



Policy Briefing: The Friends and Family Test

August 2014

Picker Institute Europe



picker
Institute Europe

Picker Institute Europe

Picker Institute Europe is a not-for-profit charity dedicated to ensuring the highest quality health and social care for all, always. We:

- Influence policy and practice so that health and social care systems are always centred around people's needs and preferences
- Inspire the delivery of the highest quality care, developing tools and services which enable all experiences to be better understood
- Empower those working in health and social care to improve experiences by effectively measuring, and acting upon, people's feedback

© Picker Institute Europe 2014

Published by and available from:

Picker Institute Europe
Buxton Court
3 West Way
Oxford, OX2 0JB
England

Tel: 01865 208100

Fax: 01865 208101

Email: Info@PickerEurope.ac.uk

Website: www.pickereurope.org

Registered Charity No 1081688

Company Limited by Registered Guarantee No 3908160

Picker Institute Europe has UKAS accredited certification for ISO20252: 2012 (GB08/74322) and ISO27001:2005 (GB10/80275).
Picker Institute Europe is registered under the Data Protection Act 1998 (Z4942556)

Friends and Family Test: Policy Briefing

The publication of NHS England's eagerly anticipated review (NHS England, 2014b) of the friends and family test (FFT), and the new guidance published alongside it (NHS England, 2014c), have put the controversial test back into the spotlight again. The FFT is no stranger to scrutiny despite being only a little over a year old. In this briefing, we look at the evolution of the FFT from its roots in the Net Promoter Score and briefly trace its history. We then consider why the FFT has been associated with so much controversy, reflect on the key recommendations in the latest review, and ask whether they go far enough to address the critics' concerns.

The origins of the FFT

The FFT's lineage is well known: it is based on the Net Promoter Score (NPS) (The Strategic Projects Team, 2012). The Net Promoter Score was popularised in the United States, where it has been promoted as a measure of customer loyalty and a predictor of business growth (Reichheld, 2003, 2006). The NPS uses a single question – “how likely is it that you would recommend our company/product/ service to a friend or colleague” – and usually collects responses on an 11 point (0-10) scale. The score itself is calculated by subtracting the proportion of ‘detractors’ (those who rate their likeliness to recommend as 0-6 on an 11 point 0-10 scale) from the proportion of ‘promoters’ (those who answer 9 or 10). Typically, it is followed by another question asking people to explain why they gave the answer they did.

The FFT in the NHS was first piloted by NHS Midlands and East, the East of England's Strategic Health Authority, in 2012 – initially using the same wording as the NPS. From day one of its pilot, ambitious claims were made about its ability to revolutionise patient experience by enabling patients to provide feedback “*like never before*” (NHS Midlands and East, 2012). Specifically, the nascent FFT sought to create a “standard approach” to allow wards and hospitals to compare themselves to one another. In April 2013, the FFT was rolled-out across the NHS following recommendations made by the Nursing and Care Quality Forum (NCQF) to the Prime Minister. The NCQF's original intended uses for the FFT were threefold (Nursing and Care Quality Forum, 2012): as a communication tool to help foster a person-centred culture within the NHS; an improvement tool to drive change locally; and a comparison tool to assess quality in different organisations.

The Department of Health's translation of the NCQF's aims for the FFT expanded on this but arguably extended the aims of FFT even further – for example by linking it to pay-for-performance (NHS England, 2014a). Even before its highly publicised national launch, a seemingly ever growing weight of expectation continued to be put on the FFT to serve an evolving set of ambitious uses including:

- a comparable performance measure to:
 - provide public feedback to inform patients' choices about hospital care;
 - facilitate intra-trust performance monitoring through the collation of ward level data; and
 - produce comparable trust data upon which to base rewards for better performing trusts; and
- an improvement tool to enhance understanding about where and how improvements could be made (Department of Health & NHS Midlands and East, 2012).

These aims were summarised by Prime Minister in a high profile announcement of the national Friends and Family Test:

“We’re going to give everyone a really clear idea of where to get the best care - and drive other hospitals to raise their game.”(Prime Minister’s Office, 2012)

Following this announcement, the FFT was rolled out nationally for inpatient and A&E care from 1st April 2013. And in October 2013, the Minister for the Cabinet Office, Francis Maude, announced that from March 2015 the FFT would be extended to all NHS services, including mental health services, community nursing and outpatient appointments (Cabinet Office & Department of Health, 2013).

Controversy and criticism

The FFT’s proposed approach prompted researchers and survey methodologists to raise serious concerns over its suitability as a comparison tool even before its national launch (see, for example, Lynn, 2013; Reeves, 2012). Concerns centred around three issues: the question being asked, the way it is administered, and the use of the data.

Firstly, the wording of the FFT – and particularly the focus on the concept of ‘recommendation’ – has come under fire. A study conducted by the Picker Institute in 2012 involving testing with more than eighty patients concluded that the Net Promoter Score was “not appropriate for use in an NHS setting”(Graham & MacCormick, 2012, p14). The term ‘recommend’ was the source of difficulties: it was either misunderstood or misinterpreted by a number of respondents. Adding the suffix “...if they needed similar care or treatment” to the NPS question, as has been done in FFT, did not adequately address these issues.

Secondly, the methodology for the test has been criticised by survey researchers (see, for example, Lynn, 2013; Reeves, 2012) for failing to provide a genuinely standardised and comparable approach. As the Picker Institute has demonstrated (Sizmur, Graham, & Walsh, 2014), the combination of the varying collection methods used by trusts (for example paper based, online, and text) and different patient demographic profiles of respondents (in terms of their age and sex) has always had the potential to significantly skew the test’s results. Yet these are only two of a range of possible confounding factors inherent in the flexible approach to the implementation of the test. Providers can administer the test at different times (on discharge or within 48 hours of discharge); using self-completion methods, neutral interviewers, or front-line staff; and so on. Taken together, these factors mean that comparisons across providers are impossible.

Thirdly, a number of issues have arisen relating to the use of FFT data based on the net promoter score. These include for example the lack of understanding and dislike of the score by both patients and professionals alike (Ipsos MORI, 2012) and issues of data reliability in cases where scores have been reported based on single responses. Critically, these have had real implications for a number of hospital wards who have been labelled as ‘failing’ on the basis of poorly understood and unreliable data (Graham, 2013; Illman, 2013).

NHS England's review of the FFT

Given the FFT's bumpy journey so far, NHS England's commitment to review the test's first six months of operation is a welcome signpost along the way. The national FFT and NHS England share the same date of birth – 1st April 2013 – so the organisation inherited a brand new programme, launched at unprecedented scale, and with complex and far reaching ambitions. Clearly, this is a difficult hand to play and its approach of engaging with, and more importantly listening to, both proponents and critics of the test in order to identify what's working well and what could be improved suggests a much needed degree of pragmatism.

So we return to our opening question - does this review, and its recommendations, go far enough to address existing concerns about the FFT and hush the voices of the test's harshest critics?

In short, our answer is yes – in part. But some concerns remain. The review covers many of the most important issues surrounding the FFT and a number of the recommendations made will undoubtedly improve the test. However our view is that the review and its findings do not go far enough to optimise the FFT as a tool.

The good news

So first for the good news. The review's headline message is that the test is most suitable as a tool for service improvement. As a result it is most valuable when accompanied by free text patient responses and most successful when its implementation is supported by a culture of patient centred care and by the resources and capacity to deliver this. This recognition is reassuring and represents a clear repositioning of FFT by NHS England away from its use for comparative performance monitoring. It acknowledges, in effect, that one simple question cannot meet the multitude of requirements that had previously been placed on its shoulders. Specifically – and helpfully – the guidance also acknowledges that “the FFT does not provide results that can be used to directly compare providers” (NHS England, 2014c, p18) and removes performance comparisons from the FFT's set of aims.

We also welcome the recommendations that NHS England should continue to clarify what the FFT can, and equally importantly can't, do and on how the test should be used going forward. In particular, the clear acknowledgement in this review that FFT is a new type of feedback mechanism rather than a robust survey tool, and that as such, relative results should neither be used as a basis for performance related payments nor presented to patients and other stakeholders in a way that suggests they are fully comparable, should go at least some way to addressing previous concerns about its use as a comparability tool. However, as we discuss below, it's not clear whether it is possible to use the FFT's reported headline score in a way that *doesn't* suggest comparability – and this remains a significant concern.

The measures directed towards improving the FFT's data quality, usage and confidentiality issues are also a welcome step in the right direction towards restoring the credibility of the test. In particular the move to mandatory collection of follow-up comments, the guidance on best practice implementation and usage of the test data and the suppression of overall results where there are less than five respondents all help to address criticisms of the FFT. The mandatory collection of 'free-text' comments is particularly important: these comments appear to be the part of the FFT most likely to support improvements, and are generally positively received by NHS staff. Critically, this contributes to a shift in the FFT's position from being a data collection (first and foremost) to a tool for service improvement.

Room for improvement?

However, it's not all good news for critics of the FFT: some of the original issues are not fully addressed. Despite the acknowledgement that FFT data is not directly comparable across providers, the review does not recommend doing away with the publication of scored results altogether. Instead, it recommends replacement of the current reported FFT headline metric with one of three new metrics. These include a star rating system akin to that used by Amazon and Trip Advisor; the percentage of people likely to recommend the service; or a positive and negative metric, representing the number likely and unlikely to recommend.

The existing FFT score, as a measure which is both methodologically unsound when used in the context of the FFT, and poorly understood by patients and carers, is clearly unsuitable and we welcome the decision to stop using it. However we remain as concerned about its replacement by any of the proposed metrics as we would be by its continued use.

Given the FFT's unsuitability as a comparative tool, a move away from reporting the results of the FFT through any single over-simplified metric would perhaps have been more warmly welcomed by critics of the test. However the biggest concern arguably lies with the potential replacement of NPS by a star rating system. Such a visual presentation of the data will inevitably encourage rather than caution against direct comparisons by patients, risking adding fuel to potentially erroneous perceptions that a 'five star' provider necessarily offers better care than a 'one star' provider.

While star ratings have their own contentious history within the NHS (Mannion, Davies, & Marshall, 2005; Snelling, 2003), any form of ratings based on the FFT are potentially problematic. As discussed above, the FFT is not suitable for organisational comparisons yet it seems highly likely, if not inevitable, that this is how the public would use ratings. These concerns are ones that NHS England itself may share, as its review includes the recommendation that it continues to investigate further ways of framing FFT data to steer away from comparability. Although this is to be welcomed, in conjunction with the recommendation that FFT data is to be published alongside other patient experience measures on NHS choices, our view is that these measures are not sufficient and that there would be merit in NHS England reconsidering the desirability of publishing a single headline metric for the FFT's results going forward.

The review also does not go far enough in addressing concerns regarding the current general lack of understanding of the FFT among the public. Whilst this lack of understanding is acknowledged, in particular the confusion faced by people concerning the framing of the test and its associated concept of *recommending* hospital care when that care may not be the result of an active choice, the review stops short of considering in detail the appropriateness of the headline FFT question. The accompanying guidance does provide scope for secondary questions to be asked in cases where respondents don't understand the initial question (for example where hard to reach respondents have special needs or vulnerabilities), and potentially different scales that may be used to aid answers. However, the review does not adequately address the methodological problems created by incorporating the resulting views of all groups into a headline score where these different methodologies may be applied.

The Picker Institute wholeheartedly endorse the underpinning principle that the FFT should provide a mechanism for *all* service users to provide feedback on issues related to their care that are most important to them. However this principle creates a tricky challenge that is yet to be solved in this review. How best to reflect the views of hard to reach groups in a single headline score when data from these groups may not be directly comparable given the need to tailor the approach to collection to individual circumstances? Whilst this issue could be seen as less important if the FFT scores are not to be used for performance assessment, it remains a factor so long as the score results are collated and published.

In fact, this issue of collating and publishing data is tightly bound up with the test's purpose and its ability to achieve its aims. Collecting data is natural whenever one wishes to measure or assess: publishing it is then essential for transparency. But with the acceptance that the FFT is not a comparative measure, the continued collection of data that cannot reliably be used for assessments or choice feels anachronistic and hard to justify. A better solution would be to collect only data that has a clear use – which in this case would mean comments only.

Outstanding questions

As described above, the review marks an important step forward for the FFT – and a number of the revisions proposed will undoubtedly make it a more useful tool for quality improvement. But there remain a few outstanding points where further development would be useful.

And so we turn to three issues in particular we're left wondering about at the end of the review:

- the mixed messages about FFT's use in quality comparisons;
- the net benefits of its implementation; and
- its place in a larger suite of patient-centred care measures.

First, the review points to a useful refocussing of the FFT as a “timely feedback tool” (NHS England, 2014c, p18), useful for service improvement and best deployed alongside – not instead of – other methods of gathering information on patients' experiences. But without removing the scored element altogether, the test retains a data collection role that it struggles to shoulder. Nowhere is this tension more clearly highlighted than in the following quote:

“in order for its use for public choice to be maximised, FFT data should be presented to the public in a much clearer and understandable manner, albeit one that does not imply direct comparability of FFT data between trusts”.(NHS England, 2014b, p43)

So, we, as patients should refer to the data to inform our choice of provider, even though the data is not comparable across providers. Proponents of the FFT's use as part of a package of patient choice tools might say here that we use information selectively to make these kind of judgements every day in other aspects of our lives, for example when considering product or service reviews on Amazon or Trip Advisor, so why not in health? And there may be merit in this argument. But, equally, there is an argument to be made that our decisions about health care are not like the ones we make about hotels, holidays, or books. The stakes are so much higher; shouldn't the information we use to choose be that much more robust? This issue is clearly one where more evidence about how FFT is used by patients in practice is required before any conclusions can be drawn.

Second, we're also left to contemplate the *incremental* benefits of the FFT to date. The review acknowledges that the test works best where a culture of considering patient experience already exists and it identifies three factors which determine the effectiveness of FFT as a feedback mechanism to drive local improvements:

- the culture and infrastructure of patient experience data that existed prior to FFT;
- the leadership on implementing FFT and learning from it; and
- the data collection and processing solutions implemented.

It also finds that:

“in wards and departments that previously lacked a culture of collecting and responding to patient experience data, FFT may take longer to become embedded and used as effectively as possible”.(NHS England, 2014b, p13)

A cynical interpretation of these findings would be that FFT works best where it is needed least. This by itself in no way negates the usefulness of the test or the opportunities for poorer performers to learn from the practices of the best performers. Indeed, the fact that many frontline staff have embraced the qualitative part of the FFT and have been able to demonstrate tangible improvements from its rollout in practice is testament to the realised benefits. However, without evidence of the actual cost of the FFT’s implementation either in the limited settings of its first six months of operation or when rolled-out across all NHS settings from later this year, warning bells start to sound. Add to this the acknowledgement in the review that more external support for larger scale data collection and analysis of qualitative feedback, and potentially further investment in technology may be required, and the bells become louder. In short, we have no indication of how much further roll-out of the FFT will cost, how this will compare to the benefits of its further implementation and over what time period we can expect these costs and benefits to accrue. In the NHS’s current financial environment, this is surely another area in which further evidence is required.

Finally, we return to where we started with the weight of expectations placed on the FFT. Here the review is clear, with the recommendation that NHS England should develop a strategy, covering the full set of complementary patient experience channels and tools, which in aggregate responds to the sector’s needs for feedback but that does not rely too heavily on any one channel or test. This is a logical and balanced approach, and an easy one to endorse. But publically, at least, the FFT still carries expectations that – given all we have learnt – feel disproportionate. In a video accompanying the review, NHS England’s National Director for Patients and Information tells us he believes the wider roll-out of the FFT later this year:

“will come to be seen as one of the most important turning points in the history of the health service”.(Kelsey, 2014b)

The value of the FFT should not be underestimated, and used appropriately it should positively contribute to the NHS’s culture of patient engagement and drive local improvements in areas that are most important to patients. But for it to be credible and for its value to be accepted, there needs to be true clarity of purpose about FFT – and the expectations should be appropriate. Healthcare quality is complex and multidimensional (Donabedian, 1980); we should not expect one single measure to provide an adequate indicator of quality in the round. Instead, we should take advantage of a range of tools with different, complementary strengths and characteristics to build up a clear picture of quality, target improvements, and provide useful information to the public.

One little test can undoubtedly achieve much, but if there's one take away lesson that should be learnt from the first six months of its operation, it's that we should also be aware of its limits. Many of the successes attributed to the FFT since its introduction have been small changes: soft-closing bins to reduce noise on wards at night, trolley services for patients too ill to visit the hospital shop for a newspaper, and the addition of Marmite to breakfast menus have all been cited as examples (Kelsey, 2014a).

Not every change needs to be revolutionary for its impact to be real and worthwhile. In order to realise and celebrate the value of the FFT, we need to be proud of the little things it can do and stop asking it to do the big things that it can't.

References

- Cabinet Office, & Department of Health. (2013, October 1). *Friends and Family Test extended across public services*. Retrieved 25 October 2013, from <https://www.gov.uk/government/news/friends-and-family-test-extended-across-public-services>
- Department of Health, & NHS Midlands and East. (2012, October 4). *NHS Friends and Family Test Implementation Guidance*. Retrieved 24 October 2013, from <http://www.england.nhs.uk/wp-content/uploads/2013/07/fft-imp-guid.pdf>
- Donabedian, A. (1980). *The definition of quality and approaches to its assessment*. Ann Arbor, Mich: Health Administration Press.
- Graham, C. (2013, August 29). *The friends and family test can make the grade*. *Health Service Journal*. Retrieved 28 May 2014, from <http://www.hsj.co.uk/comment/the-friends-and-family-test-can-make-the-grade/5062422.article>
- Graham, C., & MacCormick, S. (2012). *Overarching questions for patient surveys: development report for the Care Quality Commission (CQC)*. Oxford: Picker Institute Europe. Retrieved from http://www.nhssurveys.org/Filestore/reports/Overarching_questions_for_patient_surveys_v3.pdf
- Illman, J. (2013, August 30). *Satisfaction survey 'damages staff morale and alarms patients'*. *Health Service Journal*. Retrieved from <http://www.hsj.co.uk/news/satisfaction-survey-damages-staff-morale-and-alarms-patients/5062673.article#.U9erjfldWSo>

- Ipsos MORI. (2012). *Scoring and presenting the Friends and Family Test: a review of options*. London: Ipsos MORI. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/214942/FF-T-Ipsos-Mori-research-report.pdf
- Kelsey, T. (2014a, March 26). *The Friends and Family Test - proof of success through action. NHS England*. Retrieved from <http://www.england.nhs.uk/2014/03/26/tim-kelsey-7/>
- Kelsey, T. (2014b, July 21). *Tim Kelsey - the importance of patient feedback*. Retrieved from <https://www.youtube.com/watch?v=IQrGZ0Uu8ws&list=PL6IQwMACXkj1NEbF89RTD6oF9nUDyZKwJ&index=5>
- Lynn, P. (2013, April 9). The friends and family test is unfit for purpose. *The Guardian*. Retrieved from <http://www.theguardian.com/healthcare-network/2013/apr/09/friends-family-test-unfit-for-purpose>
- Mannion, R., Davies, H., & Marshall, M. (2005). Impact of star performance ratings in English acute hospital trusts. *Journal of Health Services Research & Policy*, 10(1), 18–24.
- NHS England. (2014a). *Commissioning for quality and innovation (CQUIN). 2014/15 guidance*. Retrieved from <http://www.england.nhs.uk/wp-content/uploads/2014/02/sc-cquin-guid.pdf>
- NHS England. (2014b). *NHS England Review of the Friends and Family Test*. London: NHS England. Retrieved from <http://www.england.nhs.uk/wp-content/uploads/2014/07/fft-rev1.pdf>
- NHS England. (2014c, July). *The Friends and Family Test. July 2014. NHS England*. Retrieved from <http://www.england.nhs.uk/wp-content/uploads/2014/07/fft-imp-guid-14.pdf>
- NHS Midlands and East. (2012). *New NHS 'friends and family' test to improve patient experience and care in hospital* (Press release). England.
- Nursing and Care Quality Forum. (2012). *Promoting and achieving accountable and empowered nurse leadership across the system*. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/216812/N-CQF-letter-to-PM.pdf
- Prime Minister's Office. (2012, May 25). *Hospital 'friends and family test' announced*. Retrieved from <https://www.gov.uk/government/news/hospital-friends-and-family-test-announced>

- Reeves, R. (2012). *Why the friends and family test won't work*. *Health Service Journal*. Retrieved 30 July 2014, from <http://www.hsj.co.uk/comment/columnists/why-the-friends-and-family-test-wont-work/5052423.article>
- Reichheld, F. F. (2003). The one number you need to grow. *Harvard Business Review*, 81(12), 46–55.
- Reichheld, F. F. (2006). The ultimate question. *Harvard Business School Press, Boston, MA*. Retrieved from http://www.jihel.be/wp-content/uploads/2006/02/Ultimate_question_excerpt.pdf
- Sizmur, S., Graham, C., & Walsh, J. (2014). Influence of patients' age and sex and the mode of administration on results from the NHS Friends and Family Test of patient experience. *Journal of Health Services Research & Policy*, 1355819614536887. doi:10.1177/1355819614536887
- Snelling, I. (2003). Do star ratings really reflect hospital performance? *Journal of Health Organization and Management*, 17(3), 210–223.
- The Strategic Projects Team. (2012). *Friends and Family Test*. Retrieved from http://www.strategicprojectseoe.co.uk/index.php?id_sec=223

