Evaluability Assessment of NHS England Quality Surveillance Team Programme

Dr Axel Kaehne and Dr Tom Simcock

October 2019

Edge Hill University

EPA
Unit for
Evaluation &
Policy Analysis

Evaluability Assessment of NHS England Quality Surveillance Team Programme

Authors:

Dr Axel Kaehne and Dr Tom Simcock
Unit for Evaluation and Policy Analysis, Edge Hill University

October 2019

A report for NHS England Quality Surveillance Team

Edge Hill University

EPA
Unit for
Evaluation &
Policy Analysis

Executive Summary

NHS England monitors the quality of all specialised commissioned and cancer services in England. The Quality Surveillance Team (QST), as part of the wider Quality Assurance Improvement Framework (QAIF), plays a key part in assessing the quality of those services.

QST uses peer reviews visits of clinical teams as part of its quality assurance processes. At the moment, QST are assessing the effectiveness of its peer review processes and have commissioned the Unit for Evaluation and Policy Analysis at Edge Hill University to conduct an evaluability assessment of this programme.

The evaluability assessment was guided by three overarching questions: does the quality of the design of the QST programme allow for an evaluation to be completed? Are the results of the QST programme verifiable based on the planed collection systems? Would an evaluation of the QST programme be feasible, credible and useful? The evaluability assessment employed a review of programme documentation, interviews with the main stakeholders, providing a detailed analysis of the programme's strengths and weaknesses.

This assessment identified findings across four categories:

1. Measuring the impact of QST;
    a. Demonstrating the effectiveness of QST's peer reviews;
    b. Designing and implementing appropriate indicators for quality surveillance as a foundation for peer review visits; and,
    c. The nature of peer review visits vis-a-vis other mechanisms to influence quality improvement of specialised services.
2. Capacity and resource issues;
3. Wider policy context of quality surveillance in the NHS; and,
4. Shared learning.

This evaluability assessment found that a reliable and credible evaluation can be undertaken of the NHS England QST programme. Based on the findings of this assessment and the previous literature review, we have identified four programme domains that the evaluation would need to examine:

1. The aims and objectives of the peer review process (the 'what');
2. The intervention itself (the 'how');
3. The theory of the intervention (the 'why should it work'); and,
4. The staff involved in the process (the 'who').

These relate to four investigative areas: the logic of intervention and development of appropriate measures of success; the role of peer reviewers and staff and their expectations; the fidelity of peer review practice across the programme; and the influence of the broader policy context. We recommend a mixed-methods evaluation

of the QST programme comprising several evaluative tasks, including interviews and focus groups with key stakeholders; observations of three peer review visits (targeted programme); and, observations of peer review training days. This evaluation programme, alongside the completed literature review and this assessment, would provide the QST with the opportunity to identify areas for improvement; to maximise the impact of QST peer-review on services as part of the quality improvement and assurance process; and to design a service that is effective and demonstrates impact.

Contents

Executive Summary 2
1. Background 5
1.1 QST framework 5
1.2 Peer Review Process 5
1.2.1 Literature Review of Peer Review Processes 5
1.2.2 QST Evaluation 6
2. Evaluability Assessment 6
2.1 Purpose and aim 6
2.2 Methodology 6
2.2.1 Review of programme documentation 6
2.2.2 Analysis of the information system 6
2.2.3 Interview of the main stakeholders 8
2.2.4 Analysis of the programme 8
3. Findings 12
3.1 Summary 12
3.2 Recommendations 13
3.3 Recommended Evaluation 13
5. References 15
6. Appendix 16
Table 1. Impact Domain Framework (Source: Smithson et al., 2019) 16

1. Background

NHS England monitors the quality of all specialised commissioned and cancer services in England. The Quality Surveillance Team (QST), as part of the broader Quality Assurance Improvement Framework (QAIF), plays a crucial part in assessing the quality of those services. To discharge its role, the QST has developed a QST framework that maps onto the 5 CQC key questions:

- Are services safe?
- Are they effective?
- Are they caring?
- Are they responsive to people's needs?
- And, are they well-led?

1.1 QST framework

The QST framework uses defined metrics to collect information from each provider on an annual basis through a self-report process. The report is based on quality indicators that are aligned to the six programmes of care in England and reflect the particular service specification. The QST implements this framework through an annual self-declaration process and a peer-review process. The self-declaration process allows QST to obtain relevant data through an established Quality Surveillance Information System (QSIS) where categories are populated by service responses, then gathered centrally and analysed by regional hubs. Aggregated reports for services are then reviewed, and actions are agreed following engagement with commissioners and service leads. Additional surveillance actions are expected where services score less than 100 per cent of their service's previously agreed quality indicators or fail their good practice compliance threshold.

Three types of actions are possible. Option 1 is routine surveillance. Option 2 is enhanced surveillance, involving either provider or commissioner action, or both. Option 3 is peer review. Options can be combined depending on the level of risk assessed by QST in consensus with provider and commissioner of a specific service.

1.2 Peer Review Process

The Quality Surveillance Team have developed a Standard Operating Procedure (SOP) for their peer-review process. This SOP outlines the objectives, processes, and the responsibilities of everyone involved.

1.2.1 Literature Review of Peer Review Processes

NHS England funded Edge Hill University as part of this programme of work to undertake a review of the existing literature on peer review processes. This research found that peer review processes remain insufficiently evidenced and differ considerably in their aims and objectives and their intended impact. A key recommendation of this research was that peer review processes suffer from poorly articulated models of change or logics of intervention that can be tested and refined.

1.2.2 QST Evaluation

The QST is keen to evaluate their current programme and use of peer review to:

- Identify areas for improvement;
- To maximise its impact on services as part of the quality improvement and assurance process; and,
- Ensure the effectiveness of the programme.

As part of a developing programme of research and evaluation between NHS England QST and Edge Hill University, the QST have commissioned the Unit for Evaluation and Policy Analysis to undertake an evaluability assessment of the QST peer-review process.

2. Evaluability Assessment

2.1 Purpose and aim

The purpose of this evaluability assessment was to examine the extent to which EHU can evaluate reliably and credibly the NHS England QST peer review programme.

The following questions guided this evaluability assessment:

1. Does the quality of the design of the QST programme allow for an evaluation to be completed?

2. Are the results of the QST programme verifiable based on the planed collection systems?

3. Would an evaluation of the QST programme be feasible, credible and useful?

2.2 Methodology

A systematic evaluability assessment of the QST programme was undertaken, supported by the framework developed by the UNODC (2019). The evaluability assessment employed a review of programme documentation, interviews with the main stakeholders, and an analysis of the programme.

2.2.1 Review of programme documentation

The research team reviewed the programme documentation provided by the stakeholder. The purpose of this element of the assessment was to identify information on the goals and organisation of the QST process and to inform the design of the semi-structured interview schedule.

2.2.2 Analysis of the information system

The review of the programme documentation provided by QST revealed that there is sufficient material available to undertake a robust assessment of the quality and impact of QST peer review. In particular, the QST have implemented a Standard

Operating Procedure (SOP) which is the basis of their work. The SOP sets out the relevant Service Compliance Processes, including the overall framework, the process for the development of Service Specific Quality Indicators and the data collection from all services through the Annual surveillance process. Also, it specifies the stages of the Annual Assessments and actions to be taken by QST as a result of the assessment.

Peer review as one possible action following a failure to comply with service-specific quality indicators is also clearly defined, and QST appear to work towards an agreed structure of the peer review visit, including prior and post-visit activities.

QST also collect and may make available all peer review reports which are produced upon completion of a peer review visit. There are, however, potential gaps in documenting service processes. In any potential evaluation this may require additional efforts in terms of data collection.

The nature, content and response to the verbal feedback to any peer-reviewed service at the time of the visit is currently not captured by the team. The verbal feedback element may require observational methods to be assessed in any future evaluation.

In addition, there is currently no data or information on the service response following the peer review visit. While the annual assessment process is likely to obtain some relevant information, which may reflect service changes implemented in the wake of the visit, the causal attribution of service changes to the peer review visit is unlikely to be sufficiently evidenced by routine annual reporting.

This highlights the lack of any documentation or processes of follow up, which makes assessing the impact of the peer review visits difficult. The lack of follow up and impact assessment of the peer review visit extends to other potentially relevant stakeholders as well as the specific service visited. This means that, in general, the role, impact and appreciation of QST's peer review visits as part of the broader quality surveillance process appears to be insufficiently documented. Any potential evaluation would need to investigate the impact of QST on the wider policy arena and its stakeholders, including specialised commissioning of NHS England. For this purpose, an evaluation would need to design appropriate data collection methods and instruments.

As for peer reviewers, QST retains a database of potential reviewers which would be available to any evaluation for analysis and interviewing. The training of peer reviewers occurs before the visit organised by regional teams and training sessions are accessible to any evaluators for observational methods if required. Training materials may differ across the regional teams due to the specific nature of services visited. However, there appear to be no difficulties to assess the training methods, ascertaining cross-regional fidelity or effectiveness of peer review training by a potential evaluation.

Interviewing peer reviewers may constitute a significant part of any potential evaluation to gain a comprehensive picture of the role, utility and shared learning processes that may result from peer-review visits to individual stakeholders and the broader system. While possible, it may be advisable that interviewing of peer reviewers takes place immediately prior or after the service visit to ensure that peer reviewers are still engaged in the process and are willing to take part in the evaluation.

2.2.3 Interview of the main stakeholders

The research team developed a semi-structured interview schedule based on the findings of the previous literature review, the review of programme documentation, and the analysis of the QST processes. The schedule was shared with key stakeholders within the QST team at NHS England to check for clarity and further suggestions. Both researchers undertook one telephone interview together, and a further two telephone interviews were undertaken by one of the researchers. One researcher undertook the analysis of the interviews. We synthesize the interview findings with the findings of the other aspects of this assessment in the next section.

2.2.4 Analysis of the programme

Analysis of interviews with key programme leads at QST revealed a range of issues that may usefully be part of any potential evaluation. We group these issues into four categories:

- Measuring the impact of QST;
- Capacity and resource issues;
- Wider policy context of quality surveillance in the NHS; and,
- Shared learning.

2.2.4.1 Measuring the impact of QST

The aspect of whether or not current data collection processes allow QST to measure its impact on services extends to three issues:

- Demonstrating the effectiveness of QST's peer reviews;
- Designing and implementing appropriate indicators for quality surveillance as a foundation for peer review visits; and,
- The nature of peer review visits vis-a-vis other mechanisms to influence the quality improvement of specialised services.

a) Demonstrating the effectiveness of QST peer review visits

At present, it appears to be difficult for QST to demonstrate how peer reviews have influenced services in the short or medium term. While monitoring of quality indicators through the dashboard may reveal service improvement, QST has no routine process in place to evidence the impact of its peer review visits. This remains a significant obstacle to evaluate the effectiveness of QST. While there is essential anecdotal and ad-hoc information about the willingness of visited services to improve, there is no

systematic engagement of QST with services post-visit which makes it difficult to assess the effect of peer review visits.

QST also appears to lack any detailed logic of the intervention (the peer review visit) which would specify through which mechanisms their peer review visit is supposed to trigger the positive changes in the service. This echoes the findings of our literature review which points to generally poorly articulated mechanisms of intervention in peer review processes in health services (Kaehne, Simcock, and Onochie, 2019). In particular, QST peer review visits appear to rely heavily on top-down managerial steering processes for service compliance and service improvement which are difficult to evidence and conflict with the wider policy direction of clinically led, professional collaborative approaches to work by staff in NHS specialised services.

There is some doubt whether frontline staff are aware of peer review visits at the time. In sum, the logic of intervention applied or assumed by QST and the potential impact of staff and their processes is a key area for investigation for any potential evaluation study and for making useful recommendations for improving the utility of peer reviews.

b) Designing and implementing service indicator lists

At present, the production of indicator lists is linked to the annual assessment process of QST through the dashboard. There is a question as to whether these indicators are capable of capturing the right components that make services deliver good care to patients. An evaluation would need to review how performance and quality indicators are identified and utilised throughout the peer review process and whether they are reflective of good quality care.

A related issue is whether the indicator list emphasises unduly compliance with basic service quality parameters and whether it places sufficient stress on potential service improvement. This strikes at the heart of the nature of the QST peer review process, which appears two-pronged in its intentions while potentially too focused on compliance. Again, this resonates with findings from our literature review on peer review processes and appears to be characteristic of other programmes too. An evaluation of QST peer review process would need to engage with all stakeholders to identify the main thrust and the desired intentions of the programme to contrast this with the methods of delivery and their effect.

A related issue is the fidelity of the programme intervention itself, the peer review visit. While the training of peer reviewers may ensure some cross-regional consistency in the processes when engaging visited teams, there is currently not sufficient information about how this consistency is assessed routinely, how unwarranted variation may be addressed systematically, or indeed, what the nature of programme fidelity would be in the context of peer review visits. Interviewees clearly expressed some concern about programme fidelity but also mentioned that it was difficult to measure in the absence of any clear conceptual framework.

c) Nature of peer review as quality improvement intervention

As mentioned above, there appears to be a dual purpose of the programme relating to, on one side, quality surveillance assessed by compliance to service quality indicators, and, on the other hand, relating to quality improvement. While quality improvement may be narrowly defined as a way to ensure the compliance of a service with a given set of quality indicators, it was clear from the interviews with key programme leads that QST aims to contribute to wider quality improvements of specialised services in line with NHS policy on ongoing service improvement.

Current information and data collected through the peer review process, however, do not allow to draw conclusions about whether or not QST peer review visits meet either or both of these objectives. There is insufficient information at present whether peer reviewers themselves are predominantly looking for compliance or focus on additional service improvement during the visit. Although there appears to be an emphasis on compliance with indicators as a measurable outcome, there were some important nuances in some of the interviews which testify to a wider quality improvement agenda amongst QST staff.

This dual objective of QST peer reviews appears, however, poorly articulated and, at present, lacks clear operationalisation in peer review visit. Any evaluation of the QST peer review programme thus needs to probe the intentions of the programme leads as to this dual programme aim and how peer reviews may be adjusted to take account of this if desired.

2.2.4.2 Capacity and resource issues

Interviewees stressed that QST staff operate under immense pressure with some capacity issues and limited resources. It was also noted that a service re-structuring is ongoing which may or may not alleviate some of the resource issues.

The scope of QST work has considerably increased over the last years with the service taking on the quality surveillance of all specialised services within the NHS (England) in addition to the traditional task to assess cancer services in England. While the SOP has been modified to take account of this change, the number of teams annually assessed stands at over 6,700.

Capacity issues may be exacerbated by national comprehensive peer review visits whose effect on services in terms of service improvement are not altogether clear. There are 2 to 3 comprehensive national peer review visits, and interviewees expressed some doubt as to their focus and impact in contrast to targeted visits which were thought to be more useful.

Any evaluation needs to gain a view on the comparable effectiveness of the national visits vis-a-vis targeted peer reviews to ascertain the relative value of each strand of work. Focusing limited resources on areas where service improvements are most likely would be a key recommendation to be expected from an evaluation study. This

may require additional observational and comparative programme evaluation methods. It is unclear that the existing data routinely collected by QST would permit meaningful impact comparison between national and targeted programme components.

2.2.4.3 Wider Policy Context

QST is operating in a field that is influenced by a multitude of NHS policies. While interviewees acknowledged the service improvement context of NHS I and NHS E, there appeared to be less engagement with some other policies or health system changes which may be of relevance to the peer-review programme. An evaluation of QST would need to arrive at recommendations for QST on how to incorporate these additional developments and how to adjust service practices to take account of the wider NHS service development and innovation policies.

A fundamental change in the NHS is the shift from top-down systems of management to clinically led, semi-autonomous professionally led health care delivery. While the QST peer review visit programme is mainly aimed at individual specialised teams of clinicians, the level of engagement with the team may rely overly on outdated top-down notions of dissemination, a hierarchical model of innovation within teams, and on conventional concepts of leadership.

Any evaluation needs to probe how the peer review programme may adapt effectively to this shift in health service organisation and what change would be required to ensure maximum impact supporting this shift.

A second change ongoing in specialised services in the NHS is the development of clinical networks across health care organisations and increased pressure to integrate or co-ordinate clinical units and teams from different trusts. In particular, hub and spoke models are increasingly developing in some specialised services across organisational boundaries which has implications for the service improvement agenda. The emergence of networks changes how service delivery is organised but also how mutual learning occurs across trusts and teams.

At present, it appears that QST peer review visits take less account of these novel and emerging forms of reciprocal organisational and individual engagement between teams and clinicians. This has consequences for how QST may conceptualise shared system learning through peer review visits. Any evaluation needs to investigate the extent to which peer review visits currently rely on conventional ideas of diffusion of clinical knowledge, mutual learning and cross-organisational clinical networks.

2.2.4.4 Shared learning

Understanding how shared learning occurs through peer-review visits appears to be a significant weakness of the current programme. Interviewees noted that each visit generates a list of key areas of good practice which is shared with Clinical Reference Groups (CRG). This illustrates an awareness of the need to establish shared learning processes, but it remained unclear whether CRGs are capable or willing to institute a

more extensive, and more effective shared learning process within QST or even within the wider health system. An evaluation would need to examine what is the best vehicle to increase shared learning from each peer review visits and how to maximise its impact for the various stakeholders of the programme: the peer reviewers, the visited teams as well as NHS England, and the health system as a whole. Also, there were some concerns that the relationship with the Care Quality Commission and its work may be under-utilised.

3. Findings

3.1 Summary

The analysis of the programme documentation and the interviews with programme leads revealed three main domains that require investigation in any future evaluation study.

First, it is not clear what logic of intervention or mechanism of intervention may underpin QST peer review visit. Assessing the impact of QST peer review visits requires a clear concept of how and why a peer review visit should make a difference to a service. This echoes findings from previous published research and represents an often-identified Achilles heel of peer reviews in health care systems (Kaehne, Simcock, and Onochie, 2019).

Second, and closely related to the first domain, QST may not collect data and information of sufficient internal validity to examine the impact of its peer review visits. In effect, this means that QST would find it difficult to demonstrate the effectiveness of its peer reviews above and beyond simple long-term compliance of targeted services through the annual assessment process. Much in this domain depends on how programme leads define programme success. If defined narrowly, an evaluation may restrict itself to examine service returns logged on the dashboard. However, the complexity of service improvement, and diffusion and normalisation processes of improved clinical practice is likely to require a more holistic and broader notion of service quality. An evaluation would then need to work with programme leads to develop a broader framework of quality improvement and make recommendations on how QST could operationalise it.

As part of an evaluative inquiry about the internal validity of current measures of success by QST, it would also be essential to initiate a discussion with staff in visited sites about the effect of visits on their routine practices.

Third, an evaluation would need to appraise the current level and nature of engagement of QST staff and peer reviewers with visited services to ascertain the extent to which peer reviewers and staff play complementary roles in the process and to gauge their relevant, and potentially differential impact.

3.2 Recommendations

A potential evaluation of the QST peer review programme would be able to draw on significant data routinely collected by QST. It would also need to conduct additional observational fieldwork to gain insight into QST practices on sites of peer reviews where there may be less comprehensive information available.

A vital task of any evaluation would be to identify, articulate and test a mechanism of intervention or logic model setting out how programme leads assume peer review visits to work, and to make an impact on services. This would include an assessment of whether or not QST has sufficient capacity to undertake effective and impactful peer review visits and whether or not its limited resources are adequately targeted to attain its objectives.

Last but not least, an evaluation would strive to make recommendations on how QST peer review visits may adapt to the changing nature of clinically-led specialised services in the NHS and new forms of shared learning and emerging delivery networks. This would speak to the rapidly changing policy and practice context in which QST operates.

3.3 Recommended Evaluation

This evaluability assessment has found that a reliable and credible evaluation can be undertaken of the NHS England QST programme.

Based on the findings of this assessment and the previous literature review, we have identified four programme domains that the evaluation would need to examine:

1. The aims and objectives of the peer review process (the 'what');
2. The intervention itself (the 'how');
3. The theory of the intervention (the 'why should it work'); and,
4. The staff involved in the process (the 'who').

These relate to four investigative areas: the logic of intervention and development of appropriate measures of success; the role of peer reviewers and staff and their expectations; the fidelity of peer review practice across the programme; and the influence of the broader policy context.

We propose that the impact framework developed by Prof Kieran Walshe and his team from a recent study of the impact of CQC would be a suitable framework to guide an evaluation (Smithson et al., 2018). This framework contains eight impact domains of anticipatory, directive, organisational, relational, informational, stakeholder, lateral, and systemic. The framework, alongside the development of a robust and clear logic model of the QST programme, would capture the diversity of potential impact of the service reviews and will allow us to identify strengths and weaknesses of each impact domain in the current peer review process operated by QST.

We recommend a mixed-methods evaluation of the QST programme comprising several evaluative tasks, such as interviews and focus groups with key stakeholders; observations of 3 peer review visits (targeted programme); and observations of peer review training days (2).

We would undertake a series of semi-structured interviews with a purposively selected number of stakeholders, including members of QST and staff (recipients of peer review reports) at commissioning and/or provider organisations are undertaken. These interviews would allow us to develop and subsequently test, the various mechanisms of action (theory of actions) that underpin the peer review process and its individual domain components. We would also propose that the evaluation includes several site visits at services that are undergoing peer review.

5. References

Kaehne, A., Simcock, T., &, Onochie, D., (2019). *A Literature Review of Peer Review Models in Health Care*. [Unpublished Report]. Ormskirk, UK: Edge Hill University.

Smithson, R., Richardson, E., Roberts, J., et al. (2018). Impact of the Care Quality Commission on provider performance: Room for improvement? London, UK: The King's Fund. Retrieved from: https://www.kingsfund.org.uk/publications/impact-cqc-provider-performance

UNODC., (2019). *Evaluability Assessment Template*. United Nations Office on Drugs and Crime.

6. Appendix

Table 1. Impact Domain Framework (Source: Smithson et al., 2019)

Impact mechanism	Description of logic/causal chain/process
Anticipatory	The peer review sets quality expectations, and providers understand those expectations and seek compliance in advance of any review interaction.
Directive	Providers take actions that they have been directed or guided to take by the QST.
Organisational	Peer review interaction leads to internal organisational developments, reflection and analysis by providers that are not related to specific directions. This leads to changes in areas such as internal team dynamics, leadership, culture, motivation and whistleblowing.
Relational	Results from the nature of relationships between QST staff and providers. Informal, soft, influencing actions have an impact on providers.
Informational	The QST collates intelligence and puts information about provider performance into the public domain or shares it with other actors who then use it for decision-making (e.g., commissioning, patient choice).
Stakeholder	Peer review actions encourage, mandate or influence other stakeholders to take action or to interact with the provider.
Lateral	Peer review interactions stimulate interorganisational interactions, such as providers working with their peers to share learning and undertake improvement work.
Systemic	Aggregated findings/ information from peer review are used to identify systemic or interorganisational issues, and to influence stakeholders and wider systems other than the providers themselves.

Edge Hill University

EPA

Unit for Evaluation & Policy Analysis