Benchmark of out of hours
*An overview across the services*

Report from the Primary Care Foundation
19 April 2012

Prepared by Henry Clay
Benchmark of out of hours - an overview

Contents

Executive summary........................................................................................................................................See stand-alone document

Benchmark of Out of Hours services - an overview ..................................................................................3
   Introduction................................................................................................................................................3
   Findings..................................................................................................................................................4

The detailed comparisons.................................................................................................................................7
   Overview.....................................................................................................................................................7
   Overall Demand..........................................................................................................................................8
      Demand for out of hours services by PCT compared with the proportion over 65.................................11
   Disposition and outcomes .........................................................................................................................12
      Towards Hospital or the Ambulance Service ..........................................................................................12
      Percentage advice, base or home visits .................................................................................................14
   Percentage urgent.......................................................................................................................................16
   Performance and responsiveness...............................................................................................................18
      Time to clinical assessment ....................................................................................................................18
      Time to the face to face consultation ....................................................................................................20
   Reporting and integration ..........................................................................................................................24
   Cost............................................................................................................................................................28
   Patient Perception........................................................................................................................................33
   Summary....................................................................................................................................................35
   Technical notes and considerations when looking at the analysis..............................................................36
      Possible errors because of the different ways that staff use the system, the way that data is extracted or the way that we have analysed it ..........................................................37
   Measurement points for time to clinical assessment ..................................................................................38
Benchmark of Out of Hours services - an overview

Introduction

This 4th round of the national out of hours benchmark is new and different. For the first time, everyone, whether a commissioner, provider or service user, can see how all services are performing on a wide range of headline indicators. This allows services to make helpful comparisons and learn from each other, driving up the quality of care across the country. The benchmark covers more than two thirds of the 151 PCTs (to be replaced by Clinical Commissioning Groups by April 2013) across England.

A detailed understanding of each service is as vital starting point for good clinical commissioning. Evidence is the currency for effective commissioning and the benchmark provides a wealth of information to enable services to highlight areas that need to be improved and implement changes. At the same time, commissioners also need to look at the bigger picture, understanding how out of hours services fit in to urgent care as a whole.

Crucially, it is not just the availability of transparent information but what people then do with it locally. That’s why we are pleased that many services have added their own commentary highlighting what actions they have already taken or plan to take in the future.

This is the fourth time that we have run the benchmark and the participant providers and commissioners have gained sufficient confidence in the data to allow the results to be shared publicly in a way that allows direct comparison. This is a significant step - there are no other areas of healthcare where direct comparison can be made of such a range of performance measures, cost and patient perception at PCT level, nor has it been done before in a way that allows the user to take account of explanatory factors such as deprivation, organisational structure or the nature of the area.

Out of hours services are run by organisations that work hard to deliver good care to their patients and they are staffed almost wholly by the GPs working in practices or as locums in the areas in which they operate. The benchmark is seen by its participants as a powerful way through which the organisations can improve still further and through which commissioners are better able to understand the capabilities and strengths of their service. Particularly because of the intense scrutiny that the sector has been under following a small number of disastrous cases the participants deserve credit for their openness in sharing this detail.

None of the 111 pilots are included in the analysis nor are any users of NHS Pathways, but we look forward to their participation in future benchmarks - when the nature of the benchmark and some of the detailed measures will also be changed.

We are very grateful for the work from many individuals in both providers and commissioners (particularly at a time when change looms, sometimes threateningly). You have been helpful in getting us the data, filling in the questionnaire, answering our questions about coding and other details have carefully checked the analysis and provided commentary on your service. Thank you for all of that and we sincerely hope that the results are recognisable and useful to you in improving your service still further.
Findings

The data from the benchmark has been presented in an innovative way that will, we hope, support participants in understanding the comparisons and identifying where there is scope for improvement. The Tableau workbook allows the user to interact with the data so that they can compare the service in their own PCT with others that are similar by selection of a number of ‘explanatory factors’. These explanatory factors include deprivation, GPs per 100,000 groupings of PCTs based on rurality and based on similarity as well as by the NHS Organisation structure. This can be accessed through a browser at www.primarycarefoundation.co.uk/benchmark

The key messages from the benchmark are:

- That the overall performance of out of hours services is improving - we describe how the performance on the most difficult measure of time to definitive clinical assessment is rising
- That access to out of hours service is easy and rapid. The access standards are demanding and they result in a very rapid response to a patient - with the median time to contact a patient when ringing them back being just over 15 minutes and the time to being seen face to face being within seven minutes of the later of their appointment or arrival time. This rapid response is reflected in the GP patient survey where, although the questions are slightly different, the indications are that across England a greater proportion of patients find it easy to contact the out of hours GP service by telephone than to contact their GP practice.
- It seems that the variations in cost are less than they were - with some services that perhaps had too little financial headroom edging their prices up and some of the more expensive services now being cheaper. Cost also reflects the reality of the variation between services - it appears to be driven by the age of the population, the level of demand and the geography which makes it more expensive to deliver out of hours care in a rural setting compared with an urban one
- That there are some clear relationships that emerge from the data some of which might be expected, such as demand being higher where more of the population is elderly and that where patients feel that care is delivered in a timely fashion they also rate the quality highly. However the data also refutes the assumption, made by some, that patients will prefer a service with a higher proportion of face to face consultations - the service with the highest proportion of cases where patients are given advice over the phone is rated very highly by patients.
- It is disappointing that there is no consistency in the reporting of telephony response times. This has been a long-standing quality requirement but we hope that, with the move to 111 fronting the service, this issue will be resolved.
- The variation on some measures is wider than seems to be explicable looking at the differences between the PCT populations, geography, service specification etc. and accepting that there will always be some random variation. Although services are improving some appear to have more catching up to do - with response times in a few services being comparatively slow.
- We were surprised that when we asked commissioners about whether the performance of the out of hours service was routinely reported to the board of the PCT and when we asked about unfilled hours being reported to the PCT not all services were following these recommendations from two reports following investigations into Take Care Now and the death of David Gray. We are pleased that in many of the commentaries it is clear either that gaps in the rota has never been an issue and/or that unfilled hours are now reported to the commissioner
• There are still weaknesses in coding (not an uncommon problem in the NHS) such that it is impossible to be confident that comparisons of the numbers of patients referred or choosing to go to hospital or 999 cannot be compared. Because of the variation in quality of this data we have chosen not to include this comparison in the Tableau workbook.

There are perhaps three other points that are particularly relevant to commissioners of these services that cannot be easily described in a chart or diagram but that emerged from working through the detail of the answers provided and the analysis of the data. There is also one detailed points particularly relevant to 111 and telephone health advice that commissioners might also like to consider that we have included here.

• Whilst some commissioners look in detail at the service and understand the picture 'in the round' others seldom get beyond a narrow look at the national quality requirements. We urge all commissioners to adopt a programmed approach to reviewing the service so that during the year the commissioner looks with the provider in depth at each of the following:
  - patient perception (using both the GPPS as well as any locally gathered information)
  - audits of clinicians, call handlers and, for example of how life-threatening conditions or common conditions are handled and what lessons have been learned
  - how well the out of hours service is integrated with others around the patient and how this might be improved (this is an area where input from the commissioner is important as it can often add weight and impetus to an initiative that by definition cannot be progressed by the out of hours service alone). The nitty-gritty detail of making this work with a range of services will be crucial to the success of 111
  - Reviewing activity levels, outcomes, staff effectiveness, the adequacy of staffing levels, activity levels at particular centres and considering if there is an opportunity to make changes to provide a more cost-effective service
  - Reviewing performance against the quality requirements - this too is important, but it seems to us to be no more important than many of the above
  - and finally we believe that lessons can be learnt from the benchmark (which we anticipate will change as we move towards 111 and we are also looking with the College of Emergency Medicine at comparing similar measures for A&E/UCCs to provide a fuller picture of the operation of the urgent care system)

To do each of these topics justice requires preparation and the discussion needs to be in some detail. We believe that this is best done by focusing each review meeting on a particular aspect of the service so that each is covered at an appropriate frequency. The issue is one that was highlighted by CQC in their report into Take Care Now when they said that "PCTs had a limited understanding of the service provided on their behalf" and it is something that they highlighted as potentially a national problem

• To understand and consider the effectiveness of any service that is provided to patients it is necessary to compile information that looks across the whole of a each episode of care. This is difficult to do when there is only one provider and all of the data is on one system. It is much more difficult, but not impossible (albeit with very careful management), to do when more than one provider or system is involved. With the way that most 111 contracts are being put out to tender it is highly likely that this split will exist at least in part of each area. Whilst the commissioner can pass the problem over to one provider as a lead contractor working with others this does not eliminate the problem. We recommend that those commissioning 111/OOH services look at how this will be done as a key part of evaluating bids.
• If commissioners are to understand the services that they are commissioning then they need good quality coding and recording so that they can understand more about the mix of presenting conditions, the referral to other services and dispositions, the nature and timeliness of any patient contact. Out of hours services are a data rich area where much information can be analysed to give a good understanding of the service that is procured - but even here there is more that the individuals in some services could do. We intend to do some further analysis of the most common clinical conditions and prescribing levels - but fear that at least in the first of these consistent coding is not carried out so that the information will be difficult to interpret for some providers. Commissioners might like to look at the specification and requirements that they agree with providers if they feel that this is important information that could and should be collected

• With the advent of 111 we strongly recommend that commissioners look to change the measures of the speed of response. The present measure of time to definitive clinical assessment from the time that the call is answered by a person to the start of the definitive clinical assessment becomes meaningless once the call-handler is empowered to carry out a clinical assessment and when calls are sometimes 'warm transferred' to a clinician (because the two times are simultaneous). What is much more meaningful is to measure to the end of the definitive clinical assessment as this is the time at which the patient is "clear of the outcome, including (where appropriate) the timescale within which further action will be taken and the location of any face-to-face consultation". To allow participants to get a feel for this measure we have included the median time to the end of the definitive assessment in one of the views in Tableau.

---

1 This wording comes from the existing definition of the call to be counted as the "definitive clinical assessment", the change proposed (in addition to the change to allow calls using suitable decision support software and trained non-clinical staff to be counted as equivalent to an assessment by a clinician) is to suggest that the meaningful point to compare services is the end of the call as it is not till that point that the patient knows what is to happen next or has had the advice fully explained to them.
The detailed comparisons

Overview

The benchmark covers the majority of the PCTs in England. A list of those that are participating is included in the Tableau packaged workbook together with some brief notes on those services where there are limitations on what can be compared. We also make use of information that is publicly available - in particular the GP Patient Survey data. Here we have chosen to include all Primary Care Organisations because we feel that users will wish to make comparison with the wider group. We have also used publicly available data to provide a number of possible 'explanatory factors' such as population density, deprivation, ONS groups etc. and we have included these in a number of the illustrations as part of this analysis.

Figure 1 - participants in the benchmark

The illustrations in the report are drawn from the Tableau packaged workbook that allows any user to manipulate the data and we encourage those who are interested to investigate this. The tool provides some fascinating insights and we believe presents the information in a powerful way to allow the user to explore and understand it (despite using summarised information to prevent access to individual records). Tableau is a very sophisticated user-friendly tool that can handle very large datasets and we recommend it to those involved in presenting and understanding data across the NHS.
The report is structured around five themes and for each one it highlights the main findings and tests some key/critical relationships. The themes are:

- Overall demand
- Disposition and outcomes
- Performance and responsiveness
- Cost
- Patient Perception

In compiling this information we are grateful for the support of providers and commissioners who have helped us to make sure that we get it right. Good providers recognise that they can distinguish their service from others through this type of detailed comparison. Commissioners value it too for the insight that it gives them into the service. We have been given a privileged insight and have been trusted us to feed back information in a way that makes sense. We have worked hard to respect and justify this trust. There will be pointers to particular areas that individual services need to improve but the data also illustrates just how responsive out of hours services are to patients - responding to a patient’s telephone call by starting a telephone assessment in 15 minutes and seeing patients within 7 minutes of arrival or their appointment time (both averages of the median values for all services for which data is available). Participants should certainly be looking to see how they might improve care for patients but they should also be congratulated on opening up their services to such detailed scrutiny.

Wherever we have concerns about the quality or comparability of the data we have included a caveat and have omitted data if it cannot be reasonably compared. Despite the thoroughness of our approach and even though the draft reports and analysis were fed back to commissioners and providers locally for validation and comment it is possible that some errors or inconsistency will emerge as providers and commissioners look further at specific aspects. Where this happens we will want to improve our process for the future and will be happy to acknowledge any error whether it is ours or arises from a misunderstanding in completing the return.

Finally, in this report we have highlighted a small number of services by name. Those we have chosen are Doncaster and the PCTs served by Devon Doctors or its associate organisation Halton Health. We think that the contrast and comparison is interesting - but it also helps to make it clearer by focusing attention on specific examples.

**Overall Demand**

The overall level of demand will be affected by a range of factors, some of which are discussed below, but also by what is included as a ‘normal’ or ‘core’ out of hours case. All out of hours services receive a small proportion of ‘walk-in’ cases and in some PCTs the provider is commissioned to provide a walk-in service alongside the out of hours service. In others a different provider may provide a walk in service and in some PCTs there is no convenient walk in service available. To allow us to make a like for like comparison of the volume of cases we have chosen to look here only at those patients whom we identified as phoning the service and have excluded the walk-in cases from the analysis in this section. Elsewhere (for example in looking at the cost per case) we have, of course, included the walk in cases if they are covered in the contract cost.
There is more than four-fold variation in the volume of cases per 1,000 population which ranges from around 50 cases per year up to over 200. Figure 2 is a map that illustrates this variation. It appears that there is a pattern of lower usage in more densely populated areas that could, for example, be connected to such factors as:

- A younger population in city centres compared with many rural areas
- The availability of a range of alternative services such as A&E departments, Walk in Centres etc. in city centres
- A preference among those in more rural areas to ring their surgery and then to contact the out of hours service compared to cities where there may be more unregistered patients or a less strong relationship with their GP practice
- Some PCTs being 'under-doctored' compared with others, as this might increase demand out of hours
- Differences in the service model that is specified (for example the volume of cases in Leicestershire County and Rutland appears to be much lower than that in other shire counties at less than 90 contacts a year per 1000 patients - but this could be because of the way that they have encouraged users to attend as walk-in patients rather than telephone first).

We strongly recommend exploring the data using the Tableau packaged workbook. We have designed the tool so that a range of possible 'explanatory factors' can be used to colour the graphs. These 'explanatory factors' allow the user to assess whether there is a relationship between a measure and that factor - as in the above example when it is clear from examination of the chart that London PCTs (green) typically have a smaller proportion of population over 65 than elsewhere in the country and also tend to have a lower usage of out of hours services when measured as cases per 1000 of population.

The range of explanatory factors already built into Tableau include:

- Deprivation (measured using IMD)
- NHS organisation structure
- Strategic Health Authorities
- PCT Clusters (though this has been a moving picture)
- Categories defined by the Office of National Statistics
- 3 groups from Urban to Rural
- 6 groups from Major Urban to 80% or more Rural
- ONS supergroups (7 groups of PCTs identified as similar to each other)
- ONS Groups (12 groups of PCTs identified as similar to each other)
- Population density (measured in population per hectare of the PCT)
- The type of provider (Commercial, Mutual or NHS plus a 'Mixed' category when the PCT is served by providers of different types)

Should users like to suggest other categories (for which data must be available across all Health areas) we would be delighted to look at widening the range of 'explanatory factors'.

---

2 We have normalised populations for the PCT back to the figures for the registered population of the PCT reconciled to the ONS population estimates - see technical notes
Figure 2 showing the volume of cases a year per 1000 patients

**Sheet 6**

Click on a dot to filter by Strategic Health Authority (to allow you to compare yourself with others locally) and highlight PC

Doncaster has 212 cases a year per 1000 population

Torbay has 200 cases a year per 1000 population
Figure 3 shows demand (estimated annual number of cases per 1000 population) against the proportion of the population over 65 for the PCTs in the benchmark. This graph is taken directly from the Tableau workbook.

**Demand for out of hours services by PCT compared with the proportion over 65**

There appears to be a relationship between those PCTs with a higher proportion of patients over 65 and higher demand. This is unsurprising given the evidence of the additional primary care resource used by the elderly (the Carr-Hill formula indicates that those over 80 will use many times more primary care resource than a 14 year-old boy). Colouring the PCTs by their strategic health authority (one of the 'explanatory factors') in the diagram above highlights that relatively few of those in London (green) are over 65.
Disposition and outcomes

Towards Hospital or the Ambulance Service

The proportion of patients that appear to go to secondary care or the ambulance service (both those that choose themselves and those that are referred) is a figure of interest to commissioners. Regrettably the information is difficult to compare as it is dependent on how well clinicians and call-handlers code outcomes.

We believe that the percentage going towards hospital or 999 is typically is 12 to 16%, higher than a many people may expect. Because of the inconsistencies in coding we have chosen not to include the names of PCTs on this graph (though this information has been made available locally) on the basis that we know that the reliability of recording is such that many of the figures cannot be sensibly compared.

Our message to commissioners and providers on this measure is a simple one. If you believe that there are opportunities for the out of hours service to reduce attendance at A&E then the first thing that you should do is ensure reliable recording of such cases. Then (and only when the count is near complete does it seem sensible to start) you can begin to audit such cases and assess how many fewer might have attended A&E if (for example):

- The call-handler had been clearer about when the doctor would ring back
- The caller had been given an opportunity to come straight to the centre and see a doctor
- The guidelines or algorithms given to the call-handler had been different
- The operating model had allowed warm-transfer to a clinician in cases of doubt
- The clinician had been quicker in getting back to the caller
- The telephone consultation had been carried out in a different way
- An appointment at the base had been offered sooner
Figure 6 below illustrates the apparent variation in percentage of patients going towards hospital. The average is over 11% but it is impossible to know what variation comes from poor coding and how much reflects real differences between services.
Percentage advice, base or home visits

There is a wide variation in the dispositions to telephone advice, attendance at a primary care base and home visits as shown in the illustration below which shows the services sorted by the percentage where the out of hours case ends with telephone advice. This latter category includes self-treatment at home, advice to wait and see how the condition develops or to see their own GP on the next working day and any advice given over the phone to contact another service (including 999).

Of particular interest in Figure 7 is Doncaster, the service which is highest in telephone advice at over 75% (and with a very low level of home visits at less than 1.5%). This is a service that, as is described later, is very well thought of by patients.

Commissioners and providers should remember that within any one service there will be much greater variation between individual clinicians as we are, for each service, looking at an average from many individual clinicians. It is for this reason that we stress the importance of strong processes to feed back to each individual information about their productivity, prescribing, dispositions, referrals, disposition and treatment of common or critical conditions. Sharing this information allows doctors and nurses to reflect on and develop their clinical practice in a way that is difficult in many other environments. But it also allows the service to improve consistency so that the performance and outcomes do not vary depending on which staff are working in a way that makes it impossible to meet performance targets.

Whilst the introduction of 111 might be expected to reduce variability in both the percentage telephone advice and the percentage going towards hospital because of the use of a consistent computer decision support tool. But it is the way that the tool is used, how questions are asked and information elicited that will influence dispositions and, without monitoring and sharing of good practice there may still be considerable variation between individuals.

It also seems inconceivable that the proportion given telephone advice and going towards hospital will not be higher in many PCTs with the introduction of 111. Whilst telephone assessment has limitations, assessment by a good clinician positions him or her to make a finer judgment taking in more of the circumstantial information than can be done by a non-clinical call-handler however good their decision support tool.

On the next page:

Figure 7. PCTs sorted in order of increasing percentage home visits showing the wide variation in dispositions from assessment. The colours indicate deprivation, splitting the PCTs into five bands.
Plymouth has 52% advice, 13% home visits and 35% base.

Doncaster has 76% advice, <3% home visits and 22% base.

Portsmouth has 52% advice, 18% home visits and 30% base.
Percentage urgent

Another disposition in the process is priority. Of particular relevance in looking at time to clinical assessment is the percentage of cases identified as urgent by the call-handlers. Figure 8 shows this - and the colours indicate the operating model that was reported as used by call-handlers. It appears that use of a computerised decision support tool (purple) is associated with higher priority compared with those using aide memoires (red) or detailed protocols supported by training (green).

Whilst some of the very low levels of urgent are clearly a worry lest any priority cases are missed the importance of prioritisation is lower the faster that cases are assessed. Clearly there are concerns lest the use of a less structured process such as an aide-memoire might mean that potentially urgent cases are missed. Equally there are dangers from introducing computer decision support tools without careful staff training to avoid the risk that the system is used in a way that shapes the response of the caller so as to increase priority.

On the next page:

Figure 8. PCTs sorted in order of decreasing percentage urgent on receipt showing the percentage urgent where definitive assessment starts n 20 minutes and, for less urgent in 60 minutes. The colours indicate the process used to assess priority by call-handlers.
In Doncaster their focus is on assessing all patients as early as possible and calls are only prioritised overnight. In 90% of cases (of all priority) the first attempt to assessment is made in 20 minutes. The equivalent figures to those above are 100% in 20 minutes for the small number of urgent cases and 99% for the less urgent in 60 minutes.
**Performance and responsiveness**

Patients calling and being seen by out of hours services enjoy a rapid response. Looking at the median (50th percentile) in the average service patients are rung back by a clinician within just over 15 minutes (it would be less if we counted to the first attempt to ring the patient, as sometimes the patient cannot be contacted) and they are seen by a clinician face to face in the centre within an average of just 7 minutes of the later of their appointment time or arrival time. This is much faster than in many GP practices and reflects the importance that the National Quality Requirements and providers place on quickly identifying the very small numbers of potentially critical cases where a speedy response can make a difference to the prognosis for the patient. These are not just the immediately life-threatening cases but also include those that will be seen and treated by the service where early clinical attention can make an important difference.

**Time to clinical assessment**

The overall performance of providers on this measure is rising - with the average service now starting the definitive (final) clinical assessment of more than 80% of potentially urgent cases in 20 minutes whereas when we carried out the second round of the benchmark the figure was only just over 60%. Note that (to ensure comparability across the different services) we are making no allowance for patient attributable delays where, through no fault of the service, it was not possible to make contact with the patient or where the patient chose not to wait for the call-back but went straight to another NHS service. For those services where the information is available, the first attempt to contact such patients is made within 20 minutes for urgent cases on 89% of occasions.

There is still appreciable variation, with Doncaster assessing 90% of ALL calls within 20 minutes, with median time to definitive assessment of less than five minutes and 30 minutes to the 95th percentile. However other services fall some way short of the national Quality Requirements, as shown in the figures below.

Figure 9 looks at the harsh measure of when definitive clinical assessment took place - so takes no account of the occasions when, through no fault of the provider, clinical assessment could not take place within the target time-frame (for example because the patient had chosen to go to A&E and so the phone was not answered till much later when the patient return home). PCTs are sorted in descending order of the percentage of urgent cases assessed in 20 minutes.

On the next page:

Figure 9 - time to definitive clinical assessment (a harsh measure without allowing for patient attributable delay) sorted in descending order of % urgent started in 20 minutes. The colours indicate the skill mix of the staff carrying out the telephone assessment as reported to us. Note that this is the same data as figure 8, sorted and coloured differently.
Doncaster (with doctor assessment) is ranked sixth, with 94% of urgent cases starting definitive assessment in 20 minutes and 98% of less urgent in 60 minutes.

Torbay, with a small proportion of nurses) is ranked a little below the middle with 85% of urgent cases starting definitive assessment in 20 minutes and 91% of less urgent in 60 minutes.
Unfortunately in this round none of the services using NHS Pathways or piloting 111 were in a position to provide data so we are unable to compare the performance of any of these pilots. The existing standard for time to definitive assessment becomes nonsensical once an assessment by the call-handler using a suitable decision support tool is counted as a clinical assessment because the start of the first call will in most cases be the start of the assessment (especially with the aspiration that calls should be ‘warm transferred’ if it is necessary to speak to a clinician). A much more discriminating and useful measurement is to measure to the end of the call - the point at which (using the words from the existing standard) the patient or carer is clear about “the outcome, including (where appropriate) the timescale within which further action will be taken and the location of any face-to-face consultation”.

In line with the likely future direction of all such quality definitions we strongly recommend that providers are asked by commissioners of 111 services to report to the median and 95th percentile on the time to the end of the definitive clinical assessment. It will be interesting to compare the performance of 111 services against out of hours services on this measure but one would expect the 111 service to be quicker because of the avoidance of the delay associated with ring-back. Doncaster may be an exception to this as the average time to the start of the telephone assessment is rather shorter than the median length of the initial call reported from the 111 pilots.

For the present we have provided a map showing the variation in the median (50th percentile) time to the start of definitive clinical assessment and use of the tool tip will allow users to look at the value (and the 95th percentile) for this measure. Out of interest - and in view of the recommendation above the tool tip in Tableau also provides the median time to the end of definitive assessment.

**Time to the face to face consultation**

The National Quality Requirement recognises three priorities - emergency to be seen in one hour, urgent to be seen in two hours and less urgent to be seen in six hours (all from the end of the definitive assessment). The graph below shows PCTs sorted in order of the percentage of urgent cases seen in 2 hours - but with the percentage of emergency cases in 1 hour on the left and percentage of less urgent in 6 hours to the right. It is immediately obvious that some services identified no cases as emergency to be seen in 1 hour during the four weeks taken as sample data. Although all services will undoubtedly have identified immediately life-threatening cases that will have been passed to 999 we worry lest the avoidance of the emergency category means that cases requiring a face to face consultation as soon as possible (for example the palliative care patient needing pain relief) are not seen as early as possible.

On the next page:

Figure 10 - time to face to face assessment from the end of the definitive clinical assessment - colours represent a scale from predominantly urban to predominantly rural
Devon sees 96% of urgent in 2 hours, 97% of less urgent in 6 hours and all of the emergency cases in 1 hour.

Torbay sees 98% of urgent in 2 hours, 98% of less urgent in 6 hours but no emergency cases were identified.

Doncaster sees 92% of urgent in an hour, 100% of less urgent in 6 hours but no emergency cases were identified.
Telephony

The graph below shows that in general patients find little difficulty in getting through by phone with over 70% of patients from almost all PCT areas reporting that it was very or fairly easy to get through by phone. This is reflected in the much better performance of out of hours services than general practice on this measure. 79% of respondents report it as being easy to get through to the out of hours services across all England compared with 69% to a similar question about ease of getting through by phone to their general practice (Figures from the GP Patient Survey).

This is still an area where some services have been unable to provided data about the numbers of abandoned calls and the performance in answering the calls within 60 seconds of the end of the message. It is also apparent that there is inconsistency in what is counted and which figure is used as the denominator in calculating the percentage. We very much hope that during the transition to 111 these issues are resolved but meanwhile the most useful comparative measure is probably the response to the GP Patient Survey.

On the next page

Figure 11 - telephony - perception of ease of access by phone compared with % answered in 60s
In Nottingham 72% of respondents indicate that it was easy or very easy to get through by phone and the service reports that 100% of calls were answered within 60 seconds of the end of the message.

In Doncaster 88% of respondents indicated that it was easy or fairly easy to get through by phone and the service reported that 96% of cases were answered within 60 seconds of the end of the message.
**Reporting and integration**

The tragic and unnecessary death of David Gray and Penny Campbell resulted in a number of recommendations from the various enquiries and reports many suggesting closer monitoring by PCTs and SHAs. Amongst these was a suggestion that safety in out of hours be a standing item on the PCT Board agenda for at least six months and that PCTs should monitor any failure to staff the rota to the planned level. We asked about both of these items and the map below indicates how many reported that this was done.

Figure 12 - answers to questions about reporting to the PCT board and about gaps in rotas

Although a number of PCTs reported that safety and performance of the out of hours service was not regularly on the board agenda it should be remembered that we were asking the question during a period of structural change. Equally it is clear from some of the feedback that some of those providers that did not regularly report unfilled hours compared with the planned rota to the PCT have no difficulty in filling all of their shifts so had not felt it necessary to report whilst others have started such a report since the time that we asked the question.

In addition we asked a number of questions about the recruitment of clinicians, the responses indicate that references are sought and followed up in the vast majority of cases - though a small number of providers made clear that they did not always follow up references when the candidate is already known to the service (for example because they have worked within it as a trainee) as they rely on the more detailed knowledge that they and the GP trainer have already gained.

Doncaster responded Yes, information is reported about OOH performance to the PCT board but No, unfilled hours are not reported to the PCT

Devon, Plymouth and Torbay responded Yes, information is reported about OOH performance to the PCT board and Yes, unfilled hours are reported to the PCT
One of the related issues about reporting is that too many contract review meetings rarely get beyond looking at performance against the National Quality Requirements. We have long promoted the view that providers should be looking at a much wider range of indicators and audits.

In our joint report into commissioning urgent care 'Breaking the mould without breaking the system' we suggested that a range of measures might be appropriate and that, in addition to routine reporting and occasional unannounced visits to the service, contract management might be carried out by planning meetings through the year to cover the full spectrum of issues. Such a programme is illustrated in the slide below:

**Monthly review plan for a year (with possible variation) but, say:**

- Three times a year. Review of NQRs (reported each month). Below par performance to include action plan and projection for improved performance
- Three times per year. Look at their audits of:
  - Individual performance - clinical and call-handler
  - Referrals (including to A&E/999 – spotted as early as possible and whether appropriate)
  - One other agreed topic but perhaps looking at specific common conditions or prescribing
- Twice a year. Patient perceptions (GPPS and own findings) - Trends, links with performance, actions taken and planned
- Once a year. Staff effectiveness – to consider productivity, rota adequacy, training, recruitment process
- Once a year. Working with others: opportunities for integration and improved care – Hospitals, community services, GP practices (one covers winter/Christmas planning)
- Once a year. Analysis of comparative performance from benchmark
- Once a year. Future direction (and formal review/feed-back) – commissioner to take the lead

If more than one provider then lead provider to ensure that they work as one and this covers the entire process

A number of providers too have commented on how narrowly the service is looked at by commissioners and, in an attempt to promote discussion about this topic we invited providers to tell us how often they reported 11 topics to the PCT and how frequently 11 topics were formally reviewed with the commissioner - including, for example considerations about more strategic questions over how to integrate services better within an area, the focus of attention from the commissioner to drive up health in the locality etc.

The picture below highlights the variation between PCTs in the frequency of reporting (the darker the blue the more frequently items are reported) and the frequency of formal review (the darker the orange, the more frequently items are formally reviewed). It shows the differences - but does not suggest what is good. We recommend that commissioners and providers jointly consider how they can best look at the various aspects of the service at an appropriate frequency - with the slide above perhaps providing a starting point for the discussion.
Figure 13 Showing the variation in the number and frequency of items reported to PCTs and formally reviewed between commissioner and provider.
Another area where commissioners can support providers is in helping them to integrate services with providers of other services. This is crucially important - we once came across two out of hours services (nearby to each other, though not in England) with very different arrangements for handing over responsibility to specialist services those suffering from a mental health crisis. In one there was an agreement that, subject to agreed criteria, the out of hours service would be able task the mental health team to take responsibility for care after a telephone assessment a doctor. In another area this process was far from easy and in one instance the doctor after seeing the patient face to face had spent many hours making over 40 telephone calls to arrange care for a teenage girl in crisis.

We therefore asked providers both about the doctors ability to access details about the patient (from the National Summary Care record, the GP record and special notes about vulnerable patients) and about whether processes were in place to "allow a speedy and efficient transfer of responsibility for a patient with all of the relevant information and without negotiation/discussion on each case". Our hope had been that we might have been able to produce a composite score to highlight PCTs where there were good examples of the integration of out of hours services with others to improve care to patients. Unfortunately the answers were insufficiently consistent for such a comparison to be made - but we nevertheless suggest that the questions that were asked may provide a good check-list for commissioners and providers looking to address this issue.
Cost

Calculating the cost per head of population and the cost per case of an out of hours service is not a trivial task. As with other aspects of the data collection, we take a considerable amount of trouble to understand the scope of the contract and what, if any costs are and are not included in the headline cost figure. With help from PCT finance staff as well as from those with the understanding of the contract we arrive at a cost for the 'core out of hours service'. This sometimes requires considered estimates of PCT on-costs to arrive at the equivalent of an 'arms length contract arrangement' or adjustments to recognise that the contract includes other services - some of which, such as the delivery of a 365 day a year walk in service, can be significant. Throughout the process (with emails and phone calls) we have made sure that both provider and commissioner understand the basis of the calculations so that we arrive at the best possible estimate of the cost per head and per case. We also adjust the reported populations to normalise them back to the ONS figures for the PCT reconciled to PCT registered populations to recognise that some practice boundaries draw patients who live in one PCT into another PCT.

At the end of this process (and after final validation with commissioners and providers) we believe that we have the best possible comparator of cost per head and cost per case. However the reader should be aware that sensible estimates have sometimes had to be made and the figures are certainly not consistently accurate to the penny. But the range and the spread is such that the results do tell a story.

We have also tried to highlight and explore some of the factors that must increase cost, all other things being equal. The most obvious of these being geography where travelling time for home visits and the number of centres opened will be much greater than in urban areas. For example in Devon 14 centres are available to see patients face to face (including those in Torquay and Plymouth which will also be used by Devon patients) compared with only one centre in Doncaster. In the diagram below the colours distinguish the predominantly rural PCTs (green) from the significantly rural (orange) and predominantly urban (blue). It is quite clear that there is a preponderance of green bars among the more costly services and this relationship is also reflected in the scatter plot of cost per head against population density (among a number of other explanatory factors) further below.

On next page:

Figure 14 cost per head and per case by PCT sorted in descending cost per head - colours represent a scale from predominantly urban to predominantly rural
Devon, a predominantly rural area has a cost per head of £10.07 and a cost per case of £60.50

Doncaster, a predominantly urban area, has a cost per head of £8.57 and a cost per case of £39.69
Figure 15 - Confirmation of the influence of geography on cost is also shown with the larger spots on the map showing the greater cost per head in the less populated areas.

Cost per head of population

All other things being equal it would be plausible to expect that in areas where the volume of cases per 1000 of population is greater the cost would be higher too - and this seems to be confirmed by the plot below (Figure 16). This plot also indicates the relationship of case volume and cost per head to the proportion of the population over 65 through the use of colour.
Cost per head plotted against cases a year per 1000 of population by PC
It is perhaps easy to rationalise why costs will be higher in areas with a lower population density, in PCTs that have a higher number of cases per 1000 and that have a greater proportion of people over 65. However the relationship between higher cost and more survey respondents rating care as good or very good, whilst undoubtedly true (as shown in in the diagram below with p<0.0005), may be driven by a number of other factors some of which are discussed in the next section. The colours hint at one of these with major urban centres being particularly predominant to the left.

**Cost per head plotted against proportion rating care as good by PC**

![Graph showing the relationship between cost per head and proportion rating care as good or very good.](image)

*Colour coding for the factor you have chosen as important to you:*
- 1. Major Urban
- 2. Large Urban
- 3. Other Urban
- 4. Significant Rural
- 5. Rural 50
- 6. Rural 80
Patient Perception

The GP Patient Survey provides valuable information about patient perceptions of the care provided by out of hours services during the previous six months. Although many of the respondents to the survey have not recently used their out of hours service, so many questionnaires are sent out that an average of over 1000 returns for each PCT are received to the questions about out of hours GP services. Since the same questions are asked across all PCTs comparison can be made in a way that is not possible with the range of different survey approaches and questionnaires used by providers independently. We have already seen some of these findings referred to above (for example in looking at telephone access).

For the fifth benchmark we will be looking to mine the data further including the more detailed questions asked as part of the CFEP UK survey that most benchmarked services participated in.

The first scatter plot indicates the close relationship in the answers to the survey between those who believe that they were seen ‘about right’ (as opposed to too slow) and the rating of care. The message is simple - if respondents feel that they have been attended to promptly they rate the care more highly.

Figure 16  GP Patient Survey results for rating of out of hours care by PCT

76% of respondents in Torbay rated care as good or very good and 76% said that it was provided ‘about right’

74% of respondents in Doncaster rated care as good or very good and 81% said that it was provided 'about right'

The next figure puts the same information about the rating of care on a map.
A clear pattern is visible that patients in London rate care less highly than in other areas generally - it is visible on the map and visible to the eye looking at the green spots in the scatter plot for PCTs in the London SHA. But the reader should beware of assuming that services must therefore be worse in London. We have already highlighted the different age profile of patients in the London area, the lower demand per head and analysis of the GP Patient Survey that we have done in the past showed a markedly lower response rate in London compared with other areas. In addition to these factors Londoners frequently rate services less highly than those from other areas with market researchers suggesting that this is because they are more time hungry, less patient and tolerant of waiting times, pointing out that they often have more choice and higher expectations than those from elsewhere as well as highlighting the different socio-economic and demographic profile.

In exploring the relationship between patient perception we plotted with Tableau the relationship between various measures from the GP Patient survey against a number of measures of response as shown in Figure 19.
Summary

Our objective in this report was to give some insight into the issues that affect out of hours services so as to put the individual report about your service into context. We also aimed to inspire you to open the Tableau packaged workbook and explore the data for yourself. Above all we hope that we have made clear how it is important to look at a range of measures in assessing the performance of any health service - whilst identifying some of the sound reasons for there being a difference between many of the services because of the geography, the population, the range of alternative services available and more.

It is not only external factors that can influence the performance and outcomes - the specification for the service, its' ethos and history can also make a big difference and it is clear that whilst superficially one out of hours service can look similar to another they can actually operate in a way that produces very different results - indeed there is more than one way of working that seems to produce a service that patients value and appreciate.

Above all it is the quality of the individual consultation that takes place that matters to a patient and it is the culmination of a host of individual decisions made by clinicians and others that results in the overall performance profile that we have tried to compare. Crucial to this is the quality of management and management processes in providing feedback to clinical staff on how they compare to others to foster a process of learning and sharing that will lead to a reduction in variability and to consistently good prompt care being provided to patients.
Technical notes and considerations when looking at the analysis

Population

There are various reasons for adjusting the reported population figures for use in the OOH benchmark - for example respondents will have obtained data from a variety of sources or times that are not always comparable and because registered populations which are based on practice lists will inevitably contain a number of patients who have moved from the area whose details have yet to be removed from the record. This means that the population figure submitted by one PCT is not necessarily comparable to that submitted by another.

There are a number of possible sources for comparable information all of which have been derived in different ways. They are described below. At present Tableau is set up to allow comparison using the first three (reported population, ONS mid year 2010 and ADS GP registered for April 2010 constrained to 2009 mid-year estimates). With advice, we recommend use of the ADS GP registered population constrained to the ONS mid year estimate as the best for comparison across PCTs.

Population as reported to us - where there is more than one provider in an area we have used this to allocate the other numbers pro-rata. Generally (though interestingly not in the example below) this is higher than the estimates from the ONS - presumably because the 'cleansing process' lags behind so that registered populations often include some 'ghost' patients who have moved away or died, but not been cleaned from the PCT record.

ONS mid-year population estimates - by PCT for mid 2010 (published Sept 2011). This has the advantage of being the best available estimate of the number of people resident in the PCT area and is the most up to date of the comparable measures and relates to the period that we are looking at in the fourth round of the benchmark. It will include all patients whether they are registered or not, but will exclude patients that may be registered with practices that are part of the PCT but live outside the geographic boundaries of the PCT.

ADS GP Registered Populations Constrained to Mid Year Estimates - by PCT for the April 2010 registered population constrained to 2009 mid year population estimates. This has the advantage of taking account of the 'import' or 'export' of patients that occurs because a practice within one PCT may serve some patients who fall outside the geographic boundaries. It also makes adjustments for to recognise that some age groups may be more or less inclined to be registered with a GP. However it will exclude unregistered patients and the figure is adjusted to the ONS figures for 2009 mid-year estimate so is for a year earlier than the one that we are interested in. Nevertheless as this provides the latest reliable comparison across PCTs for the population that will be referred to the OOH provider if they ring their practice we have chosen to use this in our calculations.
Possible errors because of the different ways that staff use the system, the way that data is extracted or the way that we have analysed it

Nurses, doctors, call handlers and despatchers working in an out of hours environment want to do their job well - they also want to do it efficiently and in a way that allows them to focus on the patient. Many of them (particularly the doctors) work in the out of hours service only part time and have not learned to use the system fully - or sometimes they have learned 'work-arounds' that may make them more effective but can hinder reporting. This can mean that cases may not be reported or that reports will be misleading. Some illustrative examples are listed below

There are also instances when the filters used to run the data extract are such that we may be missing some cases from a particular service. In a few cases where there are a number of providers involved in a PCT there is a danger that we have misunderstood that way that the local providers serve the population of the PCT. There is also always the danger that in setting up mapping tables for the different codes in each provider we have made a mistake or that we have misunderstood that operational process. Our best protection against this is the validation with each provider/PCT and the many conversations and emails that we have had when we spotted any possible anomaly.

Some examples of possible things that might have an effect on the report are listed below:

- Clinicians may adopt a different approach to consultations. One clinician may choose to open the case ahead of the consultation and start entering information as soon as the consultation starts whilst another may not make use of the system until the consultation has progressed some way. Since the time stamp is generally from the first key stroke this can affect the measurement of time to clinical assessment, consultation length and time to face to face consultation. It seems improbable however that this will have a significant impact on one service compared with another as all services will employ a mix of clinicians with different ways of working.

- Filtering of data extracts by the provider in generating the extract can also present a difficulty to us in carrying out the analysis. Some services have filtered on the doctors operating group which links, through the patient's practice to the PCT. Unfortunately this may mean that they have excluded patients who are unregistered or are visiting. To minimise the impact of this we have generally reported numbers of patients from the PCT that are seen by the service and excluded those that are unregistered so that we make the comparison as consistent as possible.

- Some of the measures rely on the quality of coding by clinicians and non-clinical staff working within the service. We mentioned in the main text the inconsistent coding of cases where the patient chose or was referred to A&E but we have also come across some call-handlers who will cancel cases where the patient was directed to another service but who fail to choose a cancellation reason that allows this to be identified. This will also mean that such cases are not properly identified.

- Some clinicians using Adastra 'lock' cases in a variety of different ways. This holds the case in the existing queue (say awaiting clinical assessment) until the doctor reopens the case and completes it on the system. This may be done for a number of reasons, for example.
to look down the list of cases awaiting assessment and cherry-pick cases to assess these over the next period. It will appear on the system as though the assessment has already started when actually no phone conversation has taken place

where clinicians have used the facility to allow them gather more information by speaking to the poisons unit or a hospital before returning the call to the patient to tell them what is planned thereby completing the clinical assessment

to serve as a reminder to contact the caller again, for example after reassuring the concerned mother they might have agreed to ring in an hour to check that the treatment has quietened the baby - this being more of a comfort call than a clinical assessment

In addition there are a number of other specific issues that we suspect are occurring that we will be raising during the validation process with the individual providers and PCTs. Fortunately there are few enough of any of these examples that we have identified to allow us to be confident in the results - but there are one or two cases where one service may be excluded from the comparison because of doubts about comparability

**Measurement points for time to clinical assessment**

Definitive clinical assessment is defined in commentary on the National Quality Requirements as being "an assessment carried out by an appropriately trained and experienced clinician (not a call-handler) on the telephone or face-to-face. The adjective ‘definitive’ has its normal English usage, i.e. ‘having the function of finally deciding or settling; decisive, determinative or conclusive, final’. In practice, it is the assessment which will result either in reassurance and advice, or in a face-to-face consultation (either in a centre or in the patient’s own home).

Providers have not always reported against this requirement in the same way. Some have ‘started the clock’ at the end of the initial call from the patient (instead of at the point of "the call being answered by a person") whilst others have stopped the clock on the first occasion that the patient is spoken to by a clinician (not the definitive, final decisive phone consultation that the standard refers to).

By working with an extract of all cases for four sample weeks we are able to ensure comparability (although for a small number of providers we cannot compare them for reasons summarised below). However the reader should recognise that the standard is not clear what should be done about cases where through no fault of the service it has been unable to contact the patient. This often happens when a patient goes to A&E and is not there to answer the phone or because they are busy on the phone when the service tries to ring. We have thus tried to compare:
Time to the definitive (final) clinical assessment, time to the first assessment by a clinician and time to the first attempt to carry out an assessment. The diagram below shows the differences between these measures.

Commissioner should note that:

Conscientious clinicians working will make sure that they do not close cases too soon - indeed in some cases they will try a considerable number of times to contact the patient and will consider whether the indications are that they should arrange a home visit (with police support to break in if necessary) when the symptoms noted by the call-handler indicate that this might be appropriate. This will mean that there are a more cases that are eventually assessed very late (for example when the patient gets back home having made their own way to A&E) compared with a service that chose to focus only on performance against the NQRs. Fortunately clinicians hold sufficient sway that the right thing is done for patients - but it will adversely affect the apparent performance.

Some services make sure that clinicians use the Adastra capability to record each attempt to carry out a phone assessment. In these cases it is possible to identify (and the Adastra report can be run to take account of the issue) that one or more attempts was made to contact the patient within the 20 or 60 minute target time. Unfortunately in other services clinicians either do not record this attempt so it is impossible to make a like for like comparison (which is why we focus on identifying the definitive assessment for comparative purposes although it is clearly much more difficult to achieve the target by counting as a failure those occasions that are no fault of the provider).

Occasionally we have spotted some clinicians who, when they fail to get through on the phone, open the case and record the failed contact as if the consultation took place. This scuppers any chance of accurate reporting and when we do spot it we have notified the provider service.
This slide shows different measurement points for time to assessment

Green arrow is time to definitive assessment, blue to first assessment and orange to the first attempt (red is the time to face to face consultation)

Standard process for a base visit

Standard process for a base visit with two assessments (say by a nurse then a doctor)

Failed contact – usually excluded as patient attributable delay

<table>
<thead>
<tr>
<th>Key</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial call or contact</td>
</tr>
<tr>
<td>Telephone assessment</td>
</tr>
<tr>
<td>Failed contact</td>
</tr>
<tr>
<td>Face to face</td>
</tr>
</tbody>
</table>

© Primary Care Foundation