Why do evaluations of integrated care not produce the results we expect?

A number of evaluations of different models of integrated care have not produced the expected results of reduced hospital admissions, and in some cases even found people receiving integrated care services using hospitals more than matched controls. Drawing on a wide range of expert opinion, we identify a number of reasons why this may be happening.

Key points

- Problems with the design of the model: these may be poorly designed or lacking logic, targeting the wrong population, or not listening to patient preferences. Models with a large number of complex work streams are a particular hazard.

- The model may be well designed, but fails at the implementation phase when the plan meets the real world. Getting multidisciplinary teams to work well, effectively involving primary care, and the likelihood of uncovering undiagnosed pathology are all issues that have caused problems.
• The expectations of those designing the models about the outcomes that are likely, and the speed that they can be achieved, are sometimes higher than can be achieved in the often short periods allowed for evaluation – not least because of the frequency of delays in implementation. The rather narrow outcomes used, and in particular the focus on hospital utilisation, are also a problem.

• The underuse of process evaluation also means that the active ingredient of models can be missed – this is an important issue where it is hoped to replicate the model.

There are a number of steps that can be taken to address these and other issues we identified:

• **For model design and implementation**: more focus on the problem that is being addressed and the evidence base about what tends to work. Rigorously testing the underlying logic is also important.

• **For commissioners of evaluation**: there is a need for more realistic timescales and openness to different methods and approaches, including more formative evaluation. Seek out help from experts and patients/professionals in the commissioning of evaluation.

• **For evaluators**: Choose outcomes and study designs wisely – consider mixed methods; describe interventions fully; collaborate with others to produce novel work; and use regular feedback loops with all stakeholders.

**Introduction**

Within our own work and elsewhere we have found a number of curious cases in which the data suggests that people receiving integrated care services were using some hospital services more than a carefully selected group of matched controls (Georghiou and Keeble, 2019; Parry and others, 2019; Roland and others, 2012). Studies like these stick out, even in the equivocal evidence base on integrated care. They raise questions about why such unexpected results are still being found decades into the integrated care journey in England, and whether the aims of integration to reduce hospital admissions will ever be consistently achieved.
What are integrated care models and what do they aim to achieve?

Integrated care programmes are invariably multi-faceted, complex interventions incorporating a range of interconnected changes. This can include changes to staff employment, working practices or systems – either by vertically integrating between acute and community and social care or horizontally across acute services (Curry and Ham, 2010). Some models also combine finance, management and governance functions (Baxter and others, 2018).

Most integrated care initiatives aim to improve quality of care and life, improve patient experience and satisfaction, and improve system efficiency (which includes reducing hospital admissions and costs) (Kodner and Spreeuwenberg, 2002). Yet it is widely recognised that the evidence on whether integrated care can reduce community and hospital-based service use or costs is equivocal (see for example: Baxter and others, 2018; Damery and others, 2016; Liljas and others, 2019; Mason and others, 2015). This compounds the difficulty of unpicking aspects of successful models.

In this briefing we explore three hypotheses as to why evaluations of integrated care are producing puzzling results. We then match these hypotheses with practical recommendations to challenge the status quo.

- Could it be that integrated care service models are poorly designed or lack logic?

- Or is delivery the issue? Is it simply difficult for teams to implement complex integrated care initiatives? Do new initiatives lead to staff identifying unmet need in local populations? Or do teams working with intervention groups find themselves being more risk averse to provide a ‘good’ service?

- Or finally, could it be that evaluators are using inappropriate evaluation approaches, that is, ones that don’t detect the change that is happening? Or is it possible that we might be expecting change that is not realistic, and often too quickly, given the nature of the initiatives?
The hypotheses and recommendations are the culmination of ideas gathered from the integrated care literature base and discussions with 50 integrated care experts who attended a seminar organised by the Nuffield Trust and The Health Foundation on 1 May 2019. Attendees represented providers, researchers and commissioners of evaluations. We appreciate their contributions to the event. We also acknowledge our colleagues Adam Briggs, Natasha Curry and Rebecca Rosen for their helpful comments on drafts of this briefing.

**Hypothesis 1: Integrated service models are poorly designed**

Experience seems to show that the underlying ‘logic’ or ‘theory of change’ in the design of integrated care models and policies is sometimes incomplete or makes assumptions and extrapolations that may not be justified by the evidence. In some contexts with complex models of delivery, it may not be possible to assume the sort of direct causation implied by logic models (Baxter and others, 2018). A common problem among many initiatives, not limited to integrated care, is that an intervention aims to reduce hospital admission rates, but is mainly targeted at individuals at low risk of hospital admission, making it less effective than if it had targeted a high risk population (Steventon and others, 2013). Eliciting and testing underlying logics (or programme theories) can address these design flaws (see for example the work of Sheaff and others, 2018).

Some argued at the seminar that integrated models are framed within the biomedical paradigm, and therefore conceptualised as biomedical interventions rather than as complex service innovations that build on community-based assets. Efforts to reduce hospital admissions are often a focus of integrated care initiatives. Reduced hospital admissions is conceptualised as a static, or absolute target, logically connected to the integrated care intervention, whereas in reality admission rates are affected by a range of contextual factors including other health system policies and initiatives. The literature also suggests that models are sometimes designed around organisational requirements, rather than the needs and preferences of the people being served, which often relate to maintaining access and continuity with particular providers (Curry and Ham, 2010; RAND Europe and Ernst & Young LLP, 2012; Sadler and others, 2019).
Another design challenge highlighted was that service models are also often overly complex, involving multiple work streams. Simpler single-faceted integrated care interventions, such as focused falls prevention services managed by a small central team, are known to make more rapid progress (Ling and others, 2012), and could be used to target common root problems.

These misdirected design principles can limit the partners involved and progress made, as well as the potential of initiatives’ positive impacts on patients and service users from the outset.

Hypothesis 2: The delivery of the model as intended is too challenging in real world settings

Integrated care can be difficult to implement as designed, as a delicate balance is needed between behavioural and infrastructural factors (Ling and others, 2012). Some of the specific challenges include: interventions being implemented but failing to have the intended effect, interventions not being fully implemented, interventions being implemented and having unexpected (additional) effects; these are described below.

Multidisciplinary teams, which often form part of integrated care models, are a potential issue. They may have been fully implemented but in reality may fail to deliver. This could be because they are lacking a specification of team objectives and clear role delineation. They could also be experiencing relational difficulties, poor team working and poor communication – the phenomena of ‘pseudo-teams’ in health is a well described problem (West and Lyubovnikova, 2012), but may not be easily visible to evaluators. They could be facing time pressures and lacking adequate training opportunities (Mason and others, 2015; Sadler and others, 2019).

The local and national contexts, including local politics, can also undermine integration efforts, and can mean delivering in a complex system will take longer than expected. The full engagement of primary care in these models is often difficult due to pressures of time and workload, a lack of standardisation of care and systems between practices, and in some cases, a failure to communicate what is required and the objectives of the programme (Ling and others, 2012; SCIE, 2017).
The inherent complexity and biomedical nature of some integrated care models may also be causing unexpected dynamics in the implementation of the model that might account for some of the more surprising findings, such as higher admissions. One such dynamic is that intensive case management models can uncover undiagnosed pathology, or more worriedly, medicalise aspects of life that individuals had been coping with. Even where models are successful at keeping specific groups out of hospital, wider system-level changes may not be detected. This is because it is likely that any spare capacity created by these models will simply allow other patients to be admitted. Recent studies suggest that the threshold for admission has risen over the last few years and therefore there are patients being sent home who might be admitted if capacity was available. ¹ Thus, measuring both intervention-level and system-level impacts is important, but needs real clarity about the purpose of each measure.

Hypothesis 3: The evaluations of integrated care programmes are complex and commissioners of evaluations may be unrealistic in their expectations

Evaluations of these types of services are complex and messy. This is due, in part, to the challenges described above associated with designing and delivering dynamic integrated care programmes, and the context in which they are introduced. This means that there are a number of challenges specific to evaluations that need to be unpicked.

The term integration is quite “elastic” and the literature has many ways of defining it (Kodner and Spreeuwenberg, 2002; Shaw and others, 2011). Yet, integrated care models are often associated with measures that frame them as static, uniform, and transferable. The reality is that a model changes over time, operates at multiple levels, and is context specific. One of the reasons the complexity of models is rarely captured in evaluations is that there is often a lack of corporate memory in design and implementation teams. Descriptions

tend to be high level and over-simplified, making it difficult to capture what really happened and spread learning, while at the same time leading non-participants to overestimate the speed and pace of change and overlook flaws associated with the initiative’s theory of change.

Commissioners of evaluations can often hold unrealistically high expectations for health improvements and cost savings as the outcome of integrated care initiatives (Damery and others, 2016). These expectations are at odds with the trend of marginal gains that are often seen in evaluations, especially in local areas where it is known that usual care (the counterfactual) is already performing well. Even more troubling, some evaluations are tendered with the expectation that they will prove the intervention works, not to find out if it works, thus undermining scientific principles of evaluation research. This optimism bias also leads to unrealistic timelines that do not allow sufficient time to accommodate implementation delays, for changes to occur during implementation and for the full realisation of benefits, which may arise after the intervention’s trial is over (Bardsley and others, 2013; Erens and others, 2016; Mason and others, 2015).

Another challenge is that evaluations of integrated care tend to focus on a limited number of outcomes. It was discussed how outcomes are often set based on the availability of health data and current policy concerns, rather than thinking more broadly on the (intended and unintended) impacts in other sectors linked to health or the other health care services that people would regularly come into contact with. They also raised concerns over the lack of existing measures to capture the impact on patient experiences of services beyond measures of reported patient satisfaction – a gap noted in a recent literature review (Baxter and others, 2018) and blog (Wellings, 2019). The system-wide priority to examine the impacts of health innovations, including integrated care programmes, has set an unhelpful precedent to aim for impact on emergency admissions – even in cases where the programme should not logically have a significant impact on them.

Furthermore, measurable outcomes can be compromised over time by changes in data collection methodology and definitions or by differences in coding practices across areas (Keeble, 2019). For example, the classification of admissions as including zero day cases is notoriously varied across hospital sites. Finding an appropriate, ‘uncontaminated’ control group is also challenging. Integrated care is happening across a number of sites in various forms, therefore locating a group of patients who have no contact with any
form of integration may be difficult. Without reliable controls, cause and effect may be hard to establish and important impacts may not be detectable. Lastly, evaluation questions and designs may be the problem.

Many commissioners of evaluations jump to comparative or controlled research designs too quickly. They underestimate the value in process evaluations and embedded approaches that are able to describe the model, explore the variable penetration and fidelity, and undertake the necessary ‘sensemaking’ before moving on to examining outcomes.

**Where does the problem lie? Can we fix it?**

As might be expected in this complex area, no single hypothesis completely explains the results that have been emerging from a range of different studies.

Integrated service models are inherently complex, requiring inputs from multiple multidisciplinary team members, and cooperation at a number of levels. Their design and delivery are not straightforward. Moreover, we have not made it easy on ourselves in terms of evaluation: we overlook ‘how integration is happening’ and focus instead on measuring ‘what is being achieved’ (or not).

We need to rethink how we evaluate integrated care and move on from the status quo. If we keep commissioning and undertaking evaluations in the same way, we will keep getting the same results, and limit the opportunities for new knowledge.

The next few sections of the paper explore practical advice for both local and national designers of service models, commissioners of evaluations and evaluators developed by participants at the seminar. There is some obvious repetition in messages, which emphasise important points of advice across all stakeholder groups.
Advice to service model designers

1. Focus on getting the basics right first

It is very tempting to focus on shiny new initiatives and models. Instead, planning should start with an understanding of what prevents integration locally and this should be used to get the absolute basics right first. For example, sort out the known barriers associated with data sharing, information technology, and management and administration that can hinder communication between providers and coordination of services.

Carefully define your local problems and health and social care needs that can be addressed through incremental changes or investments. Then think about whether integrated care is the solution, and which known enablers you will need to put in place. When discussing problems, avoid framing them within a ‘hospital admissions’ context, as this can lead to integrated care being suggested as the solution over other options, such as focus on discharge and rehabilitation. Sometimes integrated care schemes may not be the best solution for managing demand for hospital services.

2. Co-design models with patients and professionals

Aim to understand the needs of patients and service users who are visible, but also those who are less visible (Best and others, 2012), as integrated care can mean different things to different people (Shaw and others, 2011). Design the ways of working around patients, and then begin working in the behaviours of the systems and professionals, and then the structure. While many designs seem to put the patient at the centre, there may be limited opportunities to take into account each patient’s history, goals and wider needs, desires for continuity and coordination of services, particularly if there is a focus on narrow success criteria such as avoiding admission. A recent review found strong evidence that integrated care could lead to increases in patient satisfaction, perceived perceptions of quality of care, and patient access (Baxter and others, 2018) – it would be worth exploring with patients

2 See (Gardner and Sibthorpe, 2002; Heckman and others, 2013; Ling and others, 2012) for further examples of barriers.

3 See (Cameron and others, 2014; Local Government Association and Social Care Institute for Excellence, 2019) for examples of enablers.
in co-design sessions whether these are valid and plausible outcomes for the proposed model of care.

The most important drivers of change are the workforce, so it is vitally important that they are engaged, trained and supported – otherwise there is a risk they will not be able to implement the required changes (Best and others, 2012). The seminar heard examples of where staff and GPs seemed to be unaware of the goals of the initiative and were not clear how being part of the programme translated into changes in their practice and daily work. Evidence seems to suggest frontline professionals want more opportunities to feedback and influence integrated care models (Sadler and others, 2019), and where change is forced upon staff delivering interventions they are less likely to support the new activity (Ling and others, 2012). Further, a review of integrated models found their most important elements to include trusting multidisciplinary team relationships and strong understandings and commitment to the model – all embedded in a context of strong leadership shaping the organisational cultural support for the model, flexibility during implementation and having sufficient time (Kirst and others, 2017).

Even with careful design, because these models are complex systems embedded in larger, often equally complex, environments there is significant scope for unintended consequences and scope creep (e.g. people being entered into the scheme who do not meet the criteria). These need to be monitored and corrections made. The potential for tensions between conflicting aims for the model needs to be acknowledged and managed. For example, improved processes and patient experience could increase costs (Singer and others, 2018).

3. Draw on evidence and theory

Look at the existing evidence (including on what does not work) when designing a model or evaluation. Be clear about what you are trying to achieve and write out how you think it will work, for example, in a logic model or narrative. Draw on the evidence to back up your theory of change, and be careful about making claims or assumptions where there is little or weak supportive evidence. Make explicit the links between inputs, activities, outputs and outcomes in your logic model, and be sure to include unintended consequences. Be sure to share this with all key stakeholders (Bardsley and others, 2013).
4. Use local assets and invest wisely

Be clear about the resources available in both health and social care services, including the workforce and local communities. The lack of financial integration has been identified as a major barrier to the success of many integrated care schemes (Mason and others, 2015). Substantial time and resources need to be available to ensure that mutual understanding of the initiative extends beyond the core project team and to ensure clear allocation of tasks. This is especially important where an initiative’s aims are more ambitious and aim to transform the way care is delivered, or where multiple partners spanning primary, secondary and social care services are involved (Ling and others, 2012). Use asset-based approaches and ensure that investment is targeted where it is needed – and borrow from the evidence in doing so. For example, ‘step up’ and ‘step down’ intermediate care and early supported discharge are cost-effective integrated care initiatives, but are not always used strategically.

Advice to commissioners of evaluation

1. Be realistic and transparent

Be realistic in setting evaluation timelines and expectations. Where feasible, use longer evaluation time cycles to allow for change to happen and be prepared to move start and finish dates to allow for delays in implementation. In many evaluations, including the examination of the integrated care pilots, change only starts to become visible around two years post implementation (Ling and others, 2012). Consider commissioning ‘short’ and ‘long’ packages of work, and always commission an evaluation before the change occurs. Be realistic about the limits of methods to answer the research questions definitively. Working with people who have evaluated integrated care first hand can help temper potentially unrealistic expectations.

Be transparent about why you are commissioning an evaluation – is it genuinely to see if an initiative works or is it to prove that it does work or is to learn from how it is implemented. Transparency is important, as evaluations are often commissioned in a policy climate of wanting to show that the solutions work (Bardsley and others, 2013). It is also important to be upfront about how you will use evaluation results if they are unexpected and commit to transparently publishing results regardless of whether positive or negative.
2. Draw in expertise as required

If you are not a specialist in integrated care and do not have experience of commissioning evaluations, it important to seek out advice. The skills required to develop a comprehensive terms of reference and specify the approach are often underappreciated. After careful consideration and consultation, clearly set out in tender documents the research questions, resources available, the purpose of the evaluation, timelines, and intellectual property rights. It may be wise not to over-specify the methods and to remain open to negotiation during the tendering processes with evaluation teams – they may have experiences to draw on that could prove fruitful in improving the evaluation tender. This may mean considering bids that offer a different solution than in the specifications issued to bidders.

3. Commission comprehensive mixed-methods work

Economic analysis and purely quantitative methodologies are often inadequate evaluation approaches due to the complexity associated with delivering integrated care services. Encourage evaluators to describe and understand how the model of integrated care is being implemented, and unpick the local context influencing the service model (Bardsley and others, 2013). Contribute to this process by talking to local evaluation sites about the potential evaluation enablers and access to data before commissioning the research. Carefully consider the emphasis you ask for with regard to formative versus summative evaluation, and where formative work is preferred, consider approaches such as action research.

4. Create an evaluation community to drive meta-evaluation

At a minimum, ensure evaluation teams can prove that they are building on existing evaluations, and encourage a commissioning focus on adding new knowledge on the relationships between processes and outcomes. Encourage evaluation teams to seek out or develop a community of evaluators. This will allow for comparison of integrated care initiatives, reflection and overarching narratives to form (i.e. meta-evaluation).

5. Invest in co-design

Build co-design with patients and professionals into commissioning cycles to understand what they want integrated care programmes to achieve. Use local
sites’ patients and service users to frame evaluation questions and to interpret findings. Be sure to commit sufficient resources to co-production. Consider looking for embedded evaluation capacity within local teams.

6. Request regular formative feedback

Ask for frequent reporting as the study evolves, so that everyone has the opportunity to review the approach and early findings, and make changes to the model and evaluation approach if needed. However, do not get too closely involved in the detail – that should not be the role of the commissioner.

An important finding to get from the evaluation team early on is whether local actors know their role in delivery of the model and its evaluation – this will tell you whether key messages are resonating with the front line. Be mindful of how feedback can create tensions between model designers, commissioners and evaluation teams – work on ways to get through this. If there are local and national evaluations ongoing, think through how the relationships will be managed if the emerging results conflict.

Advice to evaluators

1. Be creative, but logical, in choosing outcomes

Look beyond the aspects of the model that are easy to measure and think about whether you are measuring the metrics that matter and that align with your programme theory. Borrow from models of implementation science to inform your theories.

Work with patients to determine what outcomes matter to them, and build relationships with professionals and clarify their intended outcomes, aims and objectives. Be aware that there may be tensions or conflicts in the outcomes that different stakeholders prioritise. Engage commissioners, policymakers and governments in the evaluation to shift their thinking away from avoided admissions where it is not an appropriate measure. Consider for example, whether clinical markers of health (such as control of HbA1c in people with diabetes), patient-reported health status or care experiences, staff perceptions, or early markers of change might be more appropriate rather than focusing on emergency admissions (Bardsley and others, 2013). Bardsley and others (2013) suggested that “in a number of our studies, we have noted that while a
reduction in emergency admission was not seen, changes in outpatient and elective care were observed – the latter could be early markers of change for emergency impatient care.”

However, where emergency admissions is the right measure:

- Be explicit about how admissions are counted and coded locally and in any potential control areas (i.e. zero day admissions may or may not count);
- Describe in your findings why admissions for certain populations (e.g. older frail people, homeless populations) can be appropriate and necessary; and
- Explore through primary qualitative (and additional quantitative) research whether increases in admissions resulted from identifying and addressing – and potentially over-medicalising – unmet need (or some other reason).

Keep in mind that the counterfactual may be challenging to find – many local sites are undertaking integration of some form. Ceiling effects may also limit the ability to show any improvement, as standard care is often quite good.

2. Use mixed method approaches

The limitations of some of the existing approaches means that it is worth taking risks and using more qualitative and mixed-methods work, including case studies. Primary research is needed to comprehensively evaluate patient and carer outcomes, but be prepared for access to be delayed due to research governance approvals. Case studies can provide valuable learning. Developing knowledge of a particular case within its context can help to identify questions that need to be addressed in other cases. Participatory process evaluations can allow for an exploration of complexity, and co-production of knowledge about the intervention, outcomes and context (see, for example, Eyre and others, 2015).

Think carefully about where a before and after study design will work, and where this study design does not work with the context. Be sure to describe the history and context of the starting point of the evaluation (Best and others, 2012). Follow up on findings over time and use participatory approaches.
3. Negotiate ‘what’ and ‘when’ you are evaluating with commissioners

Do fewer evaluations, but better and deeper building on what is already known. Focus on simpler, low-cost interventions that can reasonably lead to change, which may require a pivot away from evaluating system-level integration. Focus on realistic outcome measures associated with what integrated care can actually achieve. Acknowledge the effect of weak, delayed or changing implementation as a key variable associated with outcomes. Push for the timelines of evaluations to extend for long enough to see impacts.

It is also important to manage negative findings as there are examples of commissioners of evaluations trying to suppress or influence the results of evaluations. Investing in developing an open and trusting relationship with the client during the process is worthwhile.

4. Collaborate with other researchers

Think about doing more collaborations and shared events with other integrated care researchers. Consider undertaking meta-evaluations and going beyond the core offer. A community of practice will help build on others’ approaches. Build collaboration into your timelines and budgets.

5. Disseminate using feedback loops

Aim to provide ongoing, accessible feedback to inform implementation (Best and others, 2012). Focus on what was observed rather than what was not achieved. There is value in describing the work undertaken by the implementation teams who are delivering the integrated initiative (Dickinson, 2014). Articulate the active ingredients (those elements of the intervention intended to lead to change in the outcome, and without which the intervention would be ineffective), major challenges, and quantitative measures, so that the model can be spread. Beware of broad generalisations to other contexts (Bardsley and others, 2013), particularly around active ingredients, as their identification does not guarantee that others will be able to deliver the outcomes in practice.
Conclusions

This area is beset with complexity that seems to emerge from: uncertainty about the nature of the integrated models and how they work, potential lack of agreement on the objectives of integration, local context and the impact of other features such as the quality of internal or inter-organisational relationships and the ability of local leadership. This complexity is not a reason not to evaluate but it does mean that well designed approaches are required.

The seminar that produced these conclusions brought together investigators and people responsible for designing models and commissioning evaluations. There was a striking level of agreement about the need for multiple approaches in this area in terms of the design and implementation of the models, the definition of success, the approach to evaluation and the time required to do this.

NIHR or another appropriate body might consider developing an advisory service to provide support to local systems planning evaluations. Commissioning evaluations is a skilled task and additional help and guidance for NHS organisations could be beneficial.

This event also showed the value of researchers, commissioners and implementers coming together to share insights. If there was more of this, many local evaluations could be significantly improved.
References


McCleary N, Duncan EM, Stewart F and Francis JJ (2013) ‘Active ingredients are reported more often for pharmacologic than non-pharmacologic interventions: an illustrative review of reporting practices in titles and abstracts’, *Trials*, 14(1), 146.


Nuffield Trust is an independent health charity. We aim to improve the quality of health care in the UK by providing evidence-based research and policy analysis and informing and generating debate.