-5D					
Control Limits	Upper @ 99.8%	Upper @ 95%	In Control	Lower @ 95%	Lower @ 99.8%
Number of outliers	0	3	181	8	0

### Potential Negative Outliers

Table 11 lists the providers who we have identified as potential negative outliers. It provides detail on which metric this is based on and which category they fall into

Table 11: List of potential negative outliers in groin hernia surgery

PROVIDER	EQ-5D	OKS
PROVIDER A	Alarm	Alarm
PROVIDER B	Alarm	Alarm
PROVIDER C	Alert	-