Investigating Locally Commissioned Evaluations of the NHS Vanguard Programme

Citation for published version (APA):

Citing this paper
Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

General rights
Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Takedown policy
If you believe that this document breaches copyright please refer to the University of Manchester's Takedown Procedures [http://man.ac.uk/04Y6Bo] or contact uml.scholarlycommunications@manchester.ac.uk providing relevant details, so we can investigate your claim.
Investigating Locally Commissioned Evaluations of the NHS Vanguard Programme

Part I: Synthesis of the Nature and Quality of Locally Commissioned Evaluations

Part II: Exploration of Evaluation Experiences with Evaluation Leads

Authors:
Mr Paul Wilson
Professor Jenny Billings
Dr Julie MacInnes
Dr Rasa Mikelyte
Ms Elizabeth Welch
Ms Robby Richey
Professor Kath Checkland
Disclaimer:
This draft evidence synthesis is part of independent research commissioned and funded by the NIHR Policy Research Programme (Ref: PR-R16-0516-22001). The views expressed in this publication are those of the authors and do not necessarily reflect those of the Policy Research Programme, NIHR, NHS England, the Department of Health, arm’s length bodies or other government departments.

For more details contact: Paul Wilson (paul.wilson@manchester.ac.uk)

Acknowledgements:
We are very grateful to the NHS England New Care Models support and evaluation team, who helped us to gain access to the relevant documents and answered queries patiently and promptly, and to the local evaluators who gave us their time and answered our questions. We would also like to acknowledge the support and engagement of the wider research team, with whom these findings have been discussed and refined, with particular thanks to Anna Coleman and Pauline Allen.
Table of Contents

Background ........................................................................................................................................... 3

PART I: Synthesis of the Nature and Quality of Locally Commissioned Evaluations .................. 5

Aims .................................................................................................................................................. 5

Methods .............................................................................................................................................. 5

Findings from Enhanced Health in Care Homes Vanguards .......................................................... 7

Findings from the PACS Vanguards ................................................................................................. 18

Findings from the MCP Vanguards ................................................................................................. 31

Summary and Commentary on Key points ...................................................................................... 48

PART II: Exploration of Experiences of Evaluation with Evaluation leads ............................... 52

Aims .................................................................................................................................................. 52

Methods .............................................................................................................................................. 52

Findings ............................................................................................................................................. 54

Summary and Commentary on Key Points ..................................................................................... 72

Overall Conclusion ............................................................................................................................ 77
Exploration of Experiences of Evaluation with Evaluation leads

Background

In England, the 2012 Health and Social Care Act made innovation in the provision of health services a statutory duty with further impetus for major system change set out in the Five Year Forward View in 2014. In the Five Year Forward View it was argued that the divide between primary care, secondary care, community services and social care was increasingly a barrier to the personalised and coordinated health services patients need.

With innovation in service delivery increasingly viewed as crucial to the long term sustainability of health systems, NHS England launched the Vanguard programme in 2015. Fifty Vanguard sites were to act as test beds for multicomponent innovations in service delivery (see Box 1), supported by a £200 million transformation fund from NHS England.

The overarching goal of the Vanguard programme was to enable local health economies to move quickly to change the way that services are delivered, breaking down barriers between sectors and improving the coordination and delivery of care. It is intended that the new models of care will improve: population health and wellbeing; quality and equality of care; and the overall health and care system efficiency. Five ‘new care models’ were established (see Box 1).

**Box 1. NHS Vanguard new care model types**

<table>
<thead>
<tr>
<th>Model Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Integrated primary and acute care systems (PACS)</td>
<td>Nine sites joining up general practice, hospital, community and mental health services</td>
</tr>
<tr>
<td>Enhanced health in care homes</td>
<td>Six sites offering older people better, joined up health, care and rehabilitation services</td>
</tr>
<tr>
<td>Multispecialty community providers (MCPs)</td>
<td>Fourteen sites focused on moving specialist outpatient and ambulatory care out of hospitals into the community</td>
</tr>
<tr>
<td>Urgent and emergency care</td>
<td>Eight sites developing new approaches to simplify and improve the coordination of services and reduce pressure on emergency departments</td>
</tr>
<tr>
<td>Acute care collaborations</td>
<td>Thirteen sites that link together local hospitals to improve their clinical and financial viability.</td>
</tr>
</tbody>
</table>

The Vanguard programme represented a novel approach to NHS change and development. With no central blueprint for change, local Vanguards were empowered to develop local solutions to perceived local problems. At the same time, a national programme of support has sought to bring local teams together to share ideas, experiences and solutions to problems encountered.
Each Vanguard site received funding and support to move quickly to establish new ways of working. This included support to develop a ‘logic model’ to guide initial development and provide a framework for describing the underlying assumptions between the proposed change(s) and the desired outcomes. In addition, NHS England allocated around £10 million to individual vanguards to procure and fund a local evaluation from an independent evaluation partner. The expectation for local evaluation was that it would complement national interrogation of outcome metrics by examining the delivery of each Vanguard’s activities in depth (NHS England 2015; 2016). It was anticipated that local evaluations would:

- Capture and evaluate the transformation changes delivered by the Vanguards appropriately. Alongside knowing whether things have changed (through outcome metrics), it is important we understand how, and in what context, the changes have occurred.
- Understand the ‘reach’ of the Vanguard locally. With this in mind, it is important to include output data such as the number of patients affected by changes made.
- Feed the information gathered into ongoing, on-the-ground delivery, so that services are continually improved.
- Share the learning gathered between the Vanguards and more widely, to promote replicability and scale up. Doing so will also help to ensure that we tackle any barriers/issues collectively, for the benefit of the whole.
- Embed a culture of evaluation and knowledge sharing within the Vanguard.

As part of the national evaluation of the New Models of Care Vanguard Programme in England led by researchers at the University of Manchester, we have conducted an evidence synthesis (part 1) and exploration of a sample of evaluation leads (part 2) to assess the commissioning, impact and findings of the locally commissioned evaluations.
PART I: Synthesis of the Nature and Quality of Locally Commissioned Evaluations

Aims

This synthesis was originally conceived as having four aims:

- Describe the commissioning process for local evaluations including the extent to which national guidance and local logic models were reflected in local tender specifications
- Describe the nature and assess quality of the commissioned local evaluations
- Draw out and summarise common themes and findings
- Assess the extent to which locally commissioned evaluations address the important elements of local logic models

This synthesis presents a summary of the nature and quality of locally commissioned evaluations relating to three Vanguard types, namely enhanced health in care homes, the primary and acute care systems (PACS) and the Multispecialty community providers (MCPs). Funding from NHS England for local evaluations amounted to around £7 million for these three types.

Methods

The review protocol was registered in PROSPERO (Registration number: CRD42017069282).

Data sources and searches

In 2016, NHS England invested in the FutureNHS collaboration platform (Kahootz), a secure, online hub for the Vanguard and other integrated care initiatives. The platform was implemented as a direct response from Vanguards and other models of care for a place to work together more collaboratively. The repository was designated as the means by which Vanguards could store, share and access key documents in one central hub. Through this platform, registered users are able to access relevant information, documents and evaluation reports related to the Vanguard programme. Because of this, traditional review search strategies were not deemed necessary for the purpose of this synthesis.

Two researchers accessed Kahootz every month from June 2017 to September 2018 to search for documents relevant to the synthesis. At the start of 2018, the NHS England evaluation team had provided a spreadsheet of anticipated dates for the delivery of final evaluation reports. In September 2018 a cross check of downloaded/extracted documents with the local evaluation deliverables spreadsheet revealed that a number of final reports were outstanding.
Access to a shared folder on a restricted area of Kahootz was then given to the team by NHS England and all available final reports and documents were downloaded. Any reports that may have been received after September 2018 are not included in the synthesis. In addition, to Kahootz, we also searched for eligible evaluation reports on Vanguard and named evaluator websites. All identified documents were downloaded logged and stored in folders on a secure shared drive.

**Study selection**

Any report or slide set from a locally commissioned evaluation of a Vanguard was eligible for inclusion. As our focus was on locally commissioned evaluations of Vanguards, any other evaluations such as those conducted by the Health Foundation Improvement Analytics Unit were excluded from the review.

**Data extraction and quality assessment**

For each identified evaluation, details of the evaluators, questions, methodological approaches and limitations in design and/or reporting were extracted and assessed by one researcher and checked by a second. For each evaluation, we adapted a set of quality questions originally proposed by O’Cathain et al (2008) to assess the quality of mixed methods studies in health services research. As there is no definitive checklist for assessing the quality of mixed methods evaluations, we used these questions to assess the appropriateness of the design, the transparency of reporting of the quantitative and qualitative components, and the extent which what was planned was delivered.

- Is the quantitative component feasible?
- Is the qualitative component feasible?
- Is the mixed methods design feasible?
- Have both qualitative and quantitative components been completed?
- Were some quantitative methods planned but not executed?
- Were some qualitative methods planned but not executed?
- Did the mixed methods design work in practice?

**Method of synthesis**

As the included evaluations were largely mixed methods with variation in the nature and type of quantitative, qualitative and cost components, we performed a narrative synthesis of the evidence. Consistent with an integrative approach to synthesising evidence the narrative synthesis aimed to present a descriptive summary of the nature, type and general quality of evaluations within, and then to generate across Vanguard types, a number of themes relevant to the aims of this review. An iterative process of adaptation and refinement was undertaken by two researchers to generate initial themes. Emergent themes were then sense
checked against themes emerging from the evaluation leads analysis undertaken as part of Work Package 1 of this national programme evaluation.

Nature and Quality of Enhanced Health in Care Homes Vanguard Evaluations

We identified a total of 25 separate reports relating to the local evaluations of the six enhanced health in care homes Vanguards. No final evaluation report for Nottingham was available at the time of writing. Each local evaluation is presented descriptively in Table 1 with any limitations in design and/or reporting highlighted.

Use of logic models

As with all other Vanguard types, the six care home Vanguards were supported to produce logic models; a requirement of their funding in Year 1 only. Logic models describe the anticipated inputs, outputs and impacts of the care model proposed. However, these proposed impacts are only partially reflected well in the research questions proposed by the local evaluations. Only two (Nottingham and Sutton) explicitly reference logic models as part of their evaluation plans with a third (Wakefield) including its logic model as an appendix. No evaluation final report relates the findings back to the logic model.

Evaluation questions

As Table 2 presents, most Vanguards posed multiple questions to be addressed by the evaluator; questions were not explicitly stated in documentation relating to Nottingham. The way questions are framed mirrors closely the generic commissioning guidance circulated by NHS England rather than the locally developed logic models. As a consequence, many questions lack specificity. While many of the evaluations appear to address the research questions stated in the original others do not. Whilst some lack of consistency can be attributed to the iterative nature of the evaluation conducted, with others it is less clear why there has been deviation from the original intentions. Some reports also stated that they were addressing specific questions and then did not do so.

Nature of evaluations

The questions posed tended to encompass a very wide scope which if to be delivered in full would necessitate evaluations that incorporated qualitative, quantitative and economic methods. This is reflected in the approaches taken by the evaluators. All care home
evaluations were planned as mixed methods evaluations or perhaps more accurately, planned to utilise a mix of methods.

More so than the PACs or MCP Vanguards, the care home Vanguards all involved the roll out / spread of defined interventions. Given this, there was an opportunity to generate generalisable knowledge through the conduct of natural experiments that evaluated the longitudinal effects of intervention implementation. Only Gateshead, via the regional funded evaluation of the five North East Vanguards, attempted a quasi-experimental design in the form of interrupted time series (ITS).

Instead, the care home evaluations are more formative by design largely focusing on describing the implementation context and the organisation and delivery of care as it changes using a combination of the following:

- Documentary analysis and hypothesis generation
- ‘Before and after’ outcome analysis on specified service utilisation and other performance metrics
- Stakeholder reflections and experience of the Vanguard
- Patient and carer experience

Three evaluations stated their intention to conduct ‘economic’ or ‘cost’ related analysis (Gateshead, East and North Herts and Sutton). Airedale evaluators included a return on investment analysis although not originally specified. In the case of East and North Herts a lack of data on costs meant the planned analysis was not realised.

Table 3 presents a summary of our assessment of the success of execution of each local evaluation using the questions posed by O’Cathain et al (2008). Where data sharing agreements were lacking and or data access issues significantly curtailed pre specified analyses, we have classified quantitative components as not feasible. In all six cases, planned quantitative analyses were either modified or not conducted. Although not without some operational challenges and acknowledged limitations, Gateshead would appear to represent the most coherent attempt to generate generalisable knowledge beyond their own setting. It should be acknowledged that as part of the regionally funded evaluation of the five North East Vanguards1, and by pooling funds, Gateshead had access to significantly evaluation resources than the other care home Vanguards. This is not to say that the other evaluations were unsuccessful or lack value. Rather, that their value is context specific. For example, Sutton was unable to include a comparator as originally planned and instead devised a weighting approach to enable comparison across care homes. Although this approach was

---

1 The North East five Vanguards are: Newcastle Gateshead (Enhanced Health in Care Homes); North East Urgent and Emergency Care Network (Urgent and Emergency Care Vanguard); Sunderland (MCP Vanguard); Northumbria Foundation Group (Acute Care Collaboration Vanguard); Northumberland (PACs Vanguard).
novel and would have provide some insight into impact for the local audience it lacks external validity and has limited generalisability as a result.

Challenges

While a number of methodological challenges were discussed in evaluation reports it is the evaluators struggle to access routine data on health service utilisation and other outcomes that is most prominent. Despite significant efforts on the part of evaluators (in Airedale, East and North Herts, Sutton and Wakefield) a lack of data sharing agreements and information governance procedures appear to have been significant barriers to data access and to the conduct of outcome analyses. In the case of Airedale, the local evaluation team was not able to access Vanguard project metrics, utilisation data or outcome data at all; the team did support the Vanguard in its Data Access Request Service application to NHS Digital.

As well as delays in accessing data flows, some evaluators have also experienced issues around data quality. Where local sources of routine data have been available, completeness and accuracy of datasets has been a significant issue. Secondary analyses have therefore been constrained by time and missing data. All evaluations are therefore suitably circumspect in the interpretation of quantitative findings and highlight significant limitations where they apply. The time and resource spent pursuing data sharing agreements and data cleaning has significant implications for future evaluation programmes of this type.

The second greatest challenge that evaluators have faced is the pursuit of the counterfactual with most struggling to create meaningful comparators. One particular issue in doing so was the identification of care home residents themselves. As care homes do not have a unique reference number, analyses used the postcode of the care home as a proxy indicator to identify residents and their associated outcomes. Doing so increases the risk of bias in terms of overestimation of the impact on outcomes as data may include non-care home residents (who could be subject to other confounders) who share a postcode. Those that used this proxy indicator recognised its analytical shortcomings but felt it better than a rudimentary before and after analysis. Again, significant time and resource was spent by evaluators trying to generate meaningful comparators. National level comparative analyses are planned by the Improvement Analytics Unit (a partnership between NHS England and the Health Foundation) and by our own national evaluation team and these may be better placed to produce meaningful counterfactual analysis.

Reporting

No standardised reporting requirements were proposed at the outset of the Vanguard programme and as a consequence many of the local evaluation reports are lengthy and challenging to navigate. Lack of standardised reporting makes it difficult to identify the methods used and findings are often not linked back to the original research questions proposed. For the quantitative components, detail is often lacking on the planned statistical
approach for analysis; though this is often implicit in the presentation of results. For the qualitative components, sampling method and size, the method of approach, non-participation rates and methods of analysis are not systematically reported across evaluations.

A more consistent reporting style would have made the reports much more accessible and improved clarity on the methods used in the evaluations. As vanguards approached the evaluations very differently with some evaluated individual aspects or interventions and others evaluating the whole system, greater standardisation in reporting would have made efforts to bring the evidence together to generate shared learning a lot easier.
## Table 1: Design of local evaluations of enhanced health in care homes Vanguards

<table>
<thead>
<tr>
<th>Vanguard / Evaluator</th>
<th>Design</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Airedale</td>
<td>Originally conceived as a mixed-methods evaluation involving retrospective before-and-after design with controlled comparisons. Qualitative component focused on stakeholder reflections conducted post-intervention, using individual interviews, focus group interviews and online or paper surveys. YHEC conducted second attempt at assessing impact of telemedicine service using retrospective before-and-after design with controlled comparisons. Including analysis by different CCG, type of home (residential vs nursing), service usage, environment (urban vs rural). YHEC also developed a return on investment analyses for the overall programme and the sub-levels identified above.</td>
<td>Significant data sharing and information governance issues – AHSN unable to access Airedale project metrics, utilisation data or outcome data. Exact number of care homes receiving telemedicine service unclear. Initially 248 stated but then 235 (148 installed pre- and 87 during Vanguard period) 41 also de-installed. A further 50 homes did not receive the service. YHEC evaluation based on data relating to 141 homes and 25 controls. Analysis based on a data set collated by an unnamed business intelligence organisation. Data cleaning revealed a large number of anomalies and missing data. Cleaned data set was reduced from 290,000 data points to 48,000 for the analysis.</td>
</tr>
<tr>
<td>Gateshead NE Vanguard evaluation: Institute of Health and Society, Newcastle University / Centre for Public Policy and Health, Durham University / Health and Life Sciences, Northumbria University</td>
<td>Combined NE Vanguard evaluation mixed-methods design, combining qualitative and quantitative approaches, to provide contextual understanding of the organisational, technological and economic facilitators and barriers shaping the implementation of the Vanguard programme. Evaluation conducted in three phases: (1) in-depth review of local documentation, semi-structured interviews with key stakeholders involved in the implementation of each Vanguard to identify organisational and technological enablers and barriers; (2) quantitative analysis and economic evaluation; (3) overarching synthesis and identification of emerging key messages for shared learning.</td>
<td>Data collection used postcode areas around care homes - possible overestimation of the impact on outcomes. Limited data points available for the post Vanguard period - potential seasonal effects were not adjusted for. NEWS - small number of homes took part in the pilot (n=2 out of planned 6). Complete data for 84 residents only.</td>
</tr>
<tr>
<td>York Health Economics Consortium (YHEC) (4 reports)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Quantitative analysis was conducted on relevant
### Institute of Health and Society, Newcastle University

**University of Sunderland**

*(5 reports)*

- Outcomes/performance metrics for care homes obtained from the Vanguard team. Analysis utilised ITS and Cox’s regression in order to make inferences regarding outcomes. The analysis conducted separately in each locality because of different stages of rollout. The cost analysis involved an estimate of the costs of running the Vanguards /the economic impact of the Vanguard on the key performance metrics provided.

- 17 interviews, conducted with senior managers and IT managers involved in the implementation of the Vanguard, to explore perceptions and experiences of the programmes’ processes, outcomes and impact. Transcribed interview data and fieldwork notes were analysed using thematic analysis to generate category systems and repeated themes from a regional perspective.

- *Firefly/ IHS qualitative evaluation* involved semi structured interviews with 23 stakeholders, including 13 Pathway of Care (PoC) members. Interviews supplemented with observations of Pathway of Care meetings. Stakeholders directly and indirectly involved were interviewed, to capture a broad range of perceptions of the PoC. Data analysed using Framework approach.

- Northumbria NEWS evaluation: quantitative analysis of NEWS scores to assess the relationship between this score and other demographic and clinical data; qualitative exploration of professionals’ (n=13) decisions relating to admissions to hospital and experiences of NEWS within the context of care homes.

- Sunderland evaluation (informed by ‘realistic’ evaluation) involved semi structured interviews with 11 (of 29) informants directly involved in the contracts and commissioning work stream.

### East and North Hertfordshire Centre for Research in Primary and Community Care, University of Hertfordshire

*(3 reports)*

- Secondary analysis of quantitative data derived from two sources: local Trust and SUS. Descriptive statistics were used to build a picture of HomeFirst and Stroke ESD users and service use. For HomeFirst, analysis explored impact on relevant outcome (service utilisation) against matched controls. No matched controls possible for Stroke ESD.

- Qualitative component involved semi structured data sharing and information governance issues - Delayed access to local Trust and SUS data. Data provided via intermediary (MedeAnalytics) incomplete Secondary analysis constrained by time and missing data.
exploration of experiences of evaluation with evaluation leads

<table>
<thead>
<tr>
<th>Nottingham City Experience Led Care / Age UK</th>
<th>Survey undertaken by Age UK in 20 care homes (2015/2016). 79 relatives participated (25% response rate). Age UK also undertook ‘engagement’ work with 59 care home staff, 40 ‘future users’ and 39 carers. Not clear if this was undertaken in same 5 care homes. ELC commissioned to review the results of the Age UK Notts consultation. ELC then conducted baseline evaluation of resident and care home staff experience of living and working in 5 care homes (2016/17). Purposive and convenience sampling produced a sample of 50 residents (20 supported by a consultee) and 50 staff. Cordis Bright commissioned to undertake a process and impact evaluation of the Vanguard but only baseline report available. Baseline evaluation involved in-depth review of local and national documentation and a review of local data relating to the Vanguard, provided by Vanguard. Semi-structured telephone interviews with 8 key stakeholders involved in the design and/or implementation of the Vanguard were also conducted. A review of literature relating to ‘what works’ in delivering assistive technology and clinical pharmacy in care homes was undertaken.</th>
<th>Cost comparison with usual service use based on unit cost data was planned but data was deemed insufficient in detail to be able to conduct a meaningful analysis. Used matched controls based on gender an age but unable to match health status.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sutton SQW / SCIE (5 reports)</td>
<td>Mixed methods evaluation. Baseline evaluation involved in-depth review of local documentation and local data relating to the Vanguard. Quantitative analysis was conducted on relevant</td>
<td>Data used post-code areas around care homes - possible overestimation of the impact on outcomes.</td>
</tr>
</tbody>
</table>
Exploration of Experiences of Evaluation with Evaluation leads

outcomes/performance metrics derived from Local Trust and London Ambulance Service data. In absence of counterfactual, analysis used a weighting approach - each intervention was assigned a number of points based on the likelihood of it contributing to a change in outcomes. The points for all the interventions for each care home were then added together to give a final score. Outcome data then compared for care homes according to their final score to ascertain if care homes with a higher score experienced greater improvements.

Qualitative component involved semi structured interviews with those involved in support/delivery of the Vanguard (n=14) to gather views on what is working well/not well, for whom and why. Supplemented with survey of care home staff (34 respondents, no denominator) SCIE conducted interviews with family/friend/ carers (n=11) from homes with moderate to significant engagement to explore impact, the 'active ingredients' and unintended costs/ consequences. Supplemented with survey (N=5).

Wakefield
Wakefield Public Health Intelligence
HealthWatch / Niche Health and Social Care Consulting (4 reports)

Originally conceived as a before-and-after design comparing a random sample from intervention with matched controls. Unable to do so revised design to use the care home as the unit of study, with postcode as a proxy indicator for intervention. Statistical process control charts used to identify significant change in service utilisation.

HealthWatch survey included a selection of National Voices ‘I’ statements to measure the variety and integration of services going into the care homes; a validated wellbeing measure (ONS-4); and two qualitative questions regarding independence and wellbeing. Purposive sampling identified 5 care homes to take part in the baseline and follow up (1 Yr) surveys (only 3 care homes were included in Follow up) 42 resident survey interviews in May/June 2016 and 32 follow up interviews in February 2017.

Did not include an assessment of a counterfactual as was originally planned.

Discrepancies between local Trust and LAS data (conveyances and A&E attendances) Unable to conduct planned analyses using local data for ASCOT, AQP or for the link nurse intervention (paper records only)

Unable to find suitable comparator (postcode proxy used)

Lack of data sharing data sharing and information governance issues meant Public Health unable to gain access to SUS (local Trust data used instead but significant data anomalies)

SPC charts of change in unplanned admissions used 2 Sigma rather than standard 3 Sigma

Unable to establish patient level linked dataset so unable to determine which intervention elements had most impact.

Follow up survey data from 3/5 care homes
Table 2. Local evaluation questions for enhanced health in care homes Vanguards

<table>
<thead>
<tr>
<th>Vanguard</th>
<th>Local evaluation questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Airedale</td>
<td>1. How does the model of telemedicine (TM) impact on care home utilisation? Which care homes use TM and why? Which care homes don’t use TM and why? What role does TM model in place play?</td>
</tr>
<tr>
<td></td>
<td>2. What impact does TM utilisation have on key outcomes of healthcare utilisation? A&amp;E attendances, Non-elective admissions, conveyance, GP call outs</td>
</tr>
<tr>
<td></td>
<td>3. Is TM perceived by those closest to it as an appropriate, acceptable, effective healthcare delivery method? What difference does the model of TM make? What are the key benefits, challenges of TM? How can TM in care homes be improved?</td>
</tr>
<tr>
<td>Gateshead</td>
<td>(From North East Vanguard evaluation)</td>
</tr>
<tr>
<td></td>
<td>1. What new conceptual understandings have been used to develop NCMs and what theories of change underpin them?</td>
</tr>
<tr>
<td></td>
<td>2. What new opportunities have become available with the five North East Vanguard Programmes? What challenges have participating organisations experienced in implementing the Vanguard Programmes and how have these challenges been addressed?</td>
</tr>
<tr>
<td></td>
<td>3. To what extent, and in what ways, has digital innovation shaped NCMs’ aspirations and achievements?</td>
</tr>
<tr>
<td></td>
<td>4. What impact has each Vanguard had on the efficiency of the local health and care economy?</td>
</tr>
<tr>
<td></td>
<td>(From pathways of care (PoC) qualitative evaluation)</td>
</tr>
<tr>
<td></td>
<td>1. How has the PoC developed and what are the keys to success?</td>
</tr>
<tr>
<td></td>
<td>2. What are participant views on the approach to clinical leadership and engagement?</td>
</tr>
<tr>
<td></td>
<td>3. How and why has the approach changed practice and influenced resident care?</td>
</tr>
<tr>
<td></td>
<td>4. How may the model and team need to change across different settings to enhance care delivery?</td>
</tr>
<tr>
<td></td>
<td>5. What are the main anticipated barriers to implementation?</td>
</tr>
<tr>
<td></td>
<td>(From Whzan Health System for recording NEWS evaluation)</td>
</tr>
<tr>
<td></td>
<td>1. Assess the relationship between NEWS score and frailty, cognitive impairment, dependency, functional ability and treatment outcomes in the acutely ill older care home resident indicators.</td>
</tr>
<tr>
<td></td>
<td>2. Explore care home staff, NHS community nurses working into care homes and GPs’ understanding of the factors that are key determinants of the presentation of acute illness in frail older people.</td>
</tr>
<tr>
<td></td>
<td>3. Explore care home staff, NHS community nurses working in to care homes and GPs’ views, experiences and barriers to the use of NEWS</td>
</tr>
<tr>
<td></td>
<td>4. Examine the impact of the introduction of NEWS on clinical decision-making process in relation to treatment of the acutely ill older care home resident.</td>
</tr>
<tr>
<td></td>
<td>(From University of Sunderland qualitative evaluation)</td>
</tr>
</tbody>
</table>
1. Explore key stakeholders’ perceptions of current ways of working as a result of the EHCH Vanguard implementation

**East and North Hertfordshire**
1. How the HomeFirst and Stroke ESD teams worked and the perceived strengths and weaknesses of the service?
2. What is the impact of the HomeFirst and Stroke ESD on the wider system?
3. What is the patient and carer experience of the service?
4. What are the outcomes and value for money of the service?

**Nottingham City**
Questions not explicitly stated. Evaluation was to capture the inputs, activity, outputs and outcomes delivered by the Vanguard through its new technology and clinical pharmacy workstreams to deliver the aims of the programme.

**Sutton**
1. What is the context into which Sutton’s model has been implemented?
2. What key changes has the Sutton model made and who is being affected by them? How have these changes been implemented?
3. What is the change in resource use and cost for the specific interventions in the new care model? How is the Vanguard performing against its expectations and how can the care model be improved?
4. What impact is the Vanguard having on: residents’ outcomes and experiences?
   residents’ families and carers experiences?
   the competence and confidence of care staff and managers in the care homes?
   the workforce commissioned to support the care staff and managers in the care homes?
   the way in which resources are used across the local health and social care system?
5. Which components of the care model are really making a difference? In particular, which key components of the model are making the biggest difference, and which of all the components are interdependent and independent?
6. What are the ‘active ingredients’ of the care model? Which aspects, if replicated elsewhere, can be expected to give similar results, and what contextual factors are prerequisites for success?
7. What are the unintended costs and consequences (positive or negative) associated with the new model of care on the local health and social care system, and in general?

**Wakefield**
1. Is care more coordinated and seamless?
2. Has the Vanguard reduced the demand on secondary care from Care Homes?
3. What impact has the Vanguard had on patient outcomes and experience? Including:
   End of life care
   Management of long term conditions and falls
4. What has been the impact on resident wellbeing?
5. Which elements of the intervention have had the greatest impact and could they be replicated elsewhere?
6. What has been the impact of involving community anchors in the lives of care home residents?
Table 3: Assessment of the success of execution of local evaluations (adapted from O’Cathain et al, 2008)

<table>
<thead>
<tr>
<th>Feature</th>
<th>Airedale</th>
<th>Gateshead</th>
<th>Herts</th>
<th>Notts</th>
<th>Sutton</th>
<th>Wakefield</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantitative component feasible?</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Qualitative component feasible?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Mixed methods design feasible?</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Have both components been completed?</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Unclear</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Quantitative methods planned, not executed?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Unclear</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Qualitative methods planned, not executed?</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Unclear</td>
<td>Unclear</td>
<td>No</td>
</tr>
<tr>
<td>Did mixed methods design work in practice?</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Unclear</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>
Nature and Quality of the PACS Vanguard Evaluations

We identified a total of 37 separate reports relating to the local evaluations of seven PACS Vanguards (19 of which were from North East Hampshire and Farnham). Two Vanguards, Wirral and Salford, did not commission local evaluations via the new care models programme. The evaluation approach employed in each Vanguard is presented descriptively in Table 4 with any limitations in design and/or reporting highlighted.

Use of logic models

As with all other Vanguard types the PACS were supported to produce logic models as a requirement of their funding. Logic models should describe the anticipated inputs, outputs and impacts of the care model proposed. For PACS the proposed impacts are only partially reflected well in the research questions proposed by the local evaluations. Only Morecambe Bay evaluators explicitly and consistently refer back to the logic model. In doing so, they emphasise a disconnect between what was specified and then actually delivered and, highlight an apparent lack of a consensus on the ground in terms of what the Vanguard outcomes should be. Mid Notts stated that their evaluation sought to identify what impact the Vanguard programme was having on the outcomes outlined in the logic model but this is not explicitly carried through to the findings. North East Hampshire and Farnham include the logic model as an appendix and state that where possible logic models were used to develop the service evaluations conducted. No other PACS evaluation either mentions or relates the findings back to the logic model.

Evaluation questions

As Table 5 presents, most Vanguards posed multiple questions to be addressed by the evaluator. In PACS, the questions posed mirror the generic commissioning guidance circulated by NHS England rather than the locally developed logic models. Many questions lack local specificity. Both Morecambe Bay and Mid Notts evaluators were explicitly asked to address the questions set by NHS England.

Nature of evaluations

The questions posed tended to encompass a very wide scope which if to be delivered in full would necessitate evaluations that incorporated qualitative, quantitative and economic methods. This is reflected in the approaches taken by the evaluators. All evaluations were planned as mixed methods evaluations or more accurately planned to utilise a mix of methods. North East Hampshire and Farnham and Isle of Wight (both evaluations coordinated by Wessex AHSN) conducted a series of small scale evaluations of specific elements of the Vanguard programme.
Northumberland (via the regionally funded evaluation of the five North East Vanguards) and Harrogate were the only Vanguards to employ more rigorous quasi-experimental designs in the form of interrupted time series (ITS). PACS evaluations generally use a combination of the following:

- Literature and or document reviews of data sources and related interventions;
- ‘Before and after’ activity analyses largely on pre specified service utilisation metrics;
- Qualitative exploration of staff experience, and perceptions of challenges and enablers relating to implementation;
- Qualitative exploration of patient and carer experience;
- Economic or cost related impact analysis

Four PACS (Isle of Wight, Morecambe Bay, Northumberland and South Somerset) planned economic analyses that were subsequently not realised.

Table 6 presents a summary of our assessment of the success of execution of each local evaluation using the questions posed by O’Cathain et al (2008). Where data sharing agreements were lacking and or data access issues significantly curtailed pre specified analysis, we have classified quantitative components as not feasible. Northumberland represents the most coherent and successful attempt to generate generalisable knowledge. As elsewhere, we acknowledge that this regionally funded evaluation had significantly more resources that the other Vanguards. Of the others, Morecambe Bay, although not without challenges and limitations does attempt to relate findings to the original intentions of the Vanguard. Harrogate’s evaluation although severely curtailed, does attempt to situate tentative findings in relation to some existing evidence for intermediate care and integrated services. With other PACS, poor reporting makes it difficult to ascertain whether there were other analyses planned, but not executed.

Challenges

A number of methodological challenges were discussed in the PACS evaluation reports; Mid Notts did not highlight any evaluation challenges. Harrogate, Isle of Wight and South Somerset all struggled to access routine data with a lack of data sharing agreements and information governance significant barriers to access. The challenge of obtaining data sharing agreements was so significant for South Somerset that the evaluator was unable to undertake the quantitative analysis within the evaluation time frame. The activity analysis planned by Isle of Wight was not done as they were unable to obtain a data sharing agreement for analysis of social care data. Other data challenges included an absence of patient level data for Morecambe Bay and Harrogate’s unsuccessful attempts to capture key service-level data.

Harrogate struggled to create a meaningful comparator for all planned analyses. North East Hampshire and Farnham, Isle of Wight and Morecambe Bay do not appear to have
attempted comparative evaluation, opting instead for before and after activity analyses. Mid Notts activity analysis compared local attendances with national trends.

**Reporting**

No standardised reporting requirements were proposed at the outset of the Vanguard programme and as a consequence many of the local evaluation reports are lengthy and challenging to navigate. Lack of standardised reporting makes it difficult to identify the methods used and findings are often not linked back to the original research questions proposed. Dedicated methods sections are often absent; this is particularly true for summative consultancy reports.

Harrogate, Morecambe Bay and Northumberland are all clearly reported evaluations. For the quantitative components of other PACS, detail is often lacking, beyond the name, on the data sources and models developed for analysis. It is however, possible to derive the nature and size of the cohorts modelled from the presentation of results. The qualitative components for North East Hampshire and Farnham, Isle of Wight and Mid Notts are all poorly reported. Planned sampling method and size, the approaches taken, non-participation rates and methods of analysis are all not well reported across evaluations.

Stipulation of a standard reporting style would have made the reports much more accessible and improved clarity on the methods used in the evaluations. It would have also made efforts to bring the evidence together to generate shared learning easier.
Table 4: Design of local evaluations of primary and acute care systems Vanguards

<table>
<thead>
<tr>
<th>Vanguard / Evaluator</th>
<th>Design</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harrogate</td>
<td>Three work streams: 1) Theory-led qualitative evaluation, 2) Development of Evaluation &amp; Monitoring Metrics, 3) Economic Evaluation. WS1 = Literature review to identify a range of intervention methods for systems change. Documentary analysis: Retrospective documentary analysis to develop a timeline and thematic understanding of the process changes and reporting. Interviews with 24 service managers. Interviews with four GP staff and Practice management involved in the Integrated Response Service (IRS). Two workshops each with a group of six members of the ‘pop-up’ IRS team, to explore what was working well, what could be improved and discuss potential alternative strategies. Semi-structured telephone interviews with up to 10 patients to understand experiences of using the new integrated service. WS2 = Participant observation conducted during weekly Steering Group meetings. Analysis of internal documents: Informal interviews and meetings: with the members of the management team to elicit requirements for the dashboard. Review of the identification and use of data to provide suggestions for the choice and representation of metrics. Assessment of the utility and acceptability of various dashboard platforms for visual representation of data. Development of a Service-User-Record (SUR) to capture key service-level data, which was not successfully implemented. Workforce Dynamics questionnaire (WDQ) used to survey the functioning of the interdisciplinary team. WS3 = ITS analysis on key performance metrics derived from SUS data conducted using Stata 14 using the Cumby-Huizinga general test for autocorrelation.</td>
<td>Access to data limited by a number of factors - challenges of information governance, the excessive pressures on provider information teams, and the lack of means to identify patients who have interacted with the service. Evaluation restricted to aggregated data for metrics covering activity levels in Local Trust. Large level of uncertainty across the ITS analyses as a result of insensitivity of the data available. Comparator evidence unavailable for all metrics. Attempts to initiate service-level measurement tools did not result in systematic or complete data collection. Service-User-Record to capture key service-level data was not successfully implemented. Patient experience interviews not conducted.</td>
</tr>
<tr>
<td>NIHR CLAHRC Yorkshire and Humber</td>
<td></td>
<td></td>
</tr>
<tr>
<td>School of Health and Related Research, University of Sheffield; Centre for Health Economics, University of York; Sheffield Hallam University; York Health Economics Consortium. (3 Reports)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>North East Hampshire and Farnham Wessex AHSN; Centre for Implementation Science, University of Southampton.</td>
<td>A programme of evaluation covering 23 separate services conducted over two years. Common set of methods used across evaluations: R-Outcomes measure (a validated short generic PROM) used by Wessex AHSN used to measure patient and staff experience. 530 staff outcomes received and analysed across 6 collection periods</td>
<td>Overall, programme of evaluation limited through decision to undertake series of small scale evaluations. Modelling of impact on activity levels based on</td>
</tr>
</tbody>
</table>
### Exploration of Experiences of Evaluation with Evaluation leads

**Over 2 years (n=?)**: 3,300 baseline and follow-up patient responses across services over 2 years (n=?).

Staff (n=51) and patient (n=62) interviews to understand experiences and challenges of implementation and nature and extent of change across services. Supplemented by patient survey (n=193).

Team observation focus groups (n=80) and survey (n=137) of integrated care teams using NOMAD questionnaire (Normalisation Process Theory) to understand the extent to which teams were able to embed new care model.

Thematic analysis of 124 case studies. Collected by staff using short template to describe the circumstances, intervention and outcomes for 8-10 of their patients.

Modelling evidence of impact on activity levels over time to estimate potential system savings. Comparison with costs to identify a potential return on investment. Before and after activity analysis undertaken by CSU using pseudo-anonymised patient records to measure the impact of new care models on principally emergency admissions.

### Isle of Wight

**Wessex AHSN**

**CM**: Limited number of participants. Three MDT coordinators participated in focus group and five staff from GP surgeries adopting CMoTAR were interviewed.

Evaluators state no ability to directly select participants to attend the focus group, or participate in interviews. Not possible to recruit staff from surgeries that did not adopt CMoTAR. No details on sampling strategy provided.

**LAC**: No sampling or recruitment detail provided for staff interviews (n=12) or

<table>
<thead>
<tr>
<th><strong>(19 Reports)</strong></th>
<th><strong>(19 Reports)</strong></th>
<th><strong>very small numbers of patients over short time frames (3 months).</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Isle of Wight</strong></td>
<td><strong>Wessex AHSN</strong></td>
<td><strong>Cumulative R-Outcomes responses large but small at service level.</strong></td>
</tr>
<tr>
<td><strong>(5 Reports)</strong></td>
<td></td>
<td><strong>Limited reporting of qualitative elements across services: small numbers of participants over 23 services.</strong></td>
</tr>
</tbody>
</table>

Application of NOMAD questionnaire based on small numbers at service level. Unclear extent of team observation (single observation of MDT meeting or at ‘away day’ rather than routine and longitudinal).
### Exploration of Experiences of Evaluation with Evaluation leads

<table>
<thead>
<tr>
<th>Method</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Care Navigator staff interviews</td>
<td>To understand the experiences, challenges and CN implementation.</td>
</tr>
<tr>
<td>‘Person’ interviews</td>
<td>To understand people’s general experience of receiving CN support.</td>
</tr>
<tr>
<td>Case studies</td>
<td>Provided to complement other qualitative and quantitative data collected. Staff survey to understand general experience of CNs. Data from R-Outcomes measures captured at CN entry point and at completion?</td>
</tr>
<tr>
<td>‘people’ interviews (n=9)</td>
<td>No detail on selection or who completed the 22 case studies. No denominator for survey (n=23) or R-Outcomes measures (n=40 service entry, n=45 for follow up).</td>
</tr>
<tr>
<td>Economic assessment</td>
<td>Of impact of CN service on the use of health and social care services and to calculate potential returns on investment for a number of scenarios.</td>
</tr>
<tr>
<td>CN: No sampling or recruitment detail provided for staff interviews (n=8) or ‘people’ interviews (n=7). No detail on selection or who completed the 11 case studies. 24/32 eligible participants completed survey. No denominator for survey (n=23) or R-Outcomes measures (n=117 service entry, n=121 for follow up).</td>
<td></td>
</tr>
</tbody>
</table>

### Morecambe Bay Health and Social Care Evaluations, University of Cumbria, (3 Reports)

Realist evaluation originally comprising four elements: 1. Desk-based study of the geographical, demographic and cultural contexts of the Vanguard. 2. Semi-structured interviews with 54 service providers and focus groups with 34 patients to gather perspectives around what changes, and who is affected; specifically, around population-based approaches, quality and integration of care, changes in culture and behaviour, and distribution of resources. 3. A large-scale survey of service providers and users to provide a broader set of qualitative and quantitative measures to test the evaluation hypotheses. 4. Assessment of the economic benefits of the Vanguard to identify changes in resource use and cost for specific interventions and evaluating the Vanguard’s overall performance against metrics developed for the Morecambe Bay Accountable Care System.

Lack of consensus about what ‘what success looks like’ and lack of overarching evaluation strategy on part of Vanguard.

Main data challenge was developmental nature of Vanguard - interventions targeted relatively small populations and required dataset would be at patient level. Difficult to obtain any data at this level.

Large scale survey not...
Exploration of Experiences of Evaluation with Evaluation leads

| Northumberland NE Vanguard evaluation: Institute of Health and Society, Newcastle University / Centre for Public Policy and Health, Durham University / Health and Life Sciences, Northumberland University (2 Reports) | Integrated Performance Report (Elective Care = outpatient first attendances and follow-up: Out of Hospital = A&E attendances, non-elective admissions, non-elective bed days and ward closures).

The year 2 phase of evaluation to focus upon three pathways (respiratory, paediatrics and frailty) in three integrated care communities (ICCs) (Barrow Town, Bay and East) revised to focus on respiratory across all ICCs. Evaluation comprised 37 (17 conducted) interviews with staff involved in the development, leadership and delivery of the respiratory pathway. In-depth qualitative interviews with patients on experiences of care and the perceived impact of pathway. Observation of MDTs (yet to commence). Economic impact assessment to assess cost-effectiveness of the new way of working.

| Northumberland NE Vanguard evaluation: Institute of Health and Society, Newcastle University / Centre for Public Policy and Health, Durham University / Health and Life Sciences, Northumberland University (2 Reports) | delivered (n=13 responses).

Data availability issues meant not able to include cost inputs in the economic analysis.

Year 2 assessment of the economic benefits of the Vanguard not undertaken.

Year 2 patient interviews postponed as too few patients had experienced pathway.

Overall realist approach only partially delivered and undermined by lack of a coherent programme theory (from the Vanguard’s perspective) from the outset.

Northumberland NE Vanguard evaluation: Institute of Health and Society, Newcastle University / Centre for Public Policy and Health, Durham University / Health and Life Sciences, Northumberland University (2 Reports) | Combined NE Vanguard evaluation mixed-methods design, combining qualitative and quantitative approaches, to provide contextual understanding of the organisational, technological and economic facilitators and barriers shaping the implementation of the Vanguards programme. Evaluation conducted in three phases: (1) in-depth review of local documentation, semi-structured interviews with key stakeholders involved in the implementation of each Vanguard to identify organisational and technological enablers and barriers; (2) quantitative analysis and economic evaluation; (3) overarching synthesis and identification of emerging key messages for shared learning.

Quantitative analysis was conducted on A&E attendances and emergency admissions for all Northumberland CCG patients from April 2014 to December 2016. (14-months pre and 21-months post). The performance data was split further by hospital Trust in order to view those patients from Northumberland who attended emergency care provision outside of the Northumberland area. Analysis utilised ITS and Cox’s regression in order to make inferences regarding outcomes.

| Northumberland NE Vanguard evaluation: Institute of Health and Society, Newcastle University / Centre for Public Policy and Health, Durham University / Health and Life Sciences, Northumberland University (2 Reports) | Possible patient duplication in data set which could have occurred if they were sent to Northumberland Specialist Emergency Care Hospital after visiting an urgent care centre.

Analysis does not address operational issues relating to the roll out of the Vanguard, including redirection of patients to other services for limited periods.

Analysis focuses on Northumberland CCG population only and is not representative of total activity at Northumberland Specialist Emergency Care Hospital.

24
The cost analysis involved an estimate of the economic impact of the Vanguard on the key performance metrics provided. Despite requesting data regarding the cost of Vanguard provision for the PACS, the evaluation team did not receive this.

13 interviews, conducted with senior managers and IT managers involved in the implementation of the Vanguard, to explore perceptions and experiences of the programmes’ processes, outcomes and impact. Transcribed interview data and fieldwork notes were analysed using thematic analysis to generate category systems and repeated themes from a regional perspective.

<table>
<thead>
<tr>
<th>Organization</th>
<th>Phase 1</th>
<th>Phase 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mid Nottinghamshire Capita Transformation; School of Health and Related Research, University of Sheffield. (4 Reports)</td>
<td>Phase 1: Evidence review of urgent and emergency care (UEC) interventions. Quantitative effectiveness measured using UEC Care Channel shift tool (developed by Capita for the UEC Vanguards), statistical analysis using GRETl (Horizon model) to assess the impact on attendances, admissions and bed days. CUSUM used to test of the stability of the relationship between the national and local trends. Economic modally using UEC tool to assess the productivity of interventions and hence their return on investment. Qualitative focused on experience of clinical pathways for urgent and proactive care for long term conditions - two focus groups with representatives from across all the Local Integrated Care Teams in Mansfield and Ashfield and Newark and Sherwood (n=? participants). Twenty telephone interviews with 9 NHS managers, 3 Social Services Managers and 8 GPs from across both CCGs. Observations of 2 MDT meetings. Phase 2: Continuation of quantitative analysis from Phase 1 involving capacity modelling to develop scenarios around acute and community bed capacity and associated costs. Further cost benefit modelling using the UEC model. Patient experience survey distributed by post to patients identified by Nottinghamshire Healthcare NHS Foundation Trust who had participated in the ‘Better Together’ programme within the last 12 months (n=837)</td>
<td>Phase 1 work presented in a ppt consultancy report with limited detail on methods. Evidence review and qualitative elements poorly reported. One of the few sets of reports not to highlight any evaluation challenges. Phase 2 patient survey yielded 17% response rate (n=143)</td>
</tr>
<tr>
<td>Salford</td>
<td>No Local evaluation commissioned via the new care models programme</td>
<td>N/A</td>
</tr>
<tr>
<td>Location</td>
<td>Description</td>
<td>Notes</td>
</tr>
<tr>
<td>----------</td>
<td>-------------</td>
<td>-------</td>
</tr>
<tr>
<td>South Somerset Symphony</td>
<td>Quantitative analysis using difference-in-differences with match controls to assess the evaluation of both Complex Care Teams and Enhanced Primary Care against key performance metrics (A&amp;E Attendances, Emergency Admissions, Emergency Bed Days / Excess Bed Days / Length of Stay, Outpatient Appointments)</td>
<td>Significant data access issues. Agreement to share data with Evaluation team was not approved by the NHS Digital IGUARD panel until 23rd March 2017, nine months into the project. Therefore, CHE unable to undertake quantitative analysis within the original timeframe.</td>
</tr>
<tr>
<td>South West Academic Health Science Network (SW AHSN), South Central and West Commissioning Support Unit (SCW), Centre for Health Economics, University of York (1 Report)</td>
<td>Data derived from Symphony dataset comprising information about each anonymised individual in the Somerset population for 2013-14 (577,982 individuals) and 2014-15 (583,618), used to provide baseline information and explore matching options. Planned to exploit other datasets to which SCW had access. Qualitative SW AHSN to undertake survey work with 18 GP practices in the Enhanced Primary Care intervention. Validated questionnaires administered to a cohort of up to 400 (based on 100 new per month) new patients by health coaches to measure mental wellbeing (Warwick Edinburgh Mental Wellbeing Scale) and isolation (De Jong Loneliness Scale). Collation and analysis of the results of the above questionnaires as well as existing patient activation measures. With existing EPC service, two focus groups of 8 to 10 patients from across localities, a focus group with 8 to 10 Health Coaches and face to face (or telephone) interviews with GPs (2 from each of the 4 localities) to understand views and gather learning about the programme to date. Patients were to be recruited via practices.</td>
<td>Qualitative work originally to be undertaken by SCW team. However, switched to SW AHSN start delayed of this work until April 2017, 10 months into project. At time of report no work had been undertaken.</td>
</tr>
<tr>
<td>Wirral</td>
<td>No Local evaluation commissioned via the new care models programme</td>
<td>N/A</td>
</tr>
</tbody>
</table>
Table 5: Local evaluation questions for primary and acute care systems Vanguards

<table>
<thead>
<tr>
<th>Vanguard</th>
<th>Local evaluation questions</th>
</tr>
</thead>
</table>
| Harrogate                       | Work stream 1 (Qualitative Process & Theory-led evaluation)  
1. What key changes have the Vanguards made and who is being affected by them?  
2. How have these changes been implemented?  
3. What are the ‘active ingredients’ of a care model?  
4. Which aspects, if replicated elsewhere, can be expected to give similar results and what contextual factors are prerequisites for success?  
Work stream 2 (Metrics Development)  
1. To develop a way of identifying and utilising metrics to measure impact across the dimensions of quality, cost and activity;  
2. To create a methodology that can be used into the future to enable the local system to continually evaluate the impact of the care model;  
3. To develop a dashboard to visualise the data and support operational and quality monitoring;  
4. To demonstrate the impact and to inform the development of the new care model.  
Work stream 3 (Economic evaluation)  
1. Assessment of the economic costs of the Integrated Response Service team  
| North East Hampshire and Farnham| Three overarching aims for the programme of evaluation  
1. To understand the patient, staff and system outcomes of the new models of care and how they were implemented  
2. To work with the Vanguard to use the evaluation findings in further development of the programme  
3. To share the learning from the evaluations to enable spread and adoption to other health care systems |
| Isle of Wight                   | Overarching questions not stated                                                                                                                                                                                                                                                                                                                                 |
| Case Management of those at risk (CMoTAR) in primary care  
1. What is the acceptability of this model?  
2. What are the implementation issues associated with delivery of the CMoTAR programme?  
3. What are the perceived effects of this approach to multi-disciplinary team (MDT) working on primary care and on people who receive care?  
4. What are the challenges and how could the service be improved?  
Local Area Coordinators  
1. What is the nature, extent and impact of Local Area Coordination as part of the My Life a Full Life new care model?  |
| Care Navigators                 | No explicit questions stated for the evaluation of the Care Navigator Service                                                                                                                                                                                                                                                                                 |
| Morecambe Bay                   | Aims to ‘address the question set developed by the national NCM team’:  
1. What is the context (e.g. history, culture, relationships, health inequalities, local and national policies, national legislation) in each Vanguard into which |
<table>
<thead>
<tr>
<th>Northumberland</th>
<th>From North East Vanguard evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. What new conceptual understandings have been used to develop NCMs and what theories of change underpin them?</td>
<td></td>
</tr>
<tr>
<td>2. What new opportunities have become available with the five North East Vanguard Programmes? What challenges have participating organisations experienced in implementing the Vanguard Programmes and how have these challenges been addressed?</td>
<td></td>
</tr>
<tr>
<td>3. To what extent, and in what ways, has digital innovation shaped NCMs’ aspirations and achievements?</td>
<td></td>
</tr>
<tr>
<td>4. What impact has each Vanguard had on the efficiency of the local health and care economy?</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mid Nottinghamshire</th>
<th>Aims to address the question set by NHS England:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. What impact is the Vanguard having? This should be compared against the counter factual</td>
<td></td>
</tr>
<tr>
<td>2. What key changes has the Vanguard made and who is being affected by them?</td>
<td></td>
</tr>
<tr>
<td>3. How have these been implemented?</td>
<td></td>
</tr>
<tr>
<td>4. How is the Vanguard performing against expectation and how can the care model be improved?</td>
<td></td>
</tr>
<tr>
<td>5. Where are unintended cost and consequences (positive and negative) associated with the new models of care on the local health and social care economy and beyond</td>
<td></td>
</tr>
<tr>
<td>6. How do we identify impact of individual interventions when we are implementing multiple interventions in a programme?</td>
<td></td>
</tr>
<tr>
<td>7. What is the change in resource use and cost for specific interventions that encompass the new care models programmes locally?</td>
<td></td>
</tr>
<tr>
<td>8. How are costs shared between commissioners including both the NHS and Local Authorities and how can the output/outcomes be attributed to individual commissioners?</td>
<td></td>
</tr>
<tr>
<td>9. What is the context (e.g. culture history, relationships, health inequalities, local and national policies national legislation) of the Vanguard into which new models are being implemented?</td>
<td></td>
</tr>
<tr>
<td>10. What are the “active ingredients,” of the care model? Which aspects if</td>
<td></td>
</tr>
</tbody>
</table>
replicated elsewhere, can be expected to give similar results and what contextual factors are prerequisites to success?
11. What impact is the Vanguard having on improving system resilience?

<table>
<thead>
<tr>
<th>Salford</th>
<th>No local evaluation commissioned via the new care models programme</th>
</tr>
</thead>
<tbody>
<tr>
<td>South Somerset Symphony</td>
<td>1. Is Symphony meeting the key outcome statements – is the programme seamless, person centred and well integrated?</td>
</tr>
<tr>
<td></td>
<td>2. The Symphony programme focuses on a sustainable model of primary care and has engaged staff to make it a better place to work. Is there evidence that this is the case?</td>
</tr>
<tr>
<td></td>
<td>3. The Symphony team would wish to understand the experience of delivering and receiving care under the Health Coach scheme (e.g. what is working best for them and why, what is not going so well. Which bits of what they do work best?)</td>
</tr>
<tr>
<td></td>
<td>4. It would be helpful to understand which bit of what Health Coaches do is most effective; who are the people for whom this approach is working and for whom it isn’t?</td>
</tr>
<tr>
<td></td>
<td>5. How the ‘communities worker’ model is benefitting the programme and how groups can evaluate how their projects are improving wellbeing e.g. social isolation?</td>
</tr>
<tr>
<td></td>
<td>6. Understand the experience of the extensivist</td>
</tr>
<tr>
<td></td>
<td>7. Understand the experience of a practice in delivering EPC</td>
</tr>
<tr>
<td></td>
<td>8. What elements of the programme are working well, and for those that aren’t what should be done instead?</td>
</tr>
<tr>
<td></td>
<td>9. What are the gaps? What haven’t we done that could be done in the programme?</td>
</tr>
<tr>
<td></td>
<td>10. Does the New Care Model ensure the health and care system is sustainable?</td>
</tr>
<tr>
<td></td>
<td>11. Is the wellbeing of patients being improved and sustained?</td>
</tr>
</tbody>
</table>

| Wirral                       | No local evaluation commissioned via the new care models programme |
Table 6: Assessment of the success of execution of local evaluations (adapted from O’Cathain et al, 2008)

<table>
<thead>
<tr>
<th></th>
<th>Harrogate</th>
<th>NE Hants</th>
<th>IoW</th>
<th>Morecambe</th>
<th>Northumb</th>
<th>Mid Notts</th>
<th>S Somerset</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Quantitative component feasible?</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td><strong>Qualitative component feasible?</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Mixed methods design feasible?</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td><strong>Have both components been completed?</strong></td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td><strong>Quantitative methods planned, not executed?</strong></td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Unclear</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Qualitative methods planned, not executed?</strong></td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Unclear</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Did mixed methods design work in practice?</strong></td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

Note: No local evaluations commissioned by Salford or Wirral Vanguards via the new care models programme.
Nature and Quality of the MCP Vanguard Evaluations

We identified a total of 46 separate reports relating to the local evaluations of nine MCP Vanguards. No local evaluations commissioned by Calderdale or Lakeside Vanguards. No reports were available for Fylde Coast, or West Cheshire. Stockport did not receive evaluation funding as part of the new care models programme. The evaluation approach employed in each Vanguard is presented descriptively in Table 7 with any limitations in design and/or reporting highlighted.

Use of logic models

As with all other Vanguard types the MCPs were supported to produce logic models as a requirement of their funding. Logic models should describe the anticipated inputs, outputs and impacts of the care model proposed. For MCPs the proposed impacts are only partially reflected well in the research questions proposed by the local evaluations. The evaluators of Birmingham and Sandwell, Better Local Care (S Hampshire) and Principia all include reviews of Vanguard logic models as part of an analysis of programme documentation. Principia’s baseline assessment does flag up the need to continually update the Vanguard logic model and delivery plan to reflect the evolution of the programme, but whether this actually happened in the proposed phase two of the evaluation is unclear (as no detail on phase 2 is reported). No MCP evaluation relates any findings back to the logic model.

Evaluation questions

As Table 8 presents, most Vanguards posed multiple questions to be addressed by the evaluator. As with the PACS and care homes, the questions posed most often mirror the generic commissioning guidance circulated by NHS England rather than the locally developed logic models. As a result, many questions lack local specificity and do not directly address key components of each Vanguard. Although the Dudley evaluators set overarching questions, these are not explicitly addressed and the evaluation is better understood as a series of discrete evaluations.

Nature of evaluations

The questions posed tended to encompass a very wide scope which if to be delivered in full would necessitate evaluations that incorporated qualitative, quantitative and economic methods. This is reflected in the approaches taken by the evaluators. All evaluations were planned as mixed methods evaluations or more accurately planned to utilise a mix of methods. Birmingham and Sandwell is explicitly framed as a formative evaluation. Tower Hamlets adopted an embedded ‘researcher in residence’ model. Dudley conducted a series of small scale evaluations of specific strategic and operational elements of the Vanguard programme, with no integration of findings across studies.
Sunderland (via the regionally funded evaluation of the five North East Vanguards) employed a rigorous quasi-experimental design in the form of interrupted time series (ITS). Dudley’s analyses of performance measures used time series (incorporating statistical process control) and funnel plots to assess variation in performance. MCPs evaluations generally used a combination of the following:

- Literature and or document reviews of data sources and related interventions;
- ‘Before and after’ activity analyses largely on pre specified service utilisation metrics;
- Qualitative exploration of staff experience, and perceptions of challenges and enablers relating to implementation;
- Qualitative exploration of patient and carer experience.
- Economic or cost related return on investment analysis

Encompass is one of the few evaluations to recognise both the need and then attempt to elicit professional stakeholder views beyond those strongly associated with the Vanguard.

Table 3 presents a summary of our assessment of the success of execution of each local evaluation using the questions posed by O’Cathain et al (2008). Where data sharing agreements were lacking and or data access issues significantly curtailed pre specified analysis, we have classified quantitative components as not feasible. Sunderland (again via the regionally funded evaluation) represents the most coherent and successful attempt to generate generalisable knowledge. Of the others, Encompass is one of the few evaluations to employ a theoretical framework (Evidence Integration Triangle) and to undertake an overarching synthesis of data sources that explicitly relates back to the research questions originally posed. With other MCPs, poor reporting particularly of qualitative components makes it difficult to ascertain whether some planned aspects have been sufficiently well executed.

No details on evaluation approaches taken by Fylde Coast, Stockport or West Cheshire were available at the time of this review. Phase two evaluation reports were also lacking for Erewash and Principia.

Challenges

Significant data sharing and information governance issues were experienced by the Better Local Care evaluators. West Wakefield indicated that they experienced data sharing and information governance issues though it’s unclear how detrimental these were as the planned quantitative activity analyses appear to have been undertaken. Other data challenges access to GP appointment data on a borough-wide for Tower Hamlets and Encompass were unable to link service user level data to analyse the impact of the community hubs.

Aside from Dudley, Sunderland and West Wakefield, comparative evaluations do not appear to have been attempted. Principia planned to compare local activity against with national trends but no report is available of that phase. Encompass evaluators stated that resources
needed to generate a sufficiently meaningful comparator were beyond the resources and timeframe set for the evaluation.

All MCPs appear to have experienced some challenges in engaging participation from patients, service users and indeed staff with low numbers of interviews and survey response rates a feature of all evaluations (where numbers are reported).

**Reporting**

No standardised reporting requirements were proposed at the outset of the Vanguard programme and as a consequence many of the local evaluation reports are lengthy and challenging to navigate. Lack of standardised reporting makes it difficult to identify the methods used and findings are often not linked back to the original research questions proposed. Dedicated methods sections are often absent.

Dudley (strategic evaluation), Encompass, Sunderland (North East evaluation) are all clearly reported evaluations. For the qualitative components of other MCPs, detail is often lacking on sampling method and size, the approaches taken, target and non-participation rates and methods of analysis. This makes it difficult to assess how well executed the qualitative components have been.

Stipulation of a standard reporting style would have made the reports much more accessible and improved clarity on the methods used in the evaluations. It would have also made efforts to bring the evidence together to generate shared learning easier.
Table 7: Design of local evaluations of the MCP Vanguards

<table>
<thead>
<tr>
<th>Vanguard / Evaluator</th>
<th>Design</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Better Local Care (South Hampshire)</td>
<td>Mixed methods evaluation using a ‘utilisation-focused’ approach including: a desk based literature review; analysis of publicly available secondary datasets to establish baseline levels for indicators; analysis of routine monitoring and feedback data collected by Better Local Care / Southern Health delivery teams; n=20 semi-structured telephone interviews with GPs, MCP delivery group members, project leads and other strategic stakeholders conducted between November and early December 2016 (n=18 follow up interviews conducted in March and April 2017); on-line survey of Better Local Care and wider Southern Health delivery teams (n=107/ estimated 250+ respondents across projects - targeted n=175-215 responses). N=5 ‘deep dive’ case studies on the SDAS, One Team, WebGP, MSK, Paramedic projects. These included telephone interviews with patients using EQ-5D questionnaire (where access provided – patient numbers only presented for Eastleigh Frailty Clinic n=15). Survey conducted pre- post Making Every Contact Count (MECC) training offered (n=15/21 responses from training recipients) N=4 interviews with health sector staff involved in delivering MECC training.</td>
<td>Significant data sharing and Information Governance issues. A lack of consistency in data availability and quality between projects. Local GDPR interpretation meant SUS data unavailable to the RSM team. Cost-savings analysis was hindered by the lack of a common standard on the unit cost of care by profession. No details available on methods for desk based literature review. Limited reporting of methods across all of the qualitative elements conducted. Difficult to establish numbers or respondents.</td>
</tr>
<tr>
<td>RSM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PACEC Limited (7 Reports)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birmingham and Sandwell (Connected Care Partnership)</td>
<td>Conceived as formative mixed methods evaluation: Interviews with (n=18) key informants involved in development of Vanguard and/or in the new care models programme nationally. Local sample included at least one representative from each of organisations that have formed the Partnership, the programme’s board and Steering Group, members of the programme team, commissioners and wider system stakeholders. Documentary analysis including the business case, logic models, minutes of partnership board and programme steering group meetings, detailed service specifications and quarterly reports of performance against local and national metrics – and reviews of published literature on key topics (on evidence about integrated care). Workshop with n=? key stakeholders, held March 2017, to explore options for economic evaluation.</td>
<td>Limited details on documentary analysis and reviews of published literature. Formative evaluation sets out a framework for evaluating Connected Care, including an assessment of the programme’s economic outcomes. Unclear what recommendations were actually taken forward.</td>
</tr>
<tr>
<td>Health Services Management Centre, University of Birmingham (2 Reports)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Later phase of work included online survey to explore staff views and experiences of the changes being made (n=164/400 responses) a partners survey (n=30/90 responses); interviews with (n=8) and focus groups (3 groups / 20 participants) with staff involved in specialist services, EPC and the clinical contact centre.

| Calderdale | No local evaluation commissioned | N/A |
| Dudley | Strategic level evaluation involved semi-structured interviews with n=21 local strategic stakeholders. Interviews were conducted between September and November 2017 by phone (n=12) and face-to-face (n=9). Interviews were recorded, transcribed and analysed thematically using NVivo 11 and a coding framework derived from Strategy Unit work (funded by NIHR) to synthesise the evidence for MCPs. An earlier round of interviews with n=16 stakeholders (14 of whom are interviewed later) were conducted by phone (n=7) and face-to-face (n=9) between June and July 2016. | Overarching evaluation questions but no integration / synthesis of data across projects; rather a set of discrete small scale evaluations undertaken. |
| **The Strategy Unit (formerly Midlands and Lancashire CSU)** | **ICF Consulting Services Limited** | **Health Services Management Centre, University of Birmingham (9 Reports)** |
| | Quantitative activity monitoring (undertaken quarterly?) on series of performance measures for Vanguard using time series (incorporating statistical process control) and funnel plots to assess variation in performance. Data sets from April 2012 to December 2017. | Although strategic level interviews conducted at two time points, no attempt made to present an overarching analysis of change in perceptions over time. |
| | MDT evaluation included: literature review of MDTs; n=12 MDT meeting observations at four selected GP practices (3 observations per GP practice), supplemented by n=24 telephone interviews conducted with staff who attend MDT meetings; n=11 MDT stakeholder telephone interviews with local stakeholders from Dudley’s MDT Implementation Group; online survey with all MDT staff operating within Dudley’s 46 GP practices (n= 140 responses); n=7 telephone interviews with MDT patients. Quantitative data analysis on data provided by Dudley CCG on patients registered with GP practice MDTs and data from routine (SUS) datasets to explore the functioning of the MDT model and its potential impacts (trends in non-elective admissions and resulting length of stay). | Limited reporting of quantitative and qualitative elements of MDT evaluation; no denominator for staff survey and small number of patients. No details of literature review but assumption that this was synthesis (funded by NIHR) undertaken by Strategy Unit. |
| | Long Term Conditions Framework (LTCF) evaluation (Jan to March 2016) included initial online staff survey (n=55 responses from 36 GP practices); interviews with n=29 staff (GPs, practice managers, nurses and HCAs) on framework implementation; n=46 patient interviews/ observations of experience of patient centred consultations; review of n=271 (target n=350) | Very limited reporting of PROMs and PREMs evaluation; no denominator for postal survey and small number of staff and patients interviewed. |
| | | Very limited reporting of PAM evaluation; |
Exploration of Experiences of Evaluation with Evaluation leads

- Care plans focused on the language and content of the goals set, and the extent to which these appeared patient-centred.

PROMs and PREMs evaluation included n=5 interviews with clinical and admin staff (n=5) to gather views on the language and suitability of measures, barriers and enablers for implementation and learning for the future; n=5 interviews with patients who had completed the measures; postal patient survey (n=17 responses).

Patient Activation Measure (PAM) evaluation included n=2 interviews with clinical staff and n=5 patients to judge the practicalities and acceptability of the PAM in practice. Limited activity analysis (based on n=167 who had their PAM level calculated) and a case study of a single patient.

Encompass Centre for Health Services Studies, University of Kent (3 Reports)

- Planned in two phases, utilising the iterative ‘Evidence Integration Triangle’ proposed by Glasgow (2013). Involves responsive mixed methods and use of a multi-agency steering group to agree methods and discuss emergent findings according to the pace of implementation.

The initial phase focused on four CHOCs but revised after 6 months due to slower than anticipated roll out and low patient recruitment. Two of the most ‘mature’ CHOCs became the central focus of the evaluation.

Quantitative data collection included demographic data from service users interviewed (n=13); Interprofessional Collaboration Scale administered to staff before (n=24) and after (n=22) the evaluation period; Encompass activity metrics: CHOC service user profile (n=1,747), referrals, emergency admissions, bed days, length of stay on caseload, A&E attendances, staff type attending CHOCs, investment and cost savings. Analysis of impact on local healthcare services was based on a sample of 100 service users that were seen in the CHOCs between October 2016 and March 2017. Return on investment analysis based on the reduction in the number of emergency admissions multiplied by an average admission cost (derived from the prior admissions associated with the sample of 100 service users).

Data from 10 care plans from interviewed participants would be analysed per CHOC (n=20).

- Challenge to recruit service users with target levels of participation not met and resulting data imbalance between patient and professional sources.
- Difficult to engage professional staff to help identify service users and or participate in the evaluation themselves.
- Some indication of data sharing and Information Governance issues as unable to link service user level data to analyse the impact of the CHOCs.
- Absence of care plan data meant planned analysis not undertaken. (only n=1/13 individuals who were interviewed had care plan document in place.)
Exploration of Experiences of Evaluation with Evaluation leads

Qualitative data included semi-structured interviews with service users (n=13) and professionals and senior managers (n=22); Four consecutive observations of team meetings across three CHOCs were also conducted. Qualitative interview data, including field notes were contextually and thematically analysed, grouped into relevant themes to provide an explanatory account.

Overarching analysis that triangulated data through blending and comparing the data sources within the research questions was conducted.

Evaluability assessment of Encompass social prescribing service provided by Red Zebra also undertaken.

Erewash
Cordis Bright baseline evaluation included: documentary analysis of strategic, operational and performance management information provided by the Vanguard (included the business case, activity and delivery plans, and financial data); a rapid evidence review of ‘what works’ in delivering similar integrated health and social care programmes; in-depth interviews with n=11 key Vanguard stakeholders identified by the evaluation steering group.

Peter Stone evaluation conducted in two phases; the first looking at the degree to which 4 projects had met any agreed outputs and outcomes. The second assessing the social and financial impact of each project (using SRoI analysis). Phase 1 data collection included facilitated discussions at project meetings supplemented by electronic surveys to gather a wider set of views around performance and delivery (planned follow up with telephone or email discussions of the impact of the projects and to gather data to support calculation of impact).

Economics Foundation evaluation included semi-structured interviews with n=15 people involved in the Vanguard; network mapping (undertaken during a workshop in Erewash), rapid review of similar workplace health and wellbeing programme; observations at workshops and local events. Some information also drawn from SRoI analysis completed by Peter Stone.

No final evaluation report from Cordis Bright available at time of this review. Planned to interview members of the public who have had interactions with services provided by Vanguard. No participants were available to take part in an interview during the baseline phase.

Cordis Bright rapid review limited to search of Google Scholar and JSTOR. Looked at first 50 articles for each of 306 search terms. No other details of methods provided.

Very limited reporting of phase 1 data collection for Peter Stone evaluation; unclear how many interviews/dischusions/surveys sent. For SRoI, analysis not planned at outset so no pre-defined baseline and...
<table>
<thead>
<tr>
<th>Location/Municipality</th>
<th>Evaluation Reports &amp; Details</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fylde Coast</td>
<td>No evaluation reports available at time of this review</td>
<td>Unknown</td>
</tr>
<tr>
<td>Lakeside Healthcare Northamptonshire</td>
<td>No local evaluation commissioned</td>
<td>N/A</td>
</tr>
<tr>
<td>South Notts (Principia) Capita</td>
<td>Phase 1 qualitative scoping interviews with n=12? Key informants, focus group (n=2), observation of meetings and presentations. Documentary analysis of value proposition, programme and project plans and documentation relating to the LTC and elective pathway work streams. Quantitative data sets reviewed included emergency and urgent care activity, financial and activity model of value proposition, QIPP savings tracker, Dr Foster? and Right Care? dashboards, unspecified data sources for initiatives in the LTC and elective pathway work streams. Phase 2 evaluation intended as mixed methods with quantitative analysis using GRETl (Horizon model developed by Capita for the UEC Vanguards). Patient experience survey also proposed.</td>
<td>Unclear how many phase 1 interviews conducted, or meetings and presentations observed. Phase 1 work presented in a ppt consultancy report with very limited reporting of all elements. No phase 2 evaluation report available.</td>
</tr>
<tr>
<td>Stockport</td>
<td>Stockport Together reported that they appointed Cordis Bright as their formal evaluation partner in summer 2017 but no evaluation funding made available via the new models of care programme. No evaluation reports available at time of this review.</td>
<td>Unknown</td>
</tr>
<tr>
<td>Sunderland</td>
<td>Combined NE Vanguard evaluation mixed-methods design, combining qualitative and quantitative approaches, to provide contextual understanding of the organisational, technological and economic facilitators and barriers shaping the implementation of the Vanguards programme. Evaluation conducted in three phases: (1) in-depth review of local documentation, semi-structured interviews with key stakeholders involved in the implementation of each project. Post-Vanguard data only available for 12 months; potential seasonal effects were not controlled for. Changes to service provision within Sunderland City</td>
<td>Post-Vanguard data only available for 12 months; potential seasonal effects were not controlled for. Changes to service provision within Sunderland City</td>
</tr>
</tbody>
</table>
### Cordis Bright (5 Reports)

Vanguard to identify organisational and technological enablers and barriers; (2) quantitative analysis and economic evaluation; (3) overarching synthesis and identification of emerging key messages for shared learning.

Quantitative analysis was conducted on all non-elective admissions and length of stay for the Sunderland CCG population from April 2013 - January 2017 (33 monthly data points pre and 13 post). Data for 30-day re-admissions was available from April 2015 to December 2016. Data regarding non-elective admissions included associated month and year of admission and length of stay. Analysis utilised ITS and Cox’s regression in order to make inferences regarding outcomes.

Cost analysis involved an estimate of resource use as a consequence of introducing the Vanguard – cost included staff time (e.g. GP support, community nursing, pharmacy, social workers, and administration), training, community and care home beds, and digital solutions.

11 interviews, conducted with senior managers and IT managers involved in the implementation of the Vanguard, to explore perceptions and experiences of the programmes’ processes, outcomes and impact. Transcribed interview data and fieldwork notes were analysed using thematic analysis to generate category systems and repeated themes from a regional perspective.

Cordis Bright baseline evaluation included a review of 36 strategic and operational documents provided by Vanguard. This included reviewing the business case for the programme and its main work streams, performance evaluation group notes, governance structures, and performance data. A rapid evidence review of ‘what works’ in integrated health and social care programmes. In-depth interviews with 13 key programme stakeholders from 8 organisations.

Cordis Bright 2017 evaluation built on baseline work and included documentary review of 130 strategic and operational documents provided by Vanguard. A rapid evidence review of review of the ‘principles of an MCP’. In-depth interviews with 18 key programme stakeholders from 9 organisations. Additional elements were review and analysis of performance data; ‘cost effectiveness analysis’ of Vanguard programme (actually estimate of resource use as a

Hospitals likely to have impacted upon key metrics such as non-elective A&E admissions and length of stay. Ambulatory wards were created and possible that these admissions would have increased the total number of non-elective admissions.

Cordis Bright rapid reviews limited to search of Google Scholar and JSTOR. Beyond search terms, no other details of methods provided.

Cordis Bright unable to conduct anticipated number of interviews with staff, stakeholders, or patients and service users. Not possible to develop case studies reflecting impact of service user pathways. Data on the full cost of Vanguard delivery unavailable for the evaluation.
<table>
<thead>
<tr>
<th>Location</th>
<th>Description</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tower Hamlets</td>
<td>Social prescribing: Activity analysis of 2,270 referrals made between 1 December 2016 and 31 July 2017. Only non-identifiable data was shared with the evaluation team. Analysis of costs of delivering the scheme and ‘EMIS’ search to examine changes in primary care use. Six month before and after analysis of n=890 patients GP appointment use (reduced from 3,388 to 2,970). Patient experience (n=37 from baseline data of n=173) measured with MYCAW (Measure Yourself Concerns and Wellbeing - a questionnaire designed for evaluating complementary therapies in cancer support). Two surveys conducted in August and September 2017, one with referrers (n=183 from 35/37 GP practices) and the other with VCS organisations (n=47 from 46 organisations). Complemented with 5 interviews with GPs, practice nurses and health care assistants; group interview with 3 VCS representatives; group interview with 3 members of the Wellbeing in Tower Hamlets Forum; focus group with social prescribers (n=8) and managers (n=5); 6 interviews with Steering Group members. Three follow up surveys conducted post Making Every Contact Count (MECC) training offered (n=142/1076 responses from training recipients). Locality based approaches evaluation included documentary analysis of contextual information pertaining to Vanguard (collated through negotiation and collaboration with relevant stakeholders from the programme) – used as basis for interview guides; n=20 semi-structured interviews with key stakeholders at strategic and operational levels of programme; 200 hours of observation of key board, steering workstream commissioning and monthly locality meetings; field notes of informal discussions also used as a source of data.</td>
<td>Very limited reporting of all social prescribing quantitative and qualitative elements; small numbers of participants and limited follow up of patients. Unable to analyse GP appointment data on a borough-wide basis as originally planned. Social prescribing activity and costs are benchmarked against findings from a non-systematic review with a high risk of bias (Polley et al 2017)</td>
</tr>
<tr>
<td>West Cheshire</td>
<td>No evaluation reports available at time of this review</td>
<td>Unknown</td>
</tr>
<tr>
<td>West Wakefield</td>
<td>Qualitative component included semi-structured interviews were conducted with 23 individuals across 9 GP practices. Patient experience assessed via 720 patient surveys collected during the course of visits to GP practices by Healthwatch Wakefield.</td>
<td>Some indication of data sharing and Information Governance issues though unclear how detrimental these</td>
</tr>
<tr>
<td>North of England Commissioning Support Unit (NECS)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HealthWatch / Niche Health and Social Care Consulting</td>
<td>Quantitative analysis involved building on metrics developed metrics and dashboards for the Vanguard for reporting to NHS England. Data included lengths of stay, admissions and A&amp;E attendances, walk in centre attendance rates and referral counts. In addition the number of care navigations, physio and pharmacy time used were collected. A cost-consequence analysis was used to identify the associated costs and outcomes of the HealthPod intervention. HealthWatch also conducted staff experience interviews with 35 staff (from the three Hubs) and their managers conducted in August/September 2016. Thematic analysis used to identify the key themes arising from interviews. Report also describes 680 interviews, 43% conducted with the carer present, with people who had been receiving Connecting Care services. These were conducted in 2014 so appears to be a pre-Vanguard evaluation. York Consulting conducted 56 one-to-one, semi-structured interviews with members of staff involved in Connecting Care. Staff interviewed were nominated by members of the evaluation steering group or by the operational leads for Connecting Care within the partner organisations.</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>York Consulting (4 Reports)</td>
<td>were. Issues with recording of data in some GP practices, meaning all care navigation activity may not have been captured. Some GP practices did not take up invitation to take part in the evaluation. Lack of robust data collection and reporting processes for physio and pharmacy interventions.</td>
<td></td>
</tr>
</tbody>
</table>
Table 8: Local evaluation questions for the MCP Vanguards

<table>
<thead>
<tr>
<th>Vanguard</th>
<th>Local evaluation questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calderdale</td>
<td>No local evaluation commissioned</td>
</tr>
<tr>
<td>Birmingham and Sandwell (Connected Care Partnership)</td>
<td>(from Phase 1 evaluation) 1. What key process changes has the Vanguard made and what changes is it struggling to make? How have these changes been implemented and who is being affected by them? 2. What impact is the Vanguard having on staff experience? How can it improve/measure team working and foster a sense of collective clinical ownership and accountability across our partners? 3. What impact is the Vanguard having on patient experience? How can it collect and use both patient experience and patient-reported outcomes to inform programme development, on an individual level, service level and system level? 4. What are the relative cost/benefits of individual programmes/resources and how can the Vanguard use this to refine the economics of the Connected Care programme? What programmes should be scaled-expanded and or changed/stopped based on an economics appraisal?</td>
</tr>
<tr>
<td>Dudley</td>
<td>(Overarching evaluation questions) 1. What is the context for the programme? 2. What was the perceived need for change and why was an MCP model seen as an appropriate response? 3. What were the major changes initiated by the programme and how well were they implemented? 4. How was the programme experienced across the system, e.g. by: the public, patients, staff and stakeholders? 5. What outcomes were achieved by the programme? How were these outcomes achieved? Were there any negative or unintended consequences of the programme? 6. In what ways and to what extent has the programme changed / improved the pattern of resource use within the local health economy? 7. To what extent has the programme addressed its founding rationale? 8. What lessons - for practice and policy - can be derived from Dudley’s experience? What would need to be considered in order to replicate component parts? Conversely, what can Dudley learn from analogous practice elsewhere?</td>
</tr>
<tr>
<td></td>
<td>(from the MDT evaluation) 1. What is an MDT, which services are represented within Dudley MDTs and what models operate in Dudley? 2. How have Dudley MDTs been developed to date, why were they set up and what problems or opportunities were they established to address? 3. How were the MDTs intended to operate and how are they operating in practice? 4. What factors facilitate working in a multidisciplinary way and what barriers exist that hinder this way of working? 5. What difference do Dudley’s MDTs have on patients and local services and how is this difference achieved? 6. How should Dudley MDTs develop in the future and what lessons can be drawn from the experience in Dudley to date? And,</td>
</tr>
<tr>
<td>Encompass</td>
<td>7. What measures could be used by the MDTs themselves to establish whether they are having the desired effect?</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>1. What impact are the Community Hub Operating Centres having on user outcomes and experience? (e.g., health and wellness, seamless care, access to resources, self-management and independence at home)</td>
</tr>
<tr>
<td></td>
<td>2. What are the components of the care model delivery (or ‘active/successful ingredients’) that are really making a difference? (e.g., associated with leadership, collaboration, continuity of care, nature of change, working environment, information sharing, workforce change)</td>
</tr>
<tr>
<td></td>
<td>3. What are the influencing contextual factors and how have they affected implementation and outcomes? (e.g., history, culture, relationships, working arrangements, contracts, local and national policies)</td>
</tr>
<tr>
<td></td>
<td>4. What changes to the use of resources and activity in the local health system have taken place and to what costs? (e.g., costs for specific new care model interventions, hospital use, intended and unintended costs)</td>
</tr>
<tr>
<td></td>
<td>5. What could be improved, replicated and sustained?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Erewash (from Cordis Bright baseline evaluation)</th>
<th>7. What measures could be used by the MDTs themselves to establish whether they are having the desired effect?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1. What impact are the Community Hub Operating Centres having on user outcomes and experience? (e.g., health and wellness, seamless care, access to resources, self-management and independence at home)</td>
</tr>
<tr>
<td></td>
<td>2. What are the components of the care model delivery (or ‘active/successful ingredients’) that are really making a difference? (e.g., associated with leadership, collaboration, continuity of care, nature of change, working environment, information sharing, workforce change)</td>
</tr>
<tr>
<td></td>
<td>3. What are the influencing contextual factors and how have they affected implementation and outcomes? (e.g., history, culture, relationships, working arrangements, contracts, local and national policies)</td>
</tr>
<tr>
<td></td>
<td>4. What changes to the use of resources and activity in the local health system have taken place and to what costs? (e.g., costs for specific new care model interventions, hospital use, intended and unintended costs)</td>
</tr>
<tr>
<td></td>
<td>5. What could be improved, replicated and sustained?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Erewash (from the Peter Stone evaluation)</th>
<th>7. What measures could be used by the MDTs themselves to establish whether they are having the desired effect?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1. To what extent has Wellbeing Erewash supported people in Erewash to come together, increase their understanding of health and wellbeing and take action as part of a social movement? What difference has the project made,</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Erewash (from the Economics Foundation evaluation)</th>
<th>7. What measures could be used by the MDTs themselves to establish whether they are having the desired effect?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1. To what extent has Wellbeing Erewash supported people in Erewash to come together, increase their understanding of health and wellbeing and take action as part of a social movement? What difference has the project made,</td>
</tr>
</tbody>
</table>
### Exploration of Experiences of Evaluation with Evaluation leads

<table>
<thead>
<tr>
<th>Location</th>
<th>How and for whom?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fylde Coast</td>
<td>No details available</td>
</tr>
<tr>
<td>Lakeside Healthcare Northamptonshire</td>
<td>No local evaluation commissioned</td>
</tr>
</tbody>
</table>
| **South Hampshire** | 1. Understand the mechanisms that have led to changes in outcomes, including an analysis of factors that underpin the impact and replicability of the models of care.  
2. Evaluate performance of the programme against stated aims and objectives, incorporating national and local metrics, routine health service data, as well as analysis of locally collated quantitative and qualitative data.  
3. Undertake an in-depth evaluation of the main models of care, incorporating but not limited to the outputs and outcomes in the programme and thematic logic models.  
4. Evaluate the enablers and barriers to organisational change at primary care and locality level.  
5. Evaluate the impact of the Better Local Care programme from a systems, financial and broader economic perspective.                                                                                                                                                                                                                       |
| South Notts (Prinicipia) | (from Phase 1 evaluation)  
1. How comprehensive is the MCP's logic model: To what extent are inputs (financial and people), activities, outputs and outcomes mapped, understood and quantified?  
2. To what extent is the combined/aggregated impact of interventions understood?  
3. Does the MCP have a consolidated view of when interventions are coming on stream and when their impact is anticipated?  
4. How is the MCP measuring and tracking inputs, activities, outputs and outcomes? Are there any gaps in what is currently being collected, measured and reported? Are there any concerns around data quality?  
5. Can the MCP provide a baseline position? Are there any gaps?  
6. To what extent are anticipated savings, shifts out of hospital and return on investment grounded in evidence? What assumptions have been made and are they reasonable?  
7. What factors make Rushcliffe unique from or comparable to other areas and what are the implications of this for transferability?  
8. What information is the MCP using to target interventions effectively?  
9. What are the early indications for the impact of interventions?                                                                                                                                                                                                                           |
| Stockport         | (from evaluation tender specification)  
1. Compared to the traditional system/service, is the new integrated service solution better for the patient?  
2. How does the new patient journey differ? Does it work for the majority?  
3. In what way are the 8 neighbourhoods different in terms of pathway, outcome, cost, access to borough wide services, acute interface and healthy community initiatives?  
4. Are some neighbourhoods more successful than others? What are the characteristics associated with successful outcomes?  
5. Have patient level outcomes improved across the borough/ by neighbourhood?  
6. How satisfied are people and carers? Does this differ between |
<table>
<thead>
<tr>
<th>Sunderland</th>
<th>(From North East Vanguard evaluation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5. What new conceptual understandings have been used to develop NCMs and what theories of change underpin them?</td>
<td></td>
</tr>
<tr>
<td>6. What new opportunities have become available with the five North East Vanguard Programmes? What challenges have participating organisations experienced in implementing the Vanguard Programmes and how have these challenges been addressed?</td>
<td></td>
</tr>
<tr>
<td>7. To what extent, and in what ways, has digital innovation shaped NCMs’ aspirations and achievements?</td>
<td></td>
</tr>
<tr>
<td>8. What impact has each Vanguard had on the efficiency of the local health and care economy?</td>
<td></td>
</tr>
</tbody>
</table>

(From Cordis Bright baseline evaluation)
1. Consider the context of the Sunderland Care model, including Sunderland specific approaches to MCP development.
2. Review the Recovery at Home/Young People’s Assessment and Liaison service, focussing upon the model of service operation, outcomes for patients and partners, longer term sustainability, return on investment (ROI) and replicability.
3. Review the Community Integrated Teams, including the roles of Living Well Link Workers, MDT Co-ordinators, and the clinical leadership function of GPs. Consideration of outcomes for partners and patients, longer term sustainability, return on investment (ROI) and replicability. Consider how effectively the teams have integrated.
4. Review the programme of Enhanced Primary Care, and the interventions that form part of this. Consideration of the impact of these projects in supporting patients with long term conditions, as well as the ability to release capacity for General Practice.
5. Consider the leadership and governance functions specific to Sunderland, including the CCG assurance function, Integrated Community Services Provider Board, and Programme Management Office.
6. Review the overall outcomes of the Vanguard programme, performance against expectations, and any unintended outcomes.

(From Cordis Bright 2017 evaluation)
1. The impact of ATB on outcomes for:
   - Patients/service users and relatives/families
   - Health and social care staff involved in/with ATB across Sunderland
   - The wider health and social care system across Sunderland
2. Economic impact
3. Process factors regarding the implementation and operation of the ATB Programme
4. The overall impact the programme has had on patients, service users, residents, health and social care staff and the health and social care system in Sunderland.

<table>
<thead>
<tr>
<th>Tower Hamlets Together</th>
<th>(From social prescribing evaluation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. To explore the extent to which the programme has embedded itself within the Tower Hamlets primary care system, its impact on the range of services available to users and the impact on those services.</td>
<td></td>
</tr>
</tbody>
</table>
2. To assess the improvement in health and wellbeing of users of the social prescribing programme.
3. To understand the experience of referring into and delivering the programme.
4. To assess the extent to which social prescribing can facilitate community development in terms of connecting residents with each other for support.
5. To establish the cost savings of the programme within the context of health care and wider public sector budgets
6. To recommend an ideal social prescribing model, including level of funding required

(from the locality based approaches to integrated care evaluation)
1. To explore the effectiveness of the CHS model to enable collaborative working across the partnership at the strategic, operational and service delivery levels and its perceived impact on staff and person outcomes.
2. To establish how the implementation of the Vanguard programme and CHS model has engaged front line practitioners and service users in terms of challenging their values and norms and changing their behaviours.
3. To assess the effectiveness of the Vanguard programme and CHS model in involving and engaging service users and citizens in programme activities and in their experience of using THT services.

<table>
<thead>
<tr>
<th>West Wakefield</th>
<th>[from NECS interim report]</th>
</tr>
</thead>
<tbody>
<tr>
<td>12. How would you describe the context of the programme/population including an understanding of history, cultures, relationships, health inequalities, local and national policies and national legislation?</td>
<td></td>
</tr>
<tr>
<td>13. What key changes have the Vanguards made and who is being affected by them? How have these changes been implemented?</td>
<td></td>
</tr>
<tr>
<td>14. What is the change in resource use and cost for the specific interventions that encompass the new care models programme locally? How is the Vanguard performing against their expectations and how can the care model be improved?</td>
<td></td>
</tr>
<tr>
<td>15. What impact is the Vanguard having on patient outcomes and experience, the health of the local population and way in which resources are used in the local health system?</td>
<td></td>
</tr>
<tr>
<td>16. Which components of the care model are really making a difference?</td>
<td></td>
</tr>
<tr>
<td>17. What are the ‘active ingredients’ of the care model? Which aspects, if replicated elsewhere, can be expected to give similar results and what contextual factors are prerequisites for success?</td>
<td></td>
</tr>
<tr>
<td>18. What are the unintended costs and consequences (positive or negative) associated with the new care model on the local health economy and beyond?</td>
<td></td>
</tr>
</tbody>
</table>

(from York Consulting report)
1. To obtain feedback on the implementation and early outcomes of Connecting Care from staff that are involved in its delivery
Table 9: Assessment of the success of execution of local evaluations (adapted from O'Cathain et al, 2008)

<table>
<thead>
<tr>
<th></th>
<th>BLC</th>
<th>Birmingham</th>
<th>Dudley</th>
<th>Encompass</th>
<th>Erewash</th>
<th>S Notts</th>
<th>Sunderland</th>
<th>T Hamlets</th>
<th>W Wakefield</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantitative component feasible?</td>
<td>No</td>
<td>N/A</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Qualitative component feasible?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Mixed methods design feasible?</td>
<td>No</td>
<td>Yes</td>
<td>N/A</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Have both components been completed?</td>
<td>No</td>
<td>Yes</td>
<td>N/A</td>
<td>Yes</td>
<td>No</td>
<td>Unclear</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Quantitative methods planned, not executed?</td>
<td>Yes</td>
<td>N/A</td>
<td>Unclear</td>
<td>Yes</td>
<td>Unclear</td>
<td>Unclear</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Qualitative methods planned, not executed?</td>
<td>No</td>
<td>No</td>
<td>Unclear</td>
<td>Yes</td>
<td>Yes</td>
<td>Unclear</td>
<td>No</td>
<td>Unclear</td>
<td>No</td>
</tr>
<tr>
<td>Did mixed methods design work in practice?</td>
<td>No</td>
<td>Yes</td>
<td>N/A</td>
<td>Yes</td>
<td>No</td>
<td>Unclear</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Note: No local evaluations commissioned by Calderdale or Lakeside Vanguards. No details available for Fylde Coast, Stockport or West Cheshire. Sunderland assessment based on the North East Vanguard evaluation.
Summary and Commentary on Key points

Significant investment was made available by NHS England to support independent local evaluation of the individual Vanguards. This review represents the first attempt to systematically assess the nature and quality of the evaluations commissioned and to capture methodological learning to inform future endeavours of this type. The synthesis summarises a significant grey literature of 108 local evaluation reports not all of which are publicly available. This review therefore represents the only comprehensive mapping of what was commissioned and conducted and may serve as a key resource for researchers and policy makers, both within the NHS and internationally.

Synthesis Limitations

Whilst we have included 108 local evaluation reports in this review, it is possible that we have not identified some local evaluation reports or that additional reports may have been submitted to the NHS England evaluation team after our synthesis was complete. We are also aware that evaluators have fed back their learning using means other than reports including slide sets, webinars, and face to face interactions. Despite this, we think it unlikely that any unidentified evaluations will be significantly different to those included in this review. Nor do we think the key themes we have identified would differ significantly had we been able to comprehensively capture other modes of communication utilized by evaluators.

There is no definitive checklist for assessing the quality of mixed methods evaluations, and so there are some limitations with the approach we have employed. As such, there is a degree of subjectivity in our assessments of the feasibility and appropriateness of each evaluation design. Our assessments were often hampered by a lack of methodological specificity in evaluation reports, making it was difficult to make judgements about the extent to which individual components were either feasible or indeed realised. Despite this, the criteria proposed by O’Cathain et al (2008) at least provides a structure for interpretation and comparison across Vanguards.

Evaluation characteristics

Most evaluations were both resource and time limited (12 months). As a consequence, most employed mixed methods or more accurately a mix of methods and often attempt but do not address all of the wide range of questions posed in the tender specifications. Most evaluations largely focus on describing the implementation context and capturing stakeholder reflections on and experience of the development and implementation of the Vanguard. Although a
majority state intentions to capture patient experience and conduct ‘economic’ or ‘cost’ related analysis, a combination of resource, data, time constraints mean that these components often lack depth, are often not fully realised or not conducted at all. A similar pattern can be seen with outcome analyses where some evaluators have struggled to access routine data on health service utilisation and other performance metrics which has hampered the conduct of analyses. We have identified a number of common methodological challenges and limitations across the local evaluations and these are discussed below.

Logic Models

The new care models consistently produced and reported logic models which was a requirement of their funding, and which they were supported to produce. These described the anticipated inputs, outputs and impacts of the new care model. However, the proposed impacts are only partially reflected in the research questions proposed by many of the evaluations. It was rare for any evaluators to explicitly frame their approach with reference to the logic model or indeed to refer back to the outputs and impacts original proposed as part of their reporting of evaluation findings. This reinforces the apparent disconnect between the process of producing logic models and then commissioning evaluations. Logic models therefore appear to have been used as an sense making tool for initial programme development, Any potential value to the evaluation process and as an ongoing programme management and improvement tool is not apparent in evaluation reports overall.

Questions

Local evaluations were expected to generate evidence that would inform the main national evaluation questions set out be NHS England (2015; 2016). Although originally issued as guidance, there is some evidence of a ‘lift and shift’ approach to local question formulation with very literal use of the questions circulated by NHS England. This ‘lift and shift’ may be rooted in the national emphasis on examining the entirety of a Vanguard’s programme of activity. This meant that many Vanguards were often asking for an ‘evaluation of everything’ when a more specified approach to question formulation may have led to more focused/meaningful evaluations.

Data sharing and other information governance issues

While a number of methodological challenges were discussed in evaluation reports it is the evaluators struggle to access sources of routine data on health service that is most prominent. Where data sharing agreements were lacking and or there were data access issues, this
Exploration of Experiences of Evaluation with Evaluation leads

significantly curtailed outcome analyses. As well as delays in accessing data, some evaluators also experienced significant issues around data quality. Where local sources of routine data have been available, completeness and accuracy of datasets has been a significant issue. Secondary analyses have therefore been constrained by time data cleaning and accounting for missing data. As many evaluators spent significant time and resource pursuing data sharing agreements, this has significant implications for future evaluation programmes of this type. Indeed it can be argued that without data sharing agreements in place from the outset, many plans for quantitative analysis were not feasible. Given this, the responsibility for data access may best lie with those specifying and then commissioning evaluations.

Counterfactual

The second greatest challenge that evaluators have faced is the pursuit of the counterfactual with most struggling to create meaningful comparators. Significant time and resource was spent by evaluators trying to generate meaningful comparators. As counterfactual analysis relies on generating carefully matched controls it can be argued that few local evaluation teams had the time, resource or crucially the skill set to conduct assessments of this type. Given this, and the investment in parallel national-level analysis of service utilisation and other performance metrics, future programmes may be best to leave the pursuit of counterfactual to national-level evaluations instead.

Qualitative methods

Whilst the above has focused on the key methodological challenges associated with quantitative aspects of the local evaluations, this synthesis also highlights that qualitative methods are often poorly executed (or written up in a way that suggests that). Good qualitative research offers explanatory power and nuanced insight into change mechanisms, barriers and facilitators and intervention transferability and sustainability. Too often in these evaluations, the qualitative aspects are largely descriptive accounts with no real attempt to theorize, generate themes or integrate findings with other data sources to explain or enhance the credibility of the findings. There was some evidence of intentions to use specific approaches such as Normalization Process Theory but without more detailed reporting it was difficult to gauge whether these approaches were really applied beyond the superficial. Other shortcomings across reports included a lack detail on participant selection, sampling and non participation; only one evaluation team explicitly referred to efforts to elicit stakeholder views beyond those directly involved in the development of the Vanguard programme.
Exploration of Experiences of Evaluation with Evaluation leads

Reporting in general

A key expectation of local evaluations was that efforts would be made to share local learning both between the Vanguards and more widely, to promote replicability and scale up. As with research generally, it is crucial that evaluators provide sufficient detail on their methods and the relationship between the analysis and the findings in the report so that readers can assess the credibility of the findings. No standardised reporting requirements were proposed at the outset of the Vanguard programme and as a consequence many of the local evaluation reports are lengthy and challenging to navigate. This lack of standardised reporting has made it difficult to identify the methods used and to describe the key findings. Findings are often not linked back to the research questions proposed. A more consistent reporting style would have made the reports much more accessible and improved clarity on the methods used in the evaluations.

Transparent reporting of evaluations is therefore essential if we are to understand what was done and ensure that learning is systematically captured in a generalisable format and or to enable evaluators to use or replicate the methods in other evaluations. There are a wide range of structured reporting frameworks (www.equator-network.org) that can and should be used. This should include enhancing the quality of intervention description – it was often difficult to work out what was being implemented. The Template for Intervention Description and Replication (TIDieR) is now widely used in health research and could be applied to describe the key components of emerging models of care.
PART II: Exploration of Experiences of Evaluation with Evaluation leads

Aims

This exploration had 4 principal aims:

- To provide context to the synthesis of the evaluations (Part I of the report)
- To investigate evaluation lead experiences of the evaluation process and of the Vanguard programme overall
- To identify factors that supported, as well as inhibited the evaluation (including the influence of the national support team)
- To explore communication between the evaluators and the local Vanguards in deciding a feasible evaluation plan and feeding back findings.
- To investigate evaluator views on the sustainability of the Vanguards and lessons for the future

Methods

Participants

Thirteen staff in local evaluation lead roles were interviewed either individually or in groups of two. Although most were directly involved in the evaluation of specific Vanguards, one interviewee was involved in conducting impact studies nationally, while another was able to give insights on the evaluation of Vanguards across a region.

Some evaluators worked with more than 1 Vanguard so that in total, data collected in relation to 18 Vanguards was analysed. Of these, 9 were MCPs, 5 were PACS, 2 were ECHs, 1 was an UEC and 1 an ACC. There was a broad geographical spread with representation from the North, South, West of England and the Midlands.

There was a range of local evaluation models with some evaluations conducted ‘in-house’ and others using external organisations commissioned by the Vanguards including Universities, Commissioning Support Units (CSUs), NIHR Collaborations for Leadership in Applied Health Research and Care (CLAHRCs), Academic Health Science Networks (AHSNs) and private consultancies. Some used a ‘hybrid’ approach using a combination of in-house and external evaluators, or multiple external evaluators for specific pieces of work.
Exploration of Experiences of Evaluation with Evaluation leads

Instruments

Semi-structured interviews were conducted to collect data on evaluation lead experiences. Aside from factual information on the interviewee role in local evaluations and the Vanguards evaluated, the interviews focused on four aspects of interviewee experience: (1) experience of the local evaluation, (2) feeding back and subsequent use of findings, (3) experience of the Vanguard programme more broadly and (4) opinions on Vanguard sustainability.

Procedure

A set of 42 evaluation leads representing 48 of the Vanguards received an invitation to take part in the interviews via email along with an information sheet. Census sampling was applied (i.e. all individuals connected to the Vanguard evaluations were contacted); the number of approached individuals was lower than the total number of Vanguards, because some of the evaluators led on more than one Vanguard evaluation, while two Vanguards did not commission an evaluation via the new care models programme. Thirteen evaluation leads who responded to the invitation and agreed to take part then provided informed consent by emailing the researchers a consent form. A team of four researchers conducted 2-3 interviews each, all of which took place via telephone. All of the interviews were audio-recorded.

Data analysis

Interview data were analysed thematically, grouped into relevant themes to provide an explanatory account, and then inputted into Excel. This main approach for analysis drew on Flick’s (1998) pragmatic notions of organising and thematically representing data through a pre-determined template, adapted in this case from the interview and focus group schedules and field notes. The Excel spread sheet was set up with the main questions, sub-questions and prompts as column headings, and raw coded data from each qualitative source was grouped into the relevant column. Any data not fitting the template was organised into an ‘other’ category for further thematic analysis. Regular meetings took place among the researchers to ensure a consistent approach.
Findings

Why local evaluators got involved

Evaluation leads became involved in the local evaluations for a number of reasons. Some were already working with the organisations in the Vanguard, it fitted in with their interests and the type of work they were already undertaking, and because they possessed the capacity and expertise to carry out the evaluation. More than one interviewee saw taking the lead on the local evaluation as an opportunity for professional growth. Since one evaluator was already working with the local Vanguard, giving advice and support, it was an obvious transition to take the lead in their local evaluation.

‘We were already well-known to them and already providing support prior to the local evaluation commencing...They brought us in very early on to help out and develop a logic model for them to think through their Vanguard programme. So, it was a natural transition that when the requirement for a local evaluation came round we were the obvious choice.’ (EL6)

When and how they got involved

A number of evaluators were involved right from the start of the local evaluation while others became involved much later in the Vanguard development. The commissioning process involved in procuring the work was lengthy and at times complex. While one evaluator responded to a publically advertised tender, there were mostly ‘single tender actions’ by the Clinical Commissioning Groups (CCGs) rather than open procurement. One site was selected from a shortlist after a presentation. The evaluators were required to submit fully costed bids and different evaluators reported varying experiences in terms of the complexity of the process.

One interviewee who was evaluating more than one Vanguard discussed a sharp contrast between two sites. Dealing with one site was straightforward, involving the submission of a fully costed bid and proposal, which was then agreed through discussions with the commissioning team. However, in the other site numerous staff changes meant the procurement process was lengthy and complicated:

‘[Name of Vanguard site] asked us to provide a proposal and then we had commissioning discussions. However, [at the other site] the person who commissioned the evaluation left the organisation and the people who were left didn’t understand the proposal and it took months to go through the proposal, chopping up and putting back together, so that they were clear and happy with what the intention was. But these people didn’t understand health evaluations.’ (EL2)
The purpose of the local evaluation and aims and objectives of the Vanguards

There was some conflict between the evaluators and the Vanguards in terms of the purpose of the local evaluation. The main purpose, in the evaluators view, was to determine the Vanguards level of success in achieving what they set out to achieve and being able to continue in the longer term. Evaluators described their role as:

‘To be able to provide formative evidence for the site to be able to take forward their findings and improve their services.’ (EL2)

‘Overall looking at what are the benefits, if any, that we can demonstrate with the data we have in the time we have.’ (EL2)

However, one Vanguard had a different view according to the local evaluator:

‘[it was about] performance management rather than a collaboration to provide evidence and support the development of the intervention.’ (EL6)

One evaluator described a more over-arching role:

‘Exploring whether the new care models were having the desired impact and if not why not? The need to understand the process of implementation, what were the facilitators and barriers for further adoption and spread?’ (EL4)

In addition, evaluators had mixed views on how clear the Vanguard sites were about what the local evaluation would involve and what the main aims and objectives of the Vanguard were. Those that reported clear aims stated:

‘To some extent they were doing what NHSE asked of them… which was quite clear in tenders.’ (EL9)

‘Pretty clear and based on the logic models.’ (EL5)

‘They were clearish in terms of what they wanted to achieve…They didn’t really understand the power of qualitative research.’ (EL11)

‘Effectiveness of each service… are they delivering? is it value for money?’ (EL10)

The evaluation leads that reported that the aims were vague, felt this was due to the speed with which changes were implemented, with some work packages:

‘Chucked together in such a rush that in many cases there wasn’t a written account of what they wanted to achieve or what they were trying to do.’ (EL6)
However, aims and objectives became clearer over time through collaboration and negotiation between the evaluation team and the Vanguard:

‘It wasn’t clear at the outset. The aims and expectations evolved gradually. There was not an obvious brief at the outset.’ (EL6)

‘We had to do work to try and refine that set of expectations and then we could get onto designing what we were going to do... We didn’t have to start with a set of unrealistic aims but what we did, did evolve over time.’ (EL6)

Different Vanguards needed different levels of support from the evaluators in terms of developing their aims. Some evaluators reported being “very involved” in working with the Vanguard in designing the aims and goals of the evaluation and steering the Vanguard in a more realistic direction:

‘[We] worked with them quite intensively to help them design the programme.’ (EL4)

Co-design was possible due to the Vanguards respecting and having trust in the skills of the evaluators:

‘They trusted us, they had confidence in us and put their trust in us.’ (EL6).

The fact that the aims and objectives changed over time was a constant source of frustration for the evaluators. This was described at a local level where the design needed to be tailored to the changing landscape around the Vanguard, and as a result of changes in the demands of the programme nationally.

‘[There were] a lot of changes in terms of how the programme was described so we have had to constantly adjust our evaluation questions, our evaluation design. We have had to constantly adapt. The programme changing, merging, moving that we’ve had to keep up with. A shifting context around the bigger piece has necessitated changes.’ (EL4)

**Relationships with the Vanguards**

A consistent view was that a good ‘communicative’ working relationship between the local evaluation teams and the Vanguard sites was key to success:

‘We developed a good working relationship with the lead director for evaluation.’ (EL11)

One evaluation lead went so far as to describe the relationship between them and the Vanguard as:
‘Energising and refreshing.’ (EL6)

Good relationships and a partnership between different evaluators were also valued:

‘[It] worked so well having a partnership between academic institutions, decent research consultancies and commissioning support units. That is a model I would recommend.’ (EL6)

For many, good relationships meant that practice staff working within the Vanguards understood the importance of the local evaluation, which in turn meant they were then more co-operative with the local evaluation demands. However, this was not the case for all where support for the evaluation and access to participants was challenging:

‘[It was] impossible to get access to the data and to get hold of the staff members.’ (EL2)

‘It’s been a battle the whole way to get any kind of support for our evaluation activities.’ (EL2)

‘There was a massive lack of communication, for example they closed down their intervention and set up another one without us, the evaluators.’ (EL6)

Some decisions needed to go to a ‘transformation board’ but due to poor working relations between the evaluator and the Vanguard programme management they were never given access to this board:

‘They were the ones representing us at this board so we don’t feel we got fairly represented at this board. We were never allowed access to these people, meetings.’ (EL2).

‘[There was] distrust between the two sides.’ (EL8)

It was argued that this distrust was fuelled by a lack of clarity around the roles and responsibilities of the Vanguard and the local evaluation team so that miscommunication occurred, for example ambiguity about who was responsible for patient recruitment.

One evaluation which was initially being carried out by an academic partner was ultimately conducted ‘in-house’ with bespoke pieces of work commissioned from the University, due to specific challenges:

‘They struggled with the difference between research and evaluation… We felt that we were teaching them rather than them teaching us’ (EL10)
Success of the local evaluations

There was a mixed response in terms of the perceived success of the local evaluations both in the broadest sense and also in terms of what they viewed as a successful evaluation, locally. Conceptually, successful evaluations were described as those in which learning had occurred, where relationships were enhanced within communities and where initiatives were sustainable:

‘Success for me for a Vanguard would be, firstly, have we learnt anything from this experiment? The Vanguard programme to me is a set of experiments trying to work out the new models of care. So have [Vanguard site] learnt from this experiment? Yes, they definitely have learnt and revised what they are doing. But secondly, have they made a difference? And I definitely believe they have. And finally, are there better working relationships across the system? and I would say that is also true.’ (EL6)

‘Conceptually, a successful evaluation is the one that makes sense of what is going on…. Sense-making is more difficult in the Vanguard work as it is dealing with a future focussed, developing approach.’ (EL8)

‘Some of the things that were coming through that I felt were more successful were not actually the interventions per se but attempts to respond to early challenges - improving co-production, people were learning and people were trying to design solutions. What I call success has been the much closer working as a result of these initiatives - neighbourhood, locality-based working - that was massive success of the Vanguard and benefitting patients and service users.’ (EL9)

‘Success of Vanguard is when everyone involved with it want it to continue, fighting to keep it going, prepared to put their hands in their pocket to fund it. Also impact on sustainability of primary care, wellbeing of staff, closing beds. Success is when the whole system feels much more sustainable.’ (EL5)

Most evaluators felt their local evaluation was a success but this was described within a context of limited time and resources and a changing environment:

‘Where we have had sufficient evidence we have been able to meet those aims.’ (EL4)

‘I think all parties would agree that the evaluation has delivered the vast majority of what it set out to do…[name of Vanguard site] was more challenging due to the complexity of the landscape in which the Vanguard was operating….We have been unable to deliver the number of evaluations we wanted to.’ (EL4)

‘I think given the budget that we were playing with and the time that we had, we did a reasonably good job… the usual evaluators frustration - with more money you could do more and with more time you could do more.’ (EL7)
Interestingly, there was a recognition that the outcomes of the evaluation may have contrasted with the expectations of NHSE:

“We answered some of the important questions... we did not necessarily answer some of the questions that the NCM team envisaged being answered.’ (EL8)

Data sharing issues

While a number of methodological challenges were discussed, one of the greatest frustrations was around data sharing. Sharing information between health and social care and accessing patient information were substantial challenges faced by the local evaluation teams. There was lack of clarity around who had responsibility for accessing the data and issues relating to ‘gate-keeping’:

‘There were issues over access to primary care data and there were issues with the primary care providers all recording things differently.’ (EL3)

‘Different partners could have shared information better as certain bits of information could have informed and helped other partners.’ (EL3)

New legislation around information governance in February 2017, immediately affected access to the types of data evaluators were able to get. Data agreements had to be frequently re-worked which often took months and put:

‘Incredible pressure on us and on them.’ (EL4).

Experiences with the counterfactual

Generally, the views towards the use of the counterfactual data were negative:

‘We wasted a lot of time on this.’ (EL9)

‘I’m quite opposed to its use, particularly in this area.’ (EL8)

Others remarked on initial positivity towards using counterfactuals that later diminished or offered a more balanced view of both advantages and drawbacks:

‘The word counterfactual was very much in favour at the start but waned quite a bit.’ (EL11)

‘[It is a] robust and structured evaluation concept, however that robustness limited its free-footedness.’ (EL5)
When the interviewees discussed the practical aspects of using the counterfactual, they either questioned the assumptions behind this method or spoke of finding more suitable ways of comparison:

‘We just didn’t have one…we did not get to that point. By the time we got involved assumptions that were being made in the economic modelling were really not well-founded. We started with a series of the analyses that were done [by a US consultancy firm], but not everyone was happy with those initial findings because that modelling very much favoured financial gains for general practice and financial losses to the other parts of the service. A lot of the job was questioning the assumptions and the analyses used.’ (EL7)

‘Our actual position was worse than the counterfactual. [The Vanguard had] quite a bit of local ‘do nothing’ data and so that became more meaningful than the counterfactual.’ (EL3)

**Feeding results back to the Vanguards**

Overall, evaluators spoke of routine feeding back to the Vanguards, with quarterly feedback meetings commonly mentioned, although a small number of interviewees mentioned feedback meetings taking place monthly.

Feedback workshops, open not only to senior management, but to any Vanguard stakeholders were also mentioned. However, some evaluators disclosed that stakeholder meetings were poorly attended and discontinued as a result or that the workshops steered away from their original aims:

‘Originally dissemination events [were intended]. In practice that is difficult to do when stakeholders are already doing new things and attending new meetings; getting people to come to a room is hard. In the end emergent findings were fed back to the Partnership Board and the operational steering group... but there was no big showpiece event...We tried hard to be realistic and be pragmatic.’ (EL7)

‘Workshops would start with a small presentation, but people did not care about the findings and wanted to rant to one another about what did not work.’ (EL8)

Whether feeding back principally to the senior management was the intended format or an alternative adopted following unsuccessful stakeholder workshops, evaluators expressed concern about the findings reaching frontline, operational staff:

‘The idea in principle was that emerging findings would drip through. But it became clear that wasn’t going to work… Decision making and reporting was very hierarchical.’ (EL8)
There was also some unease about the degree to which the findings were summarised and nuance lost:

‘They only wanted pictures as they said their chief executives wouldn’t understand anything else.’ (EL2)

Nonetheless, for the purposes of the evaluation per se, many of the evaluators found feedback meetings – whatever their format – useful in refining evaluation strategies and approaches:

‘At each stage it helped to influence the next step of thinking.’ (EL10)

**Vanguards reaction to the results**

Some evaluation leads found the Vanguards receptive to feedback:

‘The Vanguards have been engaged and very enthusiastic.’ (EL1)

‘Received very well even when [the findings] were negative…able to have quite open and constructive conversations. Good relationships and trust built over time [made] difficult conversations easier to have. There was a collaborative approach to communication around the findings.’ (EL4)

Although one interviewee expressed concern with the keenness to use interim findings before the final report:

‘They would have been quite happy to grab some of the early nuggets and run with them.’ (EL11)

Other evaluators, however, found reactions to findings less positive:

‘They thought they were paying us to give us positive results but in order to maintain our integrity we had to be as neutral as possible.’ (EL2)

In a case where the evaluation was carried out ‘in-house’ but with the involvement of external evaluation partners, emphasis was placed on the benefits of the independent nature of the evaluation in applying for resources:

‘Results have been so exciting especially in these challenging times for the NHS when you have to justify every single penny. It has been brilliant that they were not done by a bunch of NHS commissioners about return of investments. The fact they have been independent findings showing how successful it is has been have been very powerful.’ (EL3)
However, a consistent criticism was that the local evaluators were put under too much pressure to demonstrate the Vanguard sites success in a far too short a period of time. The pressure was felt to be ‘top-down’ from the national team. One evaluator explained they were:

‘Under pressure to demonstrate these improvements were already happening so they could get their next instalment of money to carry on the programme.’ (EL2)

Others stated:

‘We were being slated, basically [for not having an impact]. Especially on emergency bed utilisation... And we were trying to do what's best for the person, not just massaging the figures...what we were being measured on is acute hospital utilisation, not the things that mattered for people, like gaining skills to manage their own help.’ (EL5)

‘Delivery teams were under so much pressure to demonstrate real benefits really quickly with interventions that haven't even been embedded in…and there were no filters to protect us as evaluators from this pressure. The expectations of outcomes were quite exaggerated.’ (EL2)

‘The national expectations were very high and the relationship of having remuneration linked to demonstrating benefits was particularly harmful. That put us under a lot of pressure to demonstrate benefits really quickly which we couldn’t really do whilst maintaining any professional integrity. So we were between a rock and a hard place.’ (EL2)

‘There was a lot of top down pressure in terms of the robustness of the evidence that could be provided over a short period of time with the data that we had. I think it was coming from NHS England to the programme management office. So we were being told to produce a certain amount of evidence from the evaluation and we were very doubtful that we could do what they were asking in the time.’ (EL2).

‘They were probably desperate to get results quicker.’ (EL11)

‘Not much time, not much money but a lot of expectation regarding the outputs they wanted and in terms of the evidence that they wanted to demonstrate early benefit.’ (EL2)

**Feeding back difficult findings**

The interviews further investigated feeding back difficult or unfavourable findings and the Vanguards response to these.
Some interviewees reported a sense of unease in feeding back critical findings, along with some initial unwillingness of the Vanguard to accept criticism.

‘There was reluctance to make a difficult situation even more tricky.’ (EL5)

‘[Criticism] would not stick on them so well at first, but when it became more socially acceptable to voice positives and negatives, people were more inclined to listen.’ (EL8)

One interviewee also spoke of the Vanguard site misconstruing negative feedback:

‘Each time we raised the issue of communication, the ‘Comms’ team got shouted at…the report had to be very clear that the findings meant communication among everyone.’ (EL8)

A number of evaluators also spoke of being asked to reframe the results in written reports:

‘Some Vanguard have asked for partial re-writes and that is an on-going piece of contention.’ (EL1)

However, evaluation leads also spoke of a refusal to change or omit facts and focused instead on contextualising findings, emphasising the limitations of the evaluation, and changing the tone/wording of the reports:

‘There were a few things…that they wanted us not to say, and we really had to say no, this is what the evidence is telling us, we can contextualise it…’ (EL11)

‘It is about describing things in a different way and changing some of the terminology. We didn’t change the integrity of the research but we went along with how they wanted us to describe things.’ (EL1)

‘I was quite happy to take a steer on tone and focus but factual findings I’m not prepared to change’ (EL2)

‘[We needed to] strengthen some of those limitations within the report to avoid the potential for misinterpretation.’ (EL4)

‘[The Vanguard stance was] ‘when so much of the report has been so fantastic we don’t want people getting hung up on one controversial comment’ - so the wording was softened.’ (EL3)

Use of findings by the Vanguard sites

The interviews also explored how the Vanguard sites used the evaluation findings. A divergence of opinion occurred with some Vanguards extensively using the findings and
exploration of experiences of evaluation with evaluation leads

others ignoring them. This remained true for evaluation leads who worked with more than one Vanguard; they found some sites ready to use feedback much more than others.

‘Some suggestions taken on, some ignored [or not implemented].’ (EL9)

Importantly, the use of the findings at different stages was highlighted. For example, some interviewees spoke specifically of Vanguards using evaluation findings – especially early findings – to alter initiatives:

‘Results were used formatively by the team to inform the development of the service and inform the development programme for staff.’ (EL4)

‘In the first year they had a larger number of assisted technology interventions, but these proved difficult to roll out. Following the interim report they pared it back and focused more on a couple of types, where the interim report showed viability.’ (EL9)

Some interviewees also reported evaluation findings informing future sustainability of initiatives:

‘Bringing out [findings] in an independent and robust way really helped the decision-making going forward.’ (EL11)

‘The Vanguard programme based a number of commissioning decisions on the outcomes of the evaluation. A number of services became ‘business as usual’ and a number of services were decommissioned as a result of our evaluation.’ (EL4)

Others questioned the relevance of the findings given the changing nature of the Vanguard initiatives and the overall direction of travel:

‘The evaluation work has been of use to them but things are moving so much, there is so much change in the system that there is a slight concern that the model has changed anyway, they’ve already moved on.’ (EL4)

‘Data would help us to refine where we’re going next, and help us to evidence where we’re going next, but it won’t stop us from going where we’re going next.’ (EL5)

Lastly, the interviewees who reported limited use of the findings so far questioned whether future utilisation of findings was likely:

‘It’s too early to say’ may be a euphemism for ‘it’s never gonna happen.’ (EL8)

Working with the NCM team and level of support

Overall, evaluators had a positive experience of working with NHSE and valued the one-to-one support offered, the availability of the NCM team and the continuity of the relationship:
‘When things haven’t been going our way it has been like ‘let’s sit down and work through this together’, it’s been incredible.’ (EL3)

‘They went about it in the right way, they were very supportive.’ (EL7)

‘When I’ve asked for conversations, I was able to have them.’ (EL8)

‘I think they tried really hard to bring people together and give us a chance to share our issues.’ (EL9)

‘[They were] very supportive and very much open about us doing what was important for us.’ (EL11)

However, this was not always the case and the needs of Vanguards were not always met at a local level:

‘I expected a whole load more contact with the national programme people than we had.’ (EL6)

‘The entire process was built around the health economics. We weren’t really getting support on the local needs.’ (EL8)

Lack of support was in specific areas, namely workforce issues, partnership working and information governance:

‘They were as equally powerless as us in resolving IG issues.’ (EL4)

There was a sense that NHSE focused on monitoring and performance management, which conflicted with the support in some ways:

‘There’s been more of a monitoring relationship with them. They’ve monitored our progress, they’ve asked for reports. There’s been more monitoring than there was support.’ (EL4)

There was also some comment of the availability of the IAU (the Improvement Analytics Unit; a unique partnership between NHS England and the Health Foundation that evaluates complex local initiatives in health care in order to support learning and improvement). IAU’s initial focus was to evaluate Vanguards:

‘They did not visit enough, they were not on the ground with us at all.’ (EL10)
Monthly webinars, community of practice and events

The FutureNHS Collaboration platform (hosted by Kahootz), was not seen as useful due its complexity, the volume of information it contained and difficulties with navigation. It was viewed as a weak platform for understanding emerging issues and cross-Vanguard learning:

‘The Kahootz platform hasn’t been helpful in that it’s very difficult to locate information that you want to find. The other day I went to search for something that I know is on there and still couldn’t find it. I have been disengaged with that platform.’ (EL4)

The monthly webinars, communities of practice and events were mostly seen as useful, especially in terms of providing the national picture, although some would have preferred more ‘face-to-face’ meetings:

‘I’ve been on most of the webinars and most of them have been useful.’ (EL11)

‘I really liked community of practice events.’ (EL11)

Interestingly, one Vanguard was encouraged by their account manager to attend national events, not just as a learning opportunity but also as a way of making the Vanguard visible, politically:

‘[They] gave us advice on which national things would be good to engage with either from a genuine learning perspective or, let’s be honest politically, to make sure your face was seen.’ (EL10)

Sharing of information through webinars and events, led to some frustration around others taking ideas and getting the credit:

‘As we had already started before the Vanguard started, I think a lot of the time for us it felt like sharing what we have done... that we were doing all of the learning and, you know, trailblazing, and all the agony, and then people coming along and pinching your ideas. And they were getting positive feedback and we were still struggling.’ (EL5)

Some sites were overburdened with visits if they were seen as successful:

‘Hundreds of people came to see what we were doing. It was a business in itself, managing all the people who wanted to come and see!’ (EL5)

Regional links could have been stronger. For example, communication with a neighbouring Vanguard was facilitated locally rather than by NHSE.
Overall evaluation of the programme

There was some scepticism around the emerging findings at a national level, and the political messages broadcast as a result of the need to celebrate success. An ongoing challenge was perceived to be around how to embed all aspects of the evaluation, not just the headlines:

‘We are going to have to take the Westminster story with a pinch of salt.’ (EL2)

‘A&E usage is lower in Vanguards than non-Vanguards but underneath all that is a real mis-mash of pictures but I understand that ticks a huge box when you take it up to Parliament.’ (EL11)

‘They can say we spend x million into the new care models and look how much we saved.’ (EL11)

It was recognised that failures also need to be acknowledged through greater openness and critical discussion:

‘When you are trying things out, some things are not going to work. That’s a key success if something has not worked as you know not to try it again.’ (EL8)

In one region, the Vanguards pooled some of their evaluation money and procured a regional evaluation to look at organisations, MDTs and technology. There was a desire to bring the learning together and align Vanguard priorities with those of the STP / ICS (Sustainability and Transformation Partnership / Integrated Care Systems). This was possible as money was not spent early on:

‘[They] had too much money to spend, especially the big ones, MCPs and PACS had a lot of money to spend immediately.’ (EL11)

More time was necessary for the intervention themselves to embed and a longer period of evaluation was felt to be necessary. The timing of the overarching evaluation was also felt to be too late and should have been done in parallel to the local evaluations:

‘It’s a five year forward view, funded for 3 years, expected to show impact after 6 months.’ (EL5)

‘If you were to do something meaningful, something that helps us change health and social care, it needs to have time invested in it. Time more than money in some ways.’ (EL5)

‘[It] needs a much longer evaluation and longer follow up periods as it could all collapse and we wouldn’t know about it.’ (EL1)

‘A lot of us felt it was too late for you to do this [Overarching evaluation].’ (EL11)
Furthermore, there was a suggestion that the programme was divisive in terms of the more robust sites getting more money:

‘[The] broader NHS hasn’t felt the impact of the Vanguard. There is a risk it just becomes a way for already successful and robust sites getting extra funding. It does create complications in terms of how you divide up funds.’ (EL1)

**Sustainability of the Vanguards**

Many expressed concern with the end of the funding period and questioned Vanguard longevity beyond this point:

‘Some of the really great programmes are happening because they have had extra funding and it is hard to see how these will continue once the Vanguard extra funding has gone. And a lot of the developments people have talked about have been about having extra time, space and personnel and without the Vanguard money I can’t see it can be sustained in the same way. The Vanguard programme should have been 10 years rather than three, it didn’t make sense to have it as three.’ (EL1)

‘The model of short term investment… is a challenge to sustainability.’ (EL8)

‘Once the funding finishes if there is no commitment by the partner organisations then it collapses. So a lot more thought needs to be given to sustainability at the outset.’ (EL2)

Evaluators also emphasised the lack of continued national-level support as a main disabler of Vanguard sustainability:

‘[The national team] are not taking things forward and not listening…They do not care about whether this stuff is sensible. They are full of magical thinking.’ (EL6)

Others illustrated that local buy-in was crucial for Vanguard sustainability with examples of where this has worked well and where it was a major reason for discontinuation of (some of) the Vanguard initiatives:

‘It will be sustained. I’m really pleased that in these very challenging times for the NHS, the CCG sees the value of a number of the initiatives and although we can’t fund everything that was piloted at the level it was piloted the CCG has committed a really significant amount of money that will see the sustainability of some of the new services but also some of personal and community resilience work as well.’ (EL3)

‘Strong leadership… is continuing [the initiatives] beyond the life of the Vanguard.’ (EL4)
‘The commissioner... felt on the outside of this process. There wasn’t good working between the CCG and the partners. Some of the services will be sustained through sub-contracting and partner own investment, but everything else is going.’ (EL7)

Some interviewees stressed that the evaluation findings were not a surprise to the Vanguard but rather more of an independent corroboration of findings, and that local buy-in leading to Vanguard sustainability was not strongly related to evaluation outcomes:

‘Funding is finished, but some partnerships continued to meet and some frontline work has been funded by the members themselves...but that is not from seeing a graph or a chart from the evaluation team - that is from living and breathing the experience of new ways of working.’ (EL5)

In light of existing initiatives being decommissioned or reduced, few evaluators spoke of scaling up and spreading the initiatives. However, those who did mention replicating the initiative more widely, suggested caution:

‘You can’t just lift and shift something - contextual messages need to be shared.’ (EL11)

Finally, in relation to Vanguard sustainability, a small number of interviewees suggested reconceptualising what sustainability means. Instead of scaling up, spreading or maintaining the initiatives, it was suggested that sustainability should be seen in terms of applying lessons learnt and emerging knowledge:

‘The Vanguard has done its purpose in trying things out... the sustainability is not necessarily in the Vanguard project, the sustainability is in the knowledge that's emerged from that.’ (EL8)

Learning from the Vanguards feeding in to STPs/ACS/ACOs/ICSs

Interviewees were asked about the Vanguards moving to Vanguards relates to STPs, ICSs, ACS and ACOs (Sustainability and Transformation Partnerships, Accountable Care Systems, and Accountable Care Organisations). These partnerships/organisations cover a considerably bigger geographical footprint than the Vanguards and are expected to develop and agree new strategic plans for improving quality, achieving sustainable [financial] balance and delivering long-term plans. Interviewees considered it too early to tell how learning from the Vanguards relates to STPs, ACS and ACOs. There are links between STPs and the wider health and social care economy but these may not be direct:

‘[There are] some structures in place that will nicely fit with STPs but not sure if we can credit those to the Vanguard programme.’ (EL1)
Governance structures developed at the Vanguards, is a specific example of where learning fed into the larger picture:

‘Locally we are developing....governance structures in line with the national Vanguard governance structure. So the system wide approach to sharing resources and planning has been a massive success of the Vanguard and is being replicated across [the region].’ (EL3)

Input into the STPs and ACS/ICS was highly variable and dependent on key individuals being part of the Vanguard and the STP:

‘The new care models will be part of the evolving STP and ICSs.’ (EL4)

‘They have really embraced the learning from the Vanguards, I say that without any concern at all. It is very much build into their strategic planning for the ICS.’ (EL4)

‘I just don’t think [the Vanguards] are coming up very much on the [STP] radar. The STP here makes mention of the Vanguards but there is no evidence of them doing any work to support rollout and spread.’ (EL7)

‘The same Director who has been the lead for the evaluation of the Vanguard, is the new director for the evaluation in the ICS.’ (EL4)

‘The clinical lead for the care home Vanguard is now the clinical lead for the STP.’ (EL11)

Where the Vanguard did not fit into the footprint of the STP, this journey was much more challenging and there is a perception that there were jealousies related to the Vanguard funding:

‘Our move towards STPs... has been full of problems.’ (EL5)

There was a suggestion that some STPs had unrealistic expectations about what could be achieved through the learning of the new care models:

‘One of the things that I find quite worrying is that the STPs have expectations that a saving is going to be made through prevention or other types of services available in the new care models.’ (EL9)

**Lessons learnt**

Evaluators offered numerous lessons learnt and reflections on what should be done differently if conducting similar evaluations. There was a sense that the remit for the Vanguards lacked focus and was too broad:
‘What people were allowed to do was so broad that it was hard to evaluate…maybe devising a menu of options of projects based on evidence that is already around.’ (EL9)

Timing issues were frequently mentioned. Interviewees particularly stressed the need to plan the evaluation and procure evaluators at the outset of the new initiatives, so that the initiatives could be shaped by evaluation findings as early as possible.

‘Evaluation plans [should be] set up at a start of a project and not woven in half way through.’ (EL1)

‘Had to procure [the evaluation first, which took time] and therefore you weren’t working hand in hand to shape it.’ (EL11)

‘A step between the logic model and the delivery in which evaluation expertise is needed.’ (EL11)

The rapidity of change in newly implemented initiatives was also identified as a challenge, potentially requiring new approached to evaluation:

‘Problems stem when rapid evaluation is aligned with a rapid implementation…The problem with rapid evaluation is that there is not enough lead-in time to get the results right. Everything is expected to happen too quickly and so the implementation team ends up falling over their own feet to keep up and it is a waste of public money.’ (EL2)

‘[There is a] need to develop pragmatic and quick methods of evaluating these new care models because by their nature they are evolving and changing fast and evaluation needs to keep pace with that to be useful. [Then] evaluation can contribute to understanding process as well as impact…making sure the evaluation is informing spread and adoption.’ (EL4)

In relation to this, evaluators also spoke about the need to clearly communicate to the sites what is achievable and what should be focused on at the start of the project and the importance of building relationships to ensure optimal communication about the evaluation approach:

‘Focus earlier on particular aspects and interventions that they [the Vanguard] thought were successful.’ (EL9)

‘Building relationships - there’s no substitute for it. Being honest about the purposes of the evaluation... is it formative, is it learning-oriented, is it generating insights about implementation, and how you make changes and what does it mean for staff and new ways of working; against summative - are we looking at impact, are we looking at all sorts of outcomes? I think having some sort of understanding about the balance between those - and that both are achievable in the times scales available - I think the latter is almost always an issue.’ (EL7)
Having clearer data sharing and information governance guidelines were seen as important considerations, going forward:

‘In future, that is a key lesson for me, that it is really clear about who leads with IG….Being very clear about who holds the risk around IG……The data control arrangements should sit with the commissioner’ (EL4)

There was an identified need to have a longer and more realistic evaluation period:

‘Need a mechanism to capture the longer-term outcomes... People weren’t starting from the same starting point.’ (EL11)

There were recommendations for when funding should be allocated:

‘Change the way NHSE gave them the money - the process......In year 1 there was a lot of wasted money because people had to spend it for the sake of spending it.’ (EL11)

‘The transfer of knowledge did not get a chance to be embedded…I would have put money in at the very start to agree the evaluation and get baselines and data sorted.’ (EL11)

There were also lessons to be learnt around project management and the amount of initial set up time required:

‘Be very clear with commissioners what evaluation is and what they should be looking for...to have a really honest view of what is going on, although I am not sure if it is possible in hierarchical NHS reporting...Spend a lot more time trying to map out with the people delivering the programme what their theory of change was.’ (EL8)

‘Important lessons about the process. Learning from doing things differently. It’s really important to be honest when setting up these type of projects. There was not enough time to deliver, really unrealistic expectations. Rather than overegging the pudding over the achievements, they should focus on the lessons learnt.’ (EL9)

**Summary and Commentary on Key Points**

This section highlights the key points and makes connections from the interviews with evaluation leads to how the successes and difficulties of local evaluations were contextualised, and makes links to Part I of the report.
Challenges

The interviews emphasised the challenges local evaluation teams faced during the process, as well as the different sources of these challenges. As can be seen below, while some tensions occurred within the evaluation teams, others stemmed from the local Vanguards or related to the national programme.

National programme influence:

- Time available for the evaluations influenced the outputs.
  - The procurement process was at times, lengthy, which meant that most evaluations could not start at the same time as Vanguard changes were implemented.
  - Additionally, evaluation leads felt that the period of time the Vanguards were in operation was too short to demonstrate impact especially on the range of outcomes originally envisaged. This may go some way to explain why the evaluators were not able to address the range of questions posed in the tender specifications (part 1).
- National guidance on what the evaluations should focus on (e.g. non-elective admissions) underwent changes after the evaluations had commenced.
- The requirement to use a counterfactual was not seen as helpful (and/or feasible) by most evaluation leads, which echos the conclusions drawn in part 1.
- National collaboration platforms were seen as of limited help to the evaluators due to their complexity, while monthly webinars, communities of practice and events were mostly seen as helpful to the evaluators.
- A tension was noted between a national drive to advertise successes and local need to learn from difficulties and re-configure some interventions.

Local Vanguard influence:

- Local Vanguards were not always clear what their main aims and objectives were, making it hard to determine what could/should be evaluated. This may explain why the questions asked, lacked some local specificity, instead relying on the generic commissioning guidelines (part 1).
- Research questions when they were set out by the Vanguard were often unrealistic and were not developed with the input of the evaluators, especially those that were brought in after the start of the Vanguard. The initial questions required re-negotiation once the research teams were in place. Hence, the needs to deviate from the evaluation strategies as they were originally set out (part 1).
• A lack of data sharing agreements presented difficulties in obtaining data for the evaluations. This is also a consistent theme in part 1.

• Feedback processes were not always effective and some evaluators reported that stakeholder meetings were poorly attended.

• Some sites were utilising evaluation feedback more effectively than others.

Local evaluation teams:

• Evaluators spoke about the need to clearly communicate to the sites what was achievable and what should be focused on at the start of the project.

• The importance of building relationships to ensure optimal communication regarding the evaluation approach was stressed.

• Building relationships was also seen as crucial when sharing unfavorable results.

• Where multiple evaluators (i.e. organisations) worked on different aspects of the same Vanguard evaluation, coordination and communication required improvement in some cases. This is consistent with the findings in part 1 where there was a lack of integration of findings between different elements of the evaluation.

Overall, negotiating feasibility of the commissioned evaluation in light of the changing local Vanguard operations was an ongoing process, changing markedly for some evaluations. Rather than regarding a change in research questions or the extent to which original evaluation plans were executed as a shortfall of the local evaluations, interviews with evaluation leads suggested that in some cases a change in execution may have been an appropriate response to the changes in service delivery and a changing external/national context.

Evaluation Characteristics

There was a range of local evaluation models with some evaluations conducted ‘in-house’ and others using external organisations commissioned by the Vanguards including Universities, Commissioning Support Units (CSUs), NIHR Collaborations for Leadership in Applied Health Research and Care (CLAHRCs), Academic Health Science Networks (AHSNs) and private consultancies. Some used a ‘hybrid’ approach using a combination of internal and external evaluators, or multiple external evaluators for specific pieces of work.

Overall, interviewees felt the evaluations were conducted successfully despite multiple and complex challenges (see above). Evaluators described a mixed methodology using largely qualitative but also quantitative data. Metrics were particularly difficult to obtain, with
qualitative part of the evaluation comparatively easier to conduct according to the evaluation leads. In line with findings from the evaluation synthesis (part 1), obtaining data from patients was difficult for some evaluations. However, as well as struggling to access metric data (e.g. patient satisfaction), gaining qualitative data from patients was also an issue.

**Logic Models**

The interviewees did not comment on logic models extensively. They spoke either about being brought in as evaluators to develop a logic model for the local Vanguard site, or coming at a later point when logic models were already internally developed (the latter appeared to be more common). However, logic models were not mentioned in relation to the latter stages of the evaluation. This was in line with the disconnect of producing logic models and commissioning the evaluation, with little reported value of using them in the evaluation process. One purpose of the logic models, as described in part 1, was to guide the initial development of the Vanguard which perhaps reflects the relative lack of importance of the models over time.

**Questions**

As noted before, research questions outlined in evaluation tenders were not always seen as appropriate, feasible or sufficiently specific to the local Vanguards by the evaluation leads. Also, the evaluation evolved as the Vanguard developed; evaluators re-negotiated the evaluation strategy when they were brought in to make the evaluation more feasible; this in turn influenced which of the originally planned evaluation aspects could be executed. These factors may have contributed towards the synthesis findings in part 1 showing a lack of synergy between the evaluation research questions and findings.

**Data sharing and other information governance issues**

Similarly, inability to obtain some of the data – or obtain it without delay – was mentioned by many of the evaluation leads and attributed to difficulties in setting up data sharing agreements and the need to revise these in line with changing legislation. Issues with information governance had a substantial impact on whether initial evaluation plans could be executed. The way in which data was recorded across providers made combining data for analysis problematic.
Counterfactual

As well as reporting on difficulties to obtain a counterfactual, which was overly time-consuming, many of the evaluation leads were negative/opposed to using a counterfactual. Some respondents also talked about initial, perceived benefits of using the counterfactual, which later reduced when applied practically. Importantly, interviewees either questioned the assumptions behind this method or spoke of finding more suitable methods of comparison.

The quality and execution of evaluations

According to the interviewees, the qualitative part of the evaluation also presented some challenges. Access issues to speak to both professional informants and patients were mentioned. In addition to this, interviewees spoke of local sites not appreciating what qualitative evaluations could achieve when evaluation approach/questions were agreed. However, once obtained, the qualitative analysis was valued by the Vanguards. The execution of the evaluation was strongly enabled by trusting relationships between the Vanguard and the evaluators.

Vanguard sustainability

Evaluators emphasised the lack of continued national-level support and financial backing for the Vanguards as a main disabler of Vanguard sustainability. Whether the Vanguard initiatives were locally re-commissioned often depended not on the Vanguard success – or how well the evaluations captured Vanguard success - but rather on available finances. Local ‘buy-in’ and the ongoing support from clinical commissioners was also seen as important for sustainability.
Independent local evaluation was a key pillar of the evaluation plan for the new care model Vanguard programme. NHS England made significant resource available to individual Vanguards to procure and fund a local evaluation from an independent evaluation partner. This report represents a systematic assessment of the nature and quality of the evaluations commissioned and provides insights from evaluation leads into their experience of the evaluation process and of the Vanguard programme overall.

Our synthesis reveals that most local evaluations sought to address a wide range of questions many of them drawing on the questions set out in guidance from NHS England (2015; 2016). It is understandable that there was a desire to capture as much insight and learning from individual Vanguards as possible. However, the number of questions to be addressed in each evaluation may have been overly ambitious given the resources and time available.

A significant amount of money was spent on commissioning and conducting multiple local evaluations, but each was relatively small scale and context specific. It is perhaps unsurprising that the regionally funded evaluation of the five North East Vanguards represents the most coherent attempt to address a wide range of questions through use of rigorous and transparent methods. Through pooling funding, the North East evaluation had significantly more resources than the other Vanguards. Even then, this evaluation was not without some operational challenges and acknowledged limitations. Each individual North East Vanguard also commissioned additional small scale evaluations to explore aspects deemed locally important but that fell outside of the regional evaluation. This suggests that NHS England’s multi-faceted evaluation strategy including recognition of the need for ‘local evaluation for local people’ was right but that achieving the balance between delivering a detailed understanding of what was working, why, how in each context and local formative insight to shape implementation was harder to operationalise in practice.

Future evaluations of large scale service change should continue to consider multi-faceted strategies but perhaps with less prescription of what should be explored in depth at the local level. If we consider counterfactual analysis to be the preserve of national level analysis then one would expect that any local prescription would be focused around providing causal explanations relating to understanding impact in a given context. Where less local prescription is warranted would be on the formative and or on in depth exploration of local patient experience and or on capturing the costs of local change. The relative success of the combined North East evaluation in surfacing common barriers and enablers of system change offers a meso level of analysis through which generalisable knowledge can be generated. The aims of any evaluation strategy of course need to balance what’s desirable with what’s actually achievable within available time and resource constraints.
Our synthesis and interviews have highlighted a number of common issues across the local evaluations. We summarise our recommendations for each of these as follows:

### Research Questions

The evaluation synthesis found that many of the Vanguards copied most or all of the research questions provided by NHS England as a guide. This was reflected in contributions of the evaluation leads, who reported being given unclear and/or unfeasible research questions during the tender process, which did not sufficiently tap into specific interventions/initiatives of the local Vanguards. The issue may have occurred because local Vanguard teams issuing tenders for the evaluation did not have specialist understanding of research/evaluation methodologies.

**Recommendation 1:** it may be beneficial for evaluators to become involved at an earlier point of new initiative development and/or liaise directly with NHSE to determine research questions that are both locally relevant and meet national requirements.

### Data sharing and other IG issues

Data sharing agreements and information governance more broadly posed a significant barrier to obtaining relevant data (this was especially, but not exclusively true for the quantitative part of the analysis). Interviews with evaluation leads further supported this finding and emphasised that even were solutions were eventually found, information governance procedures created significant delays that subsequently compromised the feasibility of original plans.

**Recommendation 2:** information governance agreements should be in place before the evaluation commences whenever possible; alternatively evaluation plans should consider scenarios where data sharing cannot be established in determining what can be feasibly achieved.
Counterfactual

Many of the evaluations could not obtain a suitable counterfactual. The data synthesis found that few local evaluations had the time, resource (or skill set) to conduct comparative impact assessments on improvements to the quality and efficiency of care, while interviews with evaluation leads demonstrated that most interviewees did not see the counterfactual as useful and favored other comparative methods.

**Recommendation 3:** obtaining a counterfactual may not be possible for local evaluations and may be best placed with national-level evaluations instead.

Quality of research methods

Evaluation leads often focused on the challenges associated with quantitative aspects of the evaluation. However, the evaluation synthesis highlighted that qualitative methods were often poorly executed (or written up in a way that suggests this). Many of the local evaluations did not offer explanations and/or nuanced insights into the Vanguard operation, and did not integrate with other data sources to explain or enhance the credibility of the findings. However, there were difficulties in gaining access to and engaging professionals and service users as research participants which may have limited the depth of analysis.

**Recommendation 4:** evaluations using both qualitative and quantitative methods should integrate findings from both approaches in a more effective way.

Reporting in general

Extracting relevant information from local reports was difficult when synthesizing findings. It was particularly difficult to ascertain what type of interventions were evaluated. However, interviews with evaluation leads revealed that the Vanguard process also lacked clarity on what was a new initiative (i.e. what exactly should be evaluated).

**Recommendation 5:** National programmes may benefit from requesting local evaluators to use structured reporting frameworks ([www.equator-network.org](http://www.equator-network.org)), which may also assist in defining Vanguard initiatives for evaluation purposes at an earlier stage.
References


