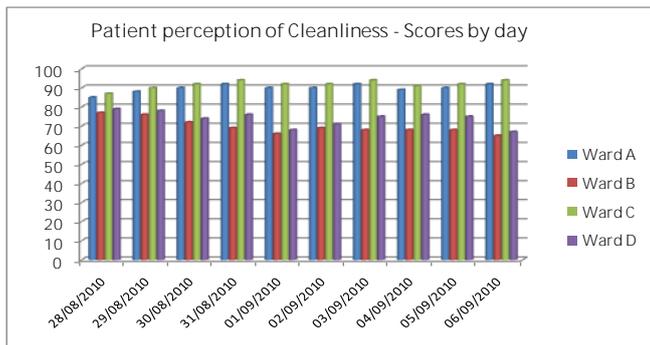


You would think that the days of paper exit questionnaires have all but disappeared when you review the market for a means of collecting experience data. Spurred on by the government's NHS health guidance for "near real-time" patient data collection, a wider selection of collection methods now exists for healthcare organisations seeking to measure their patients view of care quality.

As someone who has struggled with the local usefulness of national initiatives such as the Annual Inpatient survey each summer, whose results used to take 18 months to appear, "near real time" seems one obvious goal to aspire to.

What is meant by "near real time" ?

The key requirement is that results should be seen quickly and continuously, i.e. this is no longer a one-off sampled survey which traditionally was planned, created, distributed, completed, returned, data input, processed, data analysed and a report finally written, all on an annual basis.



This fundamental change of approach allows a number of key benefits:

1. The measurement should be continuous and the results should be capable of being seen daily
2. The coverage should aim for all patients and not a sample
3. Responses should be collected anonymously and without fear of antagonising any staff looking after patients
4. The measured population should be representative of all patients using a service, not a skewed section and should allow input from carers and others involved.
5. What is being measured should be contemporary, with local relevance, but with a capability to provide continuous performance metrics against valid external benchmarks.
6. Collection models should provide the capability for ever changing needs.
7. Finally, the model should allow the measurement to become a standard part of service delivery, not a regulatory requirement add-on.

Stuart Mathieson is the founding Director of CoMetrica Ltd. An Oxford MBA Graduate, Stuart has many years of experience as a clinician and senior manager in health-care. As a user of clinical performance and experience information himself, he recognised the frustration seen by colleagues when information was neither timely, joined up or accurate. Following significant investment in the development of a truly innovative service, CoMetrica now provides continuous experience and clinical outcomes measurement to health and care organisations as part of a portfolio of powerful performance management services. For more information, consult the CoMetrica website at www.CoMetrica.co.uk



What collection models are available to achieve this?

1) Push button devices

Recent years have seen the emergence of push devices, pre-programmed with a few satisfaction questions. These are handed to patients or are placed in strategic places such as reception desks. The apparent simplicity of such devices such as the Dr Foster PET (Patient Experience Tracker) was originally attractive in that it allowed some selected patients to provide a "point of care" view on their experience. Typically, a few devices were purchased by hospitals for distribution around wards and clinics. They are "programmed" by having the questions printed on a covering film applied to the device and the device records which button from "good to bad" was pressed for each of 5 questions. The devices were then collected up and the data uploaded to the Dr Foster software which could show the proportion of button presses in each category for each question.



Pros

- Patients can press a button during their stay (but not necessarily at the actual point of experience required)
- The questions should be simple to understand

Cons

- Limited, skewed sample as not all patients can access the device
- Limited to five very simple questions
- Questions slow and expensive to change
- Staff time consumed in managing the devices
- You know nothing about the person pressing the button.

By virtue of the previous contract between the Department of Health and Dr Foster, Dr Foster had exclusive access to hospital HES data and so many Trusts already used their performance measurement services. It was a logical extension for Trusts to buy into their PET scheme.

Most contracts were for 3 years and many are now due for renewal. However, the limited usefulness of the data, and the recognition that highly selective skewed samples were used, mean that many trusts are now seeking alternatives to PETS.

2) Improved touch screen devices and Kiosks

There are a number of other push button and touch screen devices in the market now which are capable of more detailed experience questions. These can be hand-held devices or free-standing kiosks. These allow for example, the patient to be asked their age, which ward or clinic they have attended and some other very basic demographics which increases the usefulness of the results.

The problem is, if you spend half the patient's time asking them things you already know about them, you soon come up against question fatigue and completion rates are compromised so questionnaires still have to be very short.

There are still the drawbacks of device management, minority sampling and skewing. In the sales information for push button/touch screen devices used in the clinical setting, "Real Time" i.e. measured at the time the aspect is being experienced, is strongly marketed as a distinct advantage.

In reality, patients can rarely be asked at the actual time of experiencing say, food, treatment or environmental factors. These devices tend to be used only when patients have spare time as an inpatient or as they leave which compromises confidentiality, questionnaire length & depth. MORI (2008) found that patients should only be asked when they are well enough to do so and benefit from having time for reflection and consideration of the whole pathway before responding as would be the case with a survey completed when the patient is at home.

So is there value in using such devices?

Survey systems which collect data during care emphasize speed over accuracy and rigour. Nelson et al (1991) says that "if the goal of an organisation is to measure quality with reliable and valid indicators, it will need to employ more robust methods... with increased response rates and hence to ensure results are statistically significant."

If questions can be programmed locally and quickly, for example to measure general satisfaction with a particular clinic, then having devices available on exit can give a high level view of some satisfaction aspects as long as the skewing is recognised.

Comparisons of survey speed are often made against manual paper processes. Modern electronic questionnaires, whether collected on paper or on-line can be faster than push button devices as there is no processing stage or device management and uploading time.



What do we mean by skewing?

You have to consider, for your patient population, which patients will have access to such devices, which can use them and which are inclined to use them. Given that the average age of most patients is over 70, using any technology in the health care setting is low on their list of priorities compared with their health, any worries they may have and the more familiar communication channels they use such as face to face.

Passively placed devices such as kiosks and reception desk touch pads tend to be used only by patients who are encouraged by staff to do so and for very short questionnaires. Waiting rooms could be good opportunities for capture except that usually the patient has had little of their total experience at the time they are waiting.

The result is that responses tend to come from a small sample of patients who are confident with technology, articulate, recognise the value of the collection and have free time on their hands. The experience may be measured prematurely, for example not taking into account discharge arrangements or allowing time for reflection and comparison. Measurement taken in the clinical setting can suffer from lack of confidentiality and patients may be unwilling to give poor responses while still in the care of the people they are commenting on. Response options are limited, textual and cannot probe or provide patient tailored choices.

Given the high volume of patients going through health services, it is easy to be lulled by a false sense of security through relatively high response volumes, say 50 patients per day completing such surveys. However, these could be as low as 5% of the patients using that service and critically, may be skewed away from the very group of patients you need to be measuring most i.e., the frail, elderly, those with communication barriers and those who are unwell.

So what are the alternatives?

3) **Face to Face interview programmes** can be highly effective since an interviewer can ask more detailed questions, probe and valuable commentary can be collected. Importantly, the questions asked can be more relevant to that individual patient. The interviewer should be an impartial third party and volunteers have taken up this role in some organisations.



There is some evidence that the results of questionnaires or surveys administered by individuals (as opposed to 'self-interviewing') may be more susceptible to social desirability bias – in that people provide the responses they think the 'researcher' wants (Duffy et al 2005).

A number of trusts have used such an interview scheme to supplement annual surveys. They allow higher degrees of qualitative data and are useful for exploring already identified areas of concern. They are of course limited in volume, but more importantly, are only suitable for articulate patients with time on their hands – inpatients rather than day cases for example. Post discharge interviews can be highly valuable but are only possible for very small samples.

4) **Traditional paper surveys** can have very high reach, they can for example be given or sent to all patients, but response rates have tended to be poor because some of the questions may not be relevant and half of some survey's content is wasted in collecting demographic information to improve the usefulness of the results. Response rates for traditional generic satisfaction paper surveys vary from 10% - 30%. Those given out or sent from clinical departments tend to have the higher rates.

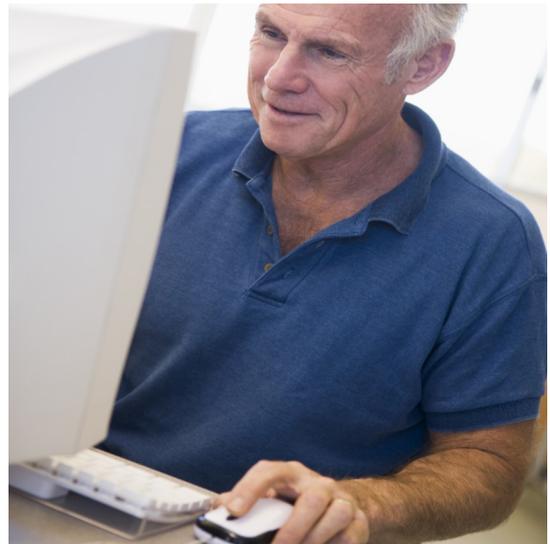


5) **Telephone surveys** can have moderate reach if attempted at various times of the day and evening and can be nearly as useful as face to face, but user acceptance is low. Elderly patients do not answer unexpected calls and phone number records are often incorrect. The telephone can be useful as a means of reminders for postal surveys, but as mainstream use are expensive and have low coverage.

6) **Automated Telephone surveys (CATI)** are very cheap to operate, but suffer from the same accessibility challenges as operator based telephone surveys with the added problem that they have extremely low user tolerance. Most calls are abandoned by recipients within a few minutes.

7) **Email surveys** are cheap to operate but have highly selective reach. Email addresses are rarely recorded in hospital data, and quickly become out of date.

8) **Web-based forums** such as the NHS Choices website and a number of voluntary and commercial web forums allow anyone to make comments about anything but the aim is for patients to comment on their experience at a specific organisation. A quick review of such forums indicates that contributions tend to be from those with bad experiences and a compensating smattering of patients eager to express their gratitude. No measure is available.



8) **Web-based surveys** are popular and cheap to run but again can only reach a minority of patients. You also have to get the invitation to complete the survey to patients in the first place and usually this is by email with the drawbacks identified above.

9) **Mobile phone based** channels such as smartphones and to a lesser extent, text systems can be used to target specific patients as long as those patients are in such technology enabled groups. These have flexibility in that they can link to on-line surveys which can be more comprehensive but start from a very low patient coverage base.



What data collection models can provide all the benefits of near real-time measurement?

Each of the above models has its virtues but it is clear that to provide maximum reach, more than one channel is necessary using different methods for different patient groups. What does not change is the need for seamless continuous communication with patients and the capability to collect information at multiple stages by multiple routes, joining up the information into one view. Organisations seeking to fulfil an engagement strategy should adopt systems which are flexible and can reach more than 95% of their patients.

A note on survey design

Often local questionnaires are designed by managers themselves. The questions asked are often the wrong ones and will not measure experience. Examples seen are often leading, closed and satisfaction based such as " How much do you agree with this statement" – with Lickert scale type response options. Taking this approach, it is easy to generate questionnaires which will result in " good" satisfaction results because the questions have insufficient challenge and anchoring.

Questionnaires should always be trialled and systems should allow questions to be changed easily in real time. Questionnaire design and attitude measurement is discipline in its own right and amateur or inexperienced attempts can give a false sense of apparent patient satisfaction which could ignore fundamental problems. To avoid this, professional guidance in existing measure selection and use, and in the development of any new measures should be sought, together with guidance in interpretation of results and validity.

Of particular importance in health and care services, ensuring equality of access and avoidance of skew in resulting should be paramount. Professional measurement advice and ongoing consultancy should be an integral part of any commercial service bought by health and care organisations.

10) Fully managed measurement services

This need for a comprehensive, fully supported service should be clear in the organisation's measurement strategy. Trusts and other organisations need to assure themselves that all patients can have genuinely equal access to measurement, that the measures are sensitive and appropriate and that results are rapid and continuous so they can become a standard part of delivery. Systems capable of fully supporting these inclusive strategies are only available as a fully managed service. A comprehensive service capable of meeting these needs is available through the COM-Q service provided by CoMetrica which includes clinical and experience measurement consultancy and provides multiple collection channels.

Language and understandability

There are often requests that informational material including survey questionnaires are made available in several languages. Local statistics can be helpful in guiding this, but recording of ethnicity can be a poor indicator of the language best understood and this should be recorded separately by organisations and used to drive the language used in the questionnaires. Systems should be capable of automatically producing questionnaires in the chosen language when this is known.

Language however is only one aspect. When you need to reach the widest possible groups of patients, traditional multiple choice questions, whether collected electronically or on paper have several barriers which need to be overcome. The first is understanding the questions, the response choices given and the gradient and width of perceptive scale. One patient's view of what constitutes " Poor " can be totally different to another patient.

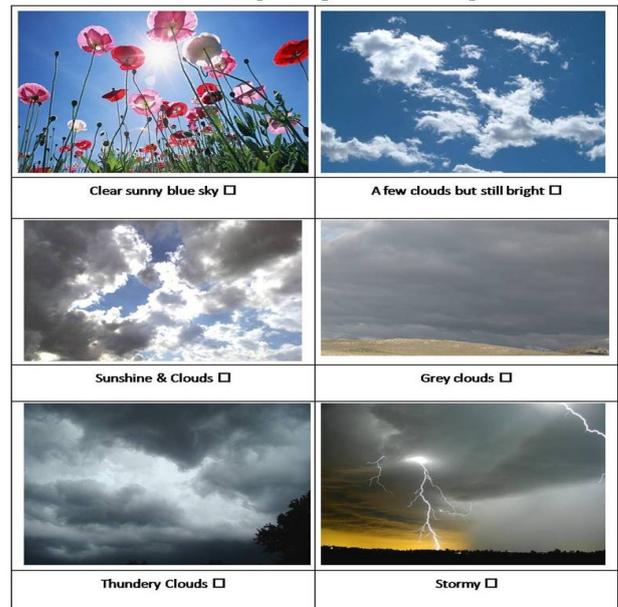
The other key aspect is capturing and retaining the respondent's in-

terest. A multi-page questionnaire consisting of many similar textual questions with Lickert type response choices e.g. " How much do you agree with the following statement: I was satisfied with the quality of food I was given – Very satisfied, Satisfied, Neither satisfied nor dissatisfied, Dissatisfied, Very dissatisfied" are extremely tedious to complete, provide no challenge or interest and no anchoring (see later). Expecting a patient to answer more than a few of these faithfully is unrealistic.

So how can reach and interest be improved?

Patients are no different to most of us, and most of us prefer visual clues of scale however subtle. Examples of this are the use of emoticons (smiley faces) to express degrees of satisfaction or pain. These can be highly effective, but should not be devalued through over-use. A minority of patients object to " infantile" scales and they should always be accompanied by text guidance in the understood language. There is however much greater scope for visual cues than smiley faces; pictures, graphics and illustrated scales are well received by patients, particularly where analogue scales allow patients to indicate " where on the scale they wish to place their cross" rather than enforcing discrete values such as " good" or " excellent.

How do you feel today?



Pictures have the greatest psychometric potential as they allow evoked responses, recollection and visual example e.g. did you see any of the following (good or bad) things when you were in hospital? Patients can be led through a recollective journey by picture prompts of places they visited or were treated at and this can reach wide groups of patients such as children, those with learning disabilities or mental illness in addition to transcending language and culture.

Personal, Proxy and carer responses

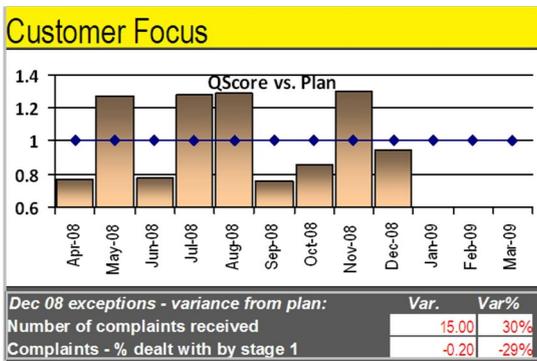
For many patients, their experience in healthcare is shared in part with those who care for them such as family, close friends and carers. It is natural that patients wish to reflect off others when considering their experience and many patients complete surveys with some degree of input or at least discussion with others socially near to them.

In addition, some patients are considered not capable of completing any surveys and these are completed by proxy through carers or family.

Such input and proxy completion where required is good, it increases response rates, provides more grounding and relativity and can sometimes allow greater complexity. What is vital however is that it is known that there has been this input so that responses can be classified as such. This is possible with paper and on-line questionnaires but not with devices. Rarely are responses made on devices by patients who need this support.

Relativity & Anchoring

Questionnaire design and attitude measurement is a scientific discipline which prompts the development of validated question instruments. This validation is intended to ensure that the measure is fit for purpose (although this can have drawbacks- see later).



Although many people know that clinical measures such as PROM (Patient Reported Outcome Measures) can have benchmarks, not many know that benchmarks are available in experience surveys such as service quality. An example of such measures is the SERVQUAL range of measures. These have been translated for healthcare setting use and are interesting in that they provide anchoring for respondents. An example would be " In the very best hospital outpatients department, how long would be reasonable to wait before being seen? 0 minutes, 1-5 minutes, 6-10 minutes, 11-20 minutes etc. Then the comparator question " When you were in outpatients at XYZ hospital, how did your wait compare with this? – etc. By providing an anchor to response choices, referenced scale is possible which allows some external benchmarking.

Measurement or feedback or both?

It is easy to confuse these. There are many means of collecting feedback, formally and informally, and some of this can be prompted by specific questions. Anyone who has reviewed such feedback will tell you that although you can ask about one aspect of the service provided, the response given by patients can often be about something else, of equal or greater importance to them.

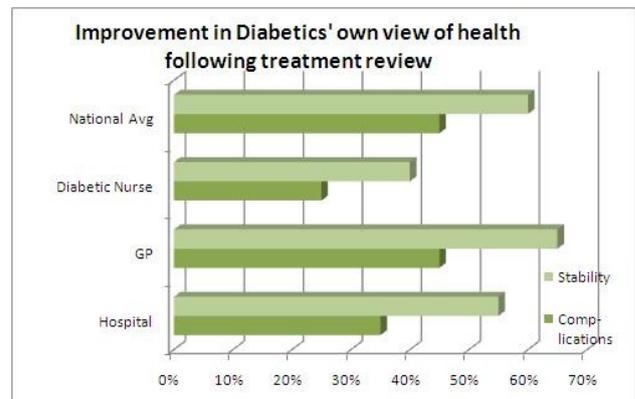
Measurement is different; it requires structured response choices, with values assigned to responses and an algorithm to calculate a value for that response. Often measures such as satisfaction with communication are achieved by calculating a composite score from a set of discrete questions, some of which may be weighted unequally.

The most useful experience measurement systems allow for continuous quantitative results AND qualitative comments, classified where possible into appropriate themes.

Statistical validity

Many of us are familiar with psychometric instruments which achieve statistical validity by asking the same question multiple times, in a slightly different way. The purpose being to challenge the respondent for consistency and weed out outlier responses which might otherwise skew the result. This is done to improve the reliability of an instrument in the case where an empirical result is required per respondent.

For similar reasons, where pharmaceutical trial outcome measures are being collected, absolute reliability, sometimes in relatively low respondent volumes is required, and question instruments are carefully validated to ensure there is no unrecognised bias. This often means that the PROM instruments developed originally for clinical RTC (Randomised Control Trials), tend to be long with many questions, repeating the same theme to improve accuracy. It is not unusual to see such questionnaires containing 50 or more questions but the respondents are usually paid to complete such questionnaires and are carefully selected for reliability in completion. Questionnaires of this length are not appropriate for most patients due to the lack of interest, time and lack of ambition to contribute to research, whether incentivised or not.



The difference with surveys carried out in healthcare settings is one of volume. Where the research surveys described above need question theme duplication to achieve reliability, questionnaires for patients can achieve similar reliability through the greatly increased volume of respondents. The results can therefore be equally valid on a cohort basis if not an individual basis.

Ensuring statistical robustness in this context is more about understanding the drivers of accuracy. This can be more important than using a highly standardised, validated survey instrument. Understanding sensitivity, volume and coverage (how many patients out of all service users have responded) can be more important than whether a survey has been validated or not when interpreting results.

For this reason, while there should always be an aim to include some validated instruments, locally unvalidated surveys developed by departments themselves can remain useful in a local context.

Patients' valuation of surveys

Getting responses from patients is important and organisations should seek to maximise response and coverage rates where possible. In addition, there is a high value placed by the patients themselves on being asked, whether they respond or not. Where services build experience measurement into service delivery as a standard for all patients rather than as a sample, an important message about inclusion is conveyed to patients.

Multiple uses and Multiple Users

If not co-ordinated by a measurement strategy and use of a comprehensive measurement system, the contemporary recognition of the need for measurement can lead to a proliferation of surveys patients are expected to complete at many stages of their treatment.

Patients should only get one questionnaire, in whatever form, at each stage of their treatment. This means that measurement needs to be co-ordinated, ideally by using one measurement system, capable of generating results for different users from the same survey. The system should allow different survey content for different patients, based on the services they have used, their treatment, diagnosis or other demographic factors.



This means that the survey system needs to be linked to patient data, and very few are, particularly the anonymous capture devices. By linking survey generation to patient information, the content of the questionnaire can automatically be matched to individuals or groups of patients. This can be in real time and has the additional benefit that all results can be linked to patient data allowing deep analysis, drill down and identification of the root of variances and problems. The benefit for patients is that they only receive questions which are relevant to them personally.

Examples of this matching would be the capability to include in a survey for another purpose, questions about diabetes management automatically for those patients who happen to be diabetic.

Clinical measurement

In measuring the total impact health services have on patients, clinical effectiveness cannot be dissociated from experience. It is well recognised that the experience patients receive can have a marked impact on the overall effectiveness of their treatment.

The value of measuring patients' perception of their clinical outcomes is recognised and the UK government's scheme to routinely collect PROMS for Hip and Knee joint replacements and some other specific treatments is being watched with interest internationally. The potential of this information is significant in planning, valuing treatments and comparing outcomes between organisations and clinicians.

By using a comprehensive system such as the COM-Q service, clinical



PROMS can be collected alongside experience measures. This provides the complete picture, how was it delivered and how effective it was.

Communication content

Given that measuring experience must include many channels, survey questionnaires are only one of these and a comprehensive strategy will include interviews, focus groups and the opportunity to consult on future policy and configuration issues. A comprehensive service will provide communication such as letters, invites to focus groups etc and that these can be targeted at the relevant patient groups without the need to use questionnaires.

Using the results

A comprehensive continuous system will provide qualitative information in the form of patient comments on a daily basis. Managers at multiple levels in the organisation should have access to these results on a daily basis on-line without relying on someone to run specific reports and send them out. Quantitative results should be reviewed at least weekly by the appropriate staff and the systems should be capable of presenting trends over short and longer term periods, comparing results and cyclical patterns such as day of week. External and internal benchmarks should be able to be placed on charts to compare results.

To encourage use of the information in supporting decision making, the system should be capable of exporting graphical reports, tables and charts directly into Word, Excel, Powerpoint and PDF documents.

A more sophisticated approach

As has been illustrated in the preceding sections, measuring patient experience properly, to produce valid results which are useful in operational as well as strategic management terms, requires a considered strategy supported by the right systems. This joining up is not always

straightforward and requires significant experience to get right. Expecting front line staff to undertake this co-ordination in addition to their " day job" is unrealistic and specialist help is usually required, even where there is local recognition of the factors and ambition. This means that in engaging expert help, the experience any supporting company has in measurement metrics, use of measurement instruments, clinical expertise and results interpretation needs to be considered. If you select a company on the basis of technical devices alone, the results could be worthless.

At the end of this article, in Appendix A, a comparison of the different measurement approaches is given.

Relative cost of different methods

In evaluating the cost of near real-time measurement, there are four factors to consider:

1. Cost of sufficient devices for all patients to be able to respond
2. Cost of staff time in managing devices and encouraging their use
3. Cost of any maintenance/ license costs
4. Cost of any ongoing question changes/ configuration/ training
5. Cost of any ongoing advice/consultancy in measurement

In the comparison table below, we based this on a hospital who needs to measure the experience of 10,000 elective inpatients before and after admission and a selection of outpatient services. Costs are all – inclusive estimates based on user experience.

Staff time

Using a collection model with low coverage, such as touch devices , in order to get a response from most patients, at least 2 devices would be needed per ward or clinic and almost 2 staff would be required in order to get all patients to respond. While some patients can use the device unaided, this will be a skewed minority and the devices still

have to be programmed, distributed, managed, data uploaded and patients encouraged to use them. This front line staff burden increases cost significantly. Interviewing is the most expensive as expected and would only be used as part of a selective programme of engagement following comprehensive quantitative analysis. Even if volunteers interview patients, staff time is spent organising this and there are expenses to cover. The cheapest is automated telephone and internet based collection. However, the coverage of both of these is extremely low, at best only 20% of the target population would be covered. Overall the fully managed service has the greatest strengths (see overleaf) at the lowest cost of the high coverage methods.

For more information about Fully Managed Patient Experience & Outcomes Measurement Services and the inclusion of clinical PROMS, see the document library at <http://www.cometrica.co.uk/document-library.php>

References

Duffy B, Smith K, Terhanian G and Bremer J. (2005) Comparing data from online and face-to-face surveys. International Journal of Market Research, 47, 6, 2005

Ipsos MORI (2008) Real Time patient feedback - information patients need and value. London. Ipsos MORI

Nelson E C et al (1991) The Patient Comment Card: A System to Gather Customer Feedback. Quality Review Bulletin, 17, 9, 1991

Data Collection methods -Cost comparison table		Simple Tracker	Developed Devices	Interviews	Telephone surveys – interviewer	Telephone surveys (automated)	Traditional paper surveys	On-line surveys/ Email	Fully managed service
Elective Inpatients pre-op p.a.	annual	10,000	10,000	10,000	10,000	10,000	10,000	10,000	10,000
Elective Inpatients post-op p.a.	annual	10,000	10,000	10,000	10,000	10,000	10,000	10,000	10,000
Outpatient volume p.a.	annual	20,000	20,000	20,000	20,000	20,000	20,000	20,000	20,000
Total measures p.a.	annual	40,000	40,000	40,000	40,000	40,000	40,000	40,000	40,000
Max questions/survey		5	20	30	20	15	25	25	30
Total Measure points		200,000	800,000	1,200,000	800,000	600,000	1,000,000	1,000,000	1,200,000
Elective wards	day	10	10	10					
Clinics	day	10	10	10					
Equip/Devices/surveys required	annual	40	40			40,000	40,000	40,000	40,000
Avg staff time to ensure collection, each survey	mins per survey	4	4	20	10		8		
Front-line staff rate	annual	£ 23,000	£ 23,000	£ 23,000	£ 23,000		£ 23,000		£ 23,000
Front-line staff cost	per minute	£ 0	£ 0	£ 0	£ 0	£ -	£ 0	£ -	£ 0
Staff Collection cost	annual	£ 38,333	£ 38,333	£ 191,667	£ 95,833	£ -	£ 76,667	£ -	£ -
Device/Survey cost	each	£ 750	£ 2,500			£ 0.75	£ 2.00	£ 0.60	£ 1.50
Device/survey cost	total	£ 30,000	£ 100,000	£ -	£ -	£ 30,000	£ 80,000	£ 24,000	£ 60,000
cost of 4 changes of question	annual	£ 10,000	£ -			£ 10,000	£ 10,000	£ 10,000	
License costs	annual	£ 8,000	£ 8,000			£ 15,000	£ 5,000	£ 5,000	£ 25,000
Consultancy	annual	£ 30,000	£ 30,000	£ 10,000	£ 10,000	£ 10,000	£ 10,000	£ 10,000	
Total	annual	£ 116,333	£ 176,333	£ 201,667	£ 105,833	£ 65,000	£ 181,667	£ 49,000	£ 85,000

