

Evaluating integrated and community-based care

How do we know what works?

Research summary

Martin Bardsley, Adam Steventon, Judith Smith and Jennifer Dixon

June 2013

The NHS is awash with service innovations, often based in the community, which attempt to tackle the challenge of delivering high-quality community-based care for increasing numbers of people with long-term conditions. Despite this we see continuing increases in emergency hospital admissions. Over the last five years, the Nuffield Trust has undertaken evaluations of over 30 different community-based interventions. In many cases we have been tasked with identifying whether service changes have led to a reduction in emergency admissions and the associated cost to the NHS. Using these indicators, the results have been almost overwhelmingly negative. The one exception was Marie Curie Nursing Services for terminally ill patients.

In this paper we outline the main community-based interventions we have evaluated and their impact, and identify nine points that may help those designing, implementing and evaluating such interventions in future. The paper could provide useful learning for the new health and social care integration 'pioneer' sites that will be appointed by the Department of Health by September 2013 (Department of Health, 2013).



Key Points

- Recognise that planning and implementing large-scale service changes takes time.
- Define the service intervention, clearly including what it is meant to achieve and how, and manage implementation well.
- Be explicit about how the desired outcomes are supposed to arise and use interim markers of success.
- Consider generalisability and context: they are very important.
- If you want to demonstrate statistically significant change, size and time matter.
- Hospital use and costs are not the only important impact measures.
- Pay attention to the process of implementation as well as outcomes.
- Carefully consider the best models for evaluation: for example, a light-touch evaluation for an initial phase, proceeding to a more comprehensive evaluation if enough progress has been made.
- Work with what you have: organisation and structural change alone may not achieve the desired outcomes.

In future, formative evaluation methods may offer benefits to sites implementing service innovation in particular methods that:

- exploit the potential of linked data sets, including greater use of GP data to develop cohort-based techniques for tracking the care of individuals with long-term conditions that include analysis of the quality of care, as well as estimated cost and service use
- enable the provision of technical and organisational development advice to sites as part of the evaluation approach
- include regular and intensive tracking of the work of clinicians and managers leading local change and include qualitative techniques that enable close integration with quantitative tracking
- develop robust methods to provide interim reporting of service changes and feedback on observations about process in a way that informs decision-making, enables learning, informs the next stage of service change, and can itself be tracked within the evaluation
- link evaluative methods with organisational development to provide a 'rapid development cycle' offer to innovative organisations in the NHS
- enable learning to be shared between sites within a series of action learning sets or similar activities designed to tackle specific issues, informed by consistent data on performance.

Background

Over the past decade, there has been international interest in improving the management of care for people with long-term and/or multiple health problems (European Observatory, 2012; Wagner, 1996). As a consequence, a range of different approaches to both the organisation and delivery of care for this group of people has been explored. The success of such models is typically assessed in terms of the extent to which they improve the quality of care services, deliver better outcomes for service users and provide services that are more cost-effective (Berwick, 2008). Worldwide, there has been a wealth of approaches but some familiar themes include:

- tools to help patients better understand and self manage their health problems
- technology devices aimed to deliver health care at a distance, such as telehealth
- better coordination of care between professional groups, for example, case management, disease management programmes, virtual wards, care pathways, hospital at home
- efforts to increase personalisation, such as personal health budgets
- integration of services across different health care providers (for example, integrated care pilots in Inner North West London)
- integration of services across health and social care providers (for example, health and social care integration in Torbay Care Trust, North East Lincolnshire Care Trust Plus)
- a range of financial incentives to encourage higher-quality and integrated care and reduced avoidable emergency admissions, including payment reform.

Meanwhile, the NHS is encouraging more innovation in service delivery, in part to reduce avoidable ill health and demand for care, particularly for people with long-term and/or multiple conditions. Many interventions operate in community settings and aim to curb the growth in emergency admissions. For example, in May 2013, the Department of Health announced it will be appointing a number of new health and social care integration ‘pioneer’ sites by September 2013 (Department of Health, 2013). The pioneers will work across the whole of their local health, public health and care and support systems. An invitation for expressions of interest is due to close at the end of June.

However, recent reviews show there is little systematic evidence of what works in terms of community-based alternatives to hospital admissions (Purdy and others, 2012). While there has been innovation and energy in developing new services, the past decade has seen a continued rise in emergency admissions to hospital (Blunt and others, 2010) – a litmus test indicator of the ability to manage chronic disease before it leads to a crisis (Bardsley and others, 2013). In the NHS in England, emergency admissions constitute 35 per cent of all admissions, with a rise of 12 per cent over the five-year period 2004–09 (driven by short-stay admissions); less than half the rise can be explained by population ageing, and the total cost is over £13.5 billion a year (Department of Health, 2012).

The Nuffield Trust is in the position of having evaluated a range of such community-based service innovations in the NHS over the past four years. Our focus has usually been on whether the service model being implemented and evaluated has had an impact on service use and costs, using quantitative methods and summative evaluation. Clearly, there are other important potential benefits from service innovation, such as better clinical

outcomes, improved safety and patient experience. However, below we focus more on evaluation of the impact on use and costs. In many cases our evaluations have included other important elements (for example, patient or staff experience, clinical metrics and ethnographic study) undertaken by Nuffield Trust staff or partner academic organisations.

Our methods have been innovative in that we have typically identified patients or ‘cases’ (anonymously), who have been provided with the new service model, and matched them with ‘controls’ – individuals who are very similar in a number of ways. We have done this using routine data collected by the NHS and a range of case-control matching techniques (Stevenson and others, 2012a). We then compared the service use and costs of ‘cases’ and ‘controls’ before and after the new service model was provided. Without these methods, evaluations risk a phenomenon called ‘regression to the mean’ which in effect means they may show a reduction of admissions that could have occurred in the absence of intervention (Roland and others, 2005). In addition to matching studies, we have also been involved in two large randomised controlled trials of telehealth and telecare (Bower, 2011).

 Our aim is to provide insights from recent service interventions in the NHS that may aid the development of future innovations

In this paper, we summarise what we have learned to date about the community-based innovations and their evaluation. Our perspective comes from our role as evaluators, not as service improvers. Improvement science has yielded many important insights as to how to develop, implement, refine and spread successful service innovation (Powell and others, 2009) that are not necessarily repeated here. Our aim is to provide insights from recent service interventions in the NHS that may aid the development of future innovations and their evaluations.

Community-based interventions and their impact

In total we have evaluated interventions in over 30 different sites. We set out here a summary of some of these projects.

Eight Partnership for Older People Projects

The Partnership for Older People Projects (POPPs) formed a series of 29 innovative projects funded by the Department of Health. They were led by local authorities, in partnership with their local primary care trusts and representatives of the voluntary, community and independent sectors. Their aim was to ‘shift resources and culture away from institutional and hospital-based crisis care for older people towards earlier, targeted interventions within their own homes and communities.’ (Department of Health, 2009).

The Nuffield Trust evaluated a subset of eight POPP interventions selected by the Department of Health (see Table 1). Four of the eight interventions were selected for evaluation by the Nuffield Trust team because they were felt most likely to have had an impact on hospital admissions. These were:

- A programme of support workers who worked alongside community matrons with people with long-term conditions.

- An intermediate care scheme supporting people on discharge from hospital.
- Multi-dimensional integrated health and social care teams.
- Daytime and out-of-hours response services.

The other four interventions aimed to signpost older people towards existing services. All eight POPP interventions are set out in Table 1.

An earlier national evaluation reported positive findings across all 29 POPP sites, in terms of a reduction in hospital bed days (Windle and others, 2009). That study used area-level data to look at the impact across a whole population, and so did not focus on the individuals who actually received the new service model. Therefore, impact detected by the previous evaluation could possibly have been due to other changes made in the areas offering POPP interventions. We were able to work with eight local authorities to link (anonymously) the individuals who had received the intervention to national hospital administrative data.

When compared to matched control patients, we did not find evidence of a reduction in emergency hospital admissions associated with any of the four POPP interventions studied (see Figure 1). In some instances, emergency admissions in the intervention group were higher than in the control group. Only one intervention reduced the number of bed days used by patients.

Figure 1: Effect of interventions on hospital admissions



*denotes statistically significant at the 5% level.

This chart shows the effect of different types of community-based interventions on hospital admissions. It is drawn from *An Evaluation of the Impact of Community-based Interventions on Hospital Use* by Adam Steventon, Martin Bardsley, John Billings, Theo Georghiou and Geraint Lewis, published by Nuffield Trust.

Table 1: The eight POPP interventions examined in the study

Intervention	
A	Support workers working under the direction of community matrons with people with one or more long-term conditions who were felt to be at risk of deterioration or were unstable. Support workers provided personal nursing and social care.
B	Intermediate care service with generic workers, which supported people on discharge from hospital.
C	Integrated health and social care teams configured around primary care teams, which focused on people with one or more long-term conditions.
D	Out-of-hours response service and daytime response service, both consisting of an integrated team comprising community alarm services, mobile wardens, generic workers, district nurses, paramedics and community psychiatric nurses.
E	Volunteer-run assessment and signposting service. Volunteers made contact with older people, carried out a home-based 'check-up', and provided advice on benefits entitlement, housing, community transport, education and leisure activities. If necessary, the volunteer acted as a personal navigator through the range of services available.
F	Short-term assessment and signposting service, which targeted older people living in the most deprived areas. A multi-agency team signposted a range of services, including health, housing, social care, benefits and community development.
G	Short-term assessment and signposting service, which involved staff visiting clients in their own environment. This initiative used the single assessment process to signpost and commission from a pre-agreed menu of community services, or referred clients to specialised services.
H	Short-term assessment and signposting service, which aimed to improve access to low-level preventive services by establishing a single point of access. Joint prevention teams consisted of health advisors, health trainers, social care workers, link workers, a team coordinator and volunteers.



Department of Health-funded integrated care pilots

This programme of integrated care pilots was a Department of Health-funded initiative. It aimed to explore different ways of providing integrated services to help drive improvements in care and wellbeing – the pilots were evaluated from 2009 to 2011. Organisations across England were invited to put forward approaches and interventions that reflected local needs and priorities; 16 were chosen for participation and a mixed methods evaluation led by RAND looked at these pilots (RAND Europe and Ernst & Young, 2012). As part of this, the Nuffield Trust undertook the analysis of the impact on hospital activity in 11 of the sites where a person-level analysis of changes in hospital admissions was felt to be most relevant. These 11 pilots proposed very different interventions, as Table 2 summarises.

Table 2: The 11 national integrated care pilots evaluated by the Nuffield Trust

Pilot site	Main integration focus/client group
Bournemouth and Poole	Structured care for people with dementia
Church View, Sunderland	Case management for at-risk older people
Cumbria	Case management for at-risk older people
Nene (Northamptonshire)	People at risk of admission to hospital (long-term conditions)
Newquay	Structured care for people with dementia
Norfolk	Case management for at-risk older people
North Tyneside	Comprehensive assessment of people at risk of falls
Northumbria	Case management for people with severe chronic obstructive pulmonary disease (COPD)
Principia, Nottinghamshire	Case management for at-risk older people and COPD
Tameside and Glossop	Primary and secondary prevention of cardiovascular disease
Tower Hamlets	Structured care for people with diabetes

The wider evaluation of all 16 pilots (Ling and others, 2010; RAND Europe and Ernst & Young, 2012) found that the integrated care interventions led to process improvements such as an increase in the use of care plans and the development of new roles for care staff. However, surveys of patients and service users indicated that they found it more difficult to see the doctor and nurse of their choice, and they reported being listened to less frequently and being less involved in decisions about their care. A central aim of many pilots was to reduce hospital utilisation, but over a six-month period of the pilots we found no evidence of a general reduction in emergency admissions. However, there were reductions in planned admissions and in outpatient attendances for some interventions that involved case management using multidisciplinary teams (Roland and others, 2012). We also found no overall significant changes in the costs of hospital use (the costs to the commissioner), but for the pilots trialling better case management there was a statistically significant net reduction in combined inpatient and outpatient costs. Thus, reduced costs for elective admissions and outpatient attendance exceeded increased costs for emergency admissions.

Whole systems demonstrator trial of the use of telehealth and telecare
Telehealth and telecare are just two examples of assistive technologies that are widely viewed as important opportunities to improve the management of care for people with long-term conditions. Though there are many people and groups advocating greater use of telehealth technologies, the evidence of its impact on patient care is ambiguous to date, mainly because of the proliferation of very small pilot studies. In 2006, the Department of Health proposed three large ‘whole systems demonstrator’ pilots that would implement assistive technologies at scale as part of efforts towards achieving more integrated care. Three areas were selected – Newham, Kent and Cornwall.

“ As the project developed, a large randomised controlled trial was put in place that compared telehealth and telecare with usual care

As the project developed, a large randomised controlled trial was put in place that compared telehealth and telecare with usual care (Bower, 2011). In the trial, telehealth was delivered for patients with long-term health conditions (diabetes, heart failure or COPD) and involved the remote transmission of vital signs (such as blood glucose levels) from the patient to a health care professional. The trial was pragmatic and compared patients from GP practices randomly assigned to either intervention or control across the three areas.

There were indications that use of the telehealth equipment was associated with a reduction in emergency admissions (Steventon and others, 2012b). However, this appeared to have been linked to an unexpected increase in admissions among the control group, rather than a reduction among telehealth patients. Further, changes in service use were not significant with respect to costs, while the intervention itself was very expensive. A formal cost-effectiveness analysis (Henderson and others, 2013)



Download the report from this evaluation at: www.nuffieldtrust.org.uk/publications/impact-telehealth-use-hospital-care-and-mortality

showed that the technologies used in the whole systems demonstrator trial were very unlikely to be cost-effective.

Moreover, studies of patient wellbeing showed no benefits associated with these technologies (Cartwright and others, 2013). Although there is some emerging evidence that some patients do benefit from telehealth, it is still not clear for which patients and under what conditions the technology might be most effective. The situation is made more complicated by the fact that the technologies themselves are evolving and associated service delivery models vary considerably.

In the trial, telecare was used for patients with social care needs and included falls sensors and bed occupancy sensors that were automatically linked to 24-hour monitoring centres. Analysis of the impact of telecare technologies also yielded negative results, with no reductions detected in the number or length of hospital admissions or in the number of admissions to care homes (Stevenson and others, 2013) despite the claims that are often made about such technology.

Virtual wards in three sites

A 'virtual ward' is a way of organising community-based teams. It involves using a predictive model to identify patients at high risk of emergency admission, then providing targeted preventive care at home using a multidisciplinary team. This targeted care intervention uses the staffing, systems and daily routines of a hospital ward to monitor and care for patients.

We have looked at a number of different variants on the virtual ward model and evaluated three sites in England operating within the periods 2007–11 (Lewis and others, 2011a; 2012). Each site had implemented the virtual ward model differently. One site, which contributed the majority of data for our evaluation, offered multidisciplinary preventive care only during a short initial pilot period, with most patients then receiving 'standard' case management from a community matron. Some of our evaluation was hampered by the fact that the virtual wards were relatively new and the volume of cases seen by the new system was not sufficient to provide a statistically significant measure of change. In another case, an established virtual ward model of care had significantly changed from the time of its inception.

 Over the three sites, we found no evidence of the expected reduction in emergency hospital use

Over the three sites, we found no evidence of the expected reduction in emergency hospital use. Nevertheless, a reduction was found in outpatient attendances and elective admissions. There was no clear explanation for why these reductions had been achieved, as they related in large part to admissions for a condition that was not targeted by the virtual wards (cancer). However, it might have been due to several factors including better care coordination, substitution of care, or better informed patient decisions. In the other sites, the numbers of patients who received the intervention during the study period were insufficient to determine significant effects. The study will be published during 2013.

Other integrated care schemes – Inner North West London and Trafford
We have worked with two sites that have attempted to develop more widespread integrated care across a whole health care economy.

The Inner North West London Integrated Care Pilot is a large-scale innovative programme designed to improve the coordination of care for people over 75 years of age, and those of all ages who are living with diabetes. The pilot, which started in June 2011, was originally intended to run until July 2012, but was extended to March 2013. A team of researchers from Imperial College London and the Nuffield Trust carried out an evaluation of the first year of operation of the new integrated care programme (Nuffield Trust and Imperial College London, 2013).

This evaluation revealed that in its first year, the pilot made substantial progress in designing and implementing a highly complex intervention, and in underpinning this progress with sophisticated governance arrangements and new financial incentives. Staff and professional satisfaction with the intervention was positive, and researchers concluded that the programme was well set for further progress in improving the coordination and delivery of care in years two and three (Curry and others, 2013).

However, the pilot was still in its early stages at the point of evaluation and it was too soon to demonstrate benefits with respect to service use and patient outcomes. After examining a cohort of patients from the first three months of the pilot our analysis showed no evidence of a reduction in emergency admissions (Bardsley and others, 2013).



We tracked and supported Trafford over a two-year period, using a range of qualitative methods

In Trafford, unlike the other sites noted above, we were not commissioned to evaluate the impact of integrated care initiatives on service use and costs. Rather we tracked and supported Trafford over a two-year period, using a range of qualitative methods to help develop the local plans. One key message here was the significant length of time needed to design and agree change across a range of organisations and the significant impact that contextual factors (such as changes in the management of a local provider trust, the departure of key personnel and wider NHS structural reform) had on implementing the desired changes to local care delivery (Shaw and Levenson, 2011).



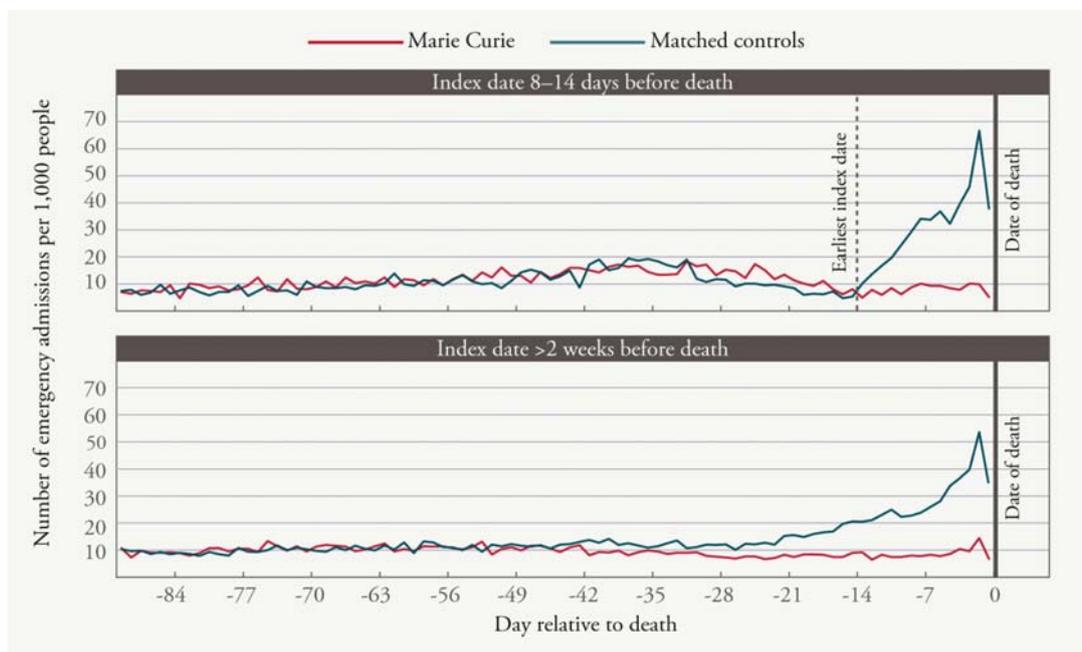
Download the reports from these evaluations at: www.nuffieldtrust.org.uk/publications/evaluation-first-year-inner-north-west-london-integrated-care-pilot
www.nuffieldtrust.org.uk/publications/towards-integrated-care-trafford

Marie Curie nursing service at the end of life

The Marie Curie nursing service is an established programme that delivers support to over 30,000 people a year at the end of life. The service is commissioned by the NHS and provided by Marie Curie and involves dedicated nursing support in the last weeks of life. The Nuffield Trust looked at whether this service was helping more people to die at home (and by implication in the place of their choosing) and whether it led to a reduction in unplanned emergency hospital use.

By using innovative data linkage and matching methods we were able to compare the patterns of care for 30,000 people receiving the Marie Curie service with a similarly matched control group. We found that the Marie Curie patients were significantly more likely to die at home and had significantly less emergency hospital use than controls (Chitnis and others, 2012; 2013). Moreover, the average costs of all hospital services used by individuals receiving the ‘intervention’ was lower, particularly so for patients who had been receiving home nursing for longer – though this was in part offset by the cost of the service. This intervention represented a well-established and widely used model of care service delivered in a standard way.

Figure 2: Number of emergency admissions per 1,000 people by day over the last three months of life



These charts illustrate the impact of the Marie Curie Nursing Service on emergency admissions in the last three months of life, at various index dates. (The index date is defined as the date of the first Marie Curie nursing visit for intervention patients, and the corresponding date for matched controls.) They are drawn from a Nuffield Trust study commissioned by Marie Curie Cancer Care *The Impact of the Marie Curie Nursing Service on Place of Death and Hospital Use at the End of Life* by Xavier Chitnis, Theo Georghiou, Adam Steventon and Martin Bardsley.



Download the report from this evaluation at: www.nuffieldtrust.org.uk/publications/marie-curie-nursing

Evaluation: points to consider

The service models evaluated to date generally appear not to be associated with reductions in emergency hospital admissions (Purdy, 2010; Purdy and others, 2012). This has also been the conclusion of others with respect to community models of care for frail older people (for example D'Souza and Guptha, 2013).

Rossi observed over several decades the failure of complex social programmes in the United States to show impact (Rossi and others, 2004), concluding that there were three main reasons for this:

- Faults in problem theory: the programme is built upon a faulty understanding of the social processes that give rise to the problem to which the social programme is ostensibly addressed.
- Faults in programme theory: the programme is built upon a faulty understanding of how to translate problem theory into specific programmes.
- Faults in programme implementation: there are faults in the organisations, resources levels and/or activities that are used to deliver the programme to its intended beneficiaries.

From our experience of evaluating new service models, we offer the following points to consider, which may be helpful to those developing and implementing innovations and thinking about how to assess impact or commission a formal evaluations.



Our work in both Trafford and Inner North West London demonstrates that bringing about change in health systems is extremely challenging and complex

Recognise that planning and implementing large-scale service changes takes time

Our work in both Trafford and Inner North West London demonstrates that bringing about change in health systems is extremely challenging and complex and requires significant time and resources. This is consistent with established evidence about how to make large-scale change in health care (Best and others, 2012; McNulty and Ferlie, 2002). Indeed, many of the models that are providing inspiration to the NHS currently (for example, Kaiser Permanente and Geisinger in the US) have been established for many years, if not decades, with stable leadership over that period. Development of both the intervention and evaluation takes time, and one year of operation is unlikely to show much result beyond the process of initial set-up and implementation.

As we noted for Trafford in 2011:

Although the basis for change was secure in Trafford, a leap of faith was necessary to invest in implementation, given the evidence underpinning integrated care. The Trafford experience suggests that two years of initial development, followed by a minimum of one year of 'live' working, and almost certainly longer, is required to show the initial effects of major changes to service organisation and provision, particularly financial savings. (Shaw and Levenson, 2011).

As we have observed in many cases, the investment in a new pilot service brings with it an expectation on the part of sponsors of the pilot that change will be detectable in a short period, say a year, and moreover that such evidence will provide a sound financial basis for the change. In such circumstances there is typically significant pressure to show a positive evaluation result in terms of activity and costs in an unreasonably short time period. Indeed, in the first year, or even two, changes to structure and process are what might reasonably be expected, and not a significant impact on service use and outcomes.

Define the intervention clearly and what it is meant to achieve and how, and implement it well

There were times in our evaluations when it was not clear to us what the intervention was (or the new service model) when compared to ‘usual’ care. This is because interventions may not be well defined, may not be completely or intensively implemented, there may be no clear start date, no clear aims for the project nor measures of success, no clear patient group to receive the new service model, and the interventions may develop (or regress) into other service models, not least if there are changes in staff. Clearly, services are dynamic and if problems emerge they must be adapted, but the basic parameters of the new service model should be properly defined, well implemented and with clear measures set out from the start (and, ideally, with such measures developed in partnership with the evaluators). Any changes need to be made according to a clear logic rather than by chance. This takes constancy of purpose, firm leadership and good project management, all of which should not be underestimated.

One aspect that is particularly important to specify relates to the selection process that governs which patients will receive the intervention and which will not. As well as explicit criteria, there may also be implicit criteria used by health care professionals, for example, in judging which types of patient might be most likely to benefit, or a patient’s preferences for certain forms of care. The way in which criteria work is fundamental to good evaluation design (Rubin, 2010). If it is not clear why some patients received the service and others did not, the evaluation results can be misleading and lead to suboptimal decisions about the future development of services.

“ It is our experience that exactly how a new service model is meant to reduce emergency admissions is often vaguely defined

Be explicit about how desired outcomes will arise, and use interim markers of success

It is our experience that exactly how a new service model is meant to reduce emergency admissions is often vaguely defined. For example, the mechanism by which telecare was supposed to lead to a reduction in admissions does not seem to have been clearly stated, especially given the potential for falls detectors to signal increased need, which can (perhaps often appropriately) lead to an emergency hospital admission. Although national patterns of service use suggest that increased social care spending is associated with lower hospital spending (Bardsley and others, 2012; Forder, 2009), the reasons for this are still unclear.

It is important for those developing a new service model to make explicit at the start the process by which it is meant to have impact, including what the main ‘drivers’ of that process will be, and to test assumptions out with all key stakeholders. Three critical factors which are often underestimated are the need for: first, support by and engagement of general practitioners, who often make the decisions about whether or not to refer patients into these services; second, effective project management; third, a long period of time for the intervention to recruit enough eligible patients and show impact.

It is important to consider the shorter term effects and changes in process that happen with a service intervention before the longer term outcomes appear. 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 inpatient care. Other interim markers might be clinical, such as improved disease control, or reflect patient measures such as patient activation scores.

Generalisability and context are important

It is important to ask how legitimate the findings of one study are in respect of transfer to another setting. Thus, for example, the whole system demonstrator trial of telehealth showed disappointing impact on hospital costs. Does this mean telehealth technology simply does not work and is not worth considering further? Not necessarily. The results of the whole system demonstrator trial in part reflect: the cases that were selected; the way telehealth devices were used and embedded within a local service; and the effectiveness of that local service. Furthermore, the results were based on a these factors in place three to five years ago. The results can perhaps signal that we should be cautious about some of the claimed benefits of the technology and think more about the wider set of services provided alongside the technology.

This caution can also work the other way – an evaluation of a new service model may show some impact in one area but not in another context. This is because the service provided may be less effective in the new context, or because the intervention is provided to a group of patients who are unlikely to benefit. A common problem is that the intervention which is hoped to reduce hospital admissions rates will not be effective if it is mainly targeted at individuals at low risk of hospital admission (Steventon and others, 2013). This is one example of a phenomenon observed by others that the spread of an innovation is associated with a gradual diminution in its effectiveness, or ‘Rossi’s iron law of evaluation’ (Rossi and others, 2004).

Put more simply, $\text{intervention} + \text{context} = \text{outcome}$ (Pawson and Tilley, 2004). A recent review of studies of over 80 large-scale changes in health care (Best and others, 2012) identified five ‘simple rules’ that were likely to enhance the success of such initiatives:

- blend designated leadership with distributed leadership
- establish feedback loops
- attend to history
- engage physicians
- include patients and families.

If you want to demonstrate statistically significant change, size and time matters

Teams developing new services often face a ‘chicken or egg problem’, as clinicians may be unwilling to refer patients into the service before there is evidence of an impact, but it is not possible to produce robust evidence before many patients have been referred. Indeed, there is a very basic question of how many patients are needed to show positive results if they occur. If the result desired is an impact of a 20 per cent reduction in emergency admissions, then an evaluation aimed at assessing whether the intervention was effective at achieving this size of impact will require a sample size of around 1,200 to 1,500 patients. However, in our experience, plausible effect sizes may be much smaller than a 20 per cent reduction in admissions, meaning that data on even more patients will be needed.

Part of the solution to this problem is in defining the intervention carefully with support from all stakeholders, developing a realistic change model and interpreting the existing evidence carefully taking generalisability into account (the last three bullet points above). There is often a temptation to set up a service to cover a wide area to serve this number of patients quickly, but a common finding is that in the rush to find enough patients as quickly as possible, eligibility criteria can loosen. This might reduce the effectiveness of the intervention if the wrong people are targeted. It is, therefore, important to risk-stratify the population to identify those individuals at highest risk of admission to hospital, and target these people appropriately using the service model to be piloted (Lewis and others, 2011b). For services where throughput is likely to be low, different metrics (other than a 20 per cent reduction in emergency admissions) will be needed to demonstrate change for the better. In all such studies, parallel qualitative study of the experience of staff and users can help to identify what is working or not, and why.

“ The importance of reducing avoidable hospital costs means that it attracts a high degree of policy attention and profile

Hospital use and costs are not the only impact measures

Many of our summative evaluations at the Nuffield Trust have focused on emergency admissions, hospital use and costs because these have been seen to be some of the key desired indicators of impact. While an avoidable emergency admission is clearly undesirable for patients (and taxpayers) there are also many other legitimate outcome indicators that might be used. These include changes in clinical markers of health (such as control of HbA1c in people with diabetes), patient-reported health status or care experiences, and staff perceptions. These markers of success may be achieved earlier than the desired change in hospital use and can add important dimensions to studies such as our analysis of the commissioning of care for people with long-term conditions (Smith and others, 2013) as well evaluations in Inner North West London, the national integrated care pilots and the whole-system demonstrator programme.

The importance of reducing avoidable hospital costs means that it attracts a high degree of policy attention and profile. Indeed, a target of reducing admissions is often a prerequisite for funding and support for pilots. As a consequence, we have observed that these targets may be included in business cases even though they are not plausible. In some cases, those related to hospital costs might be related to the primary aims of the

people developing and running the pilots, who may be more interested in health improvements or in the long-term strategy for transforming care. The result is that pilots comprise interventions that are more likely to influence outcomes other than admissions (at least in the short term), but are evaluated primarily in terms of impacts on admissions. It is generally easier to evaluate the impact of interventions on hospital use than other outcomes, as data are readily available. A more honest approach towards specifying impacts from the outset might help with the long-term sustainability of schemes.

“ Part of the reason why community-based service interventions may not have demonstrated impact is because the strength of implementation of the intervention is thought to be 90 to 100 per cent, but in fact it is under 30 per cent

Pay attention to the process of implementation as well as outcome

Part of the reason why community-based service interventions may not have demonstrated impact is because the strength of implementation of the intervention is thought to be 90 to 100 per cent, but in fact it is under 30 per cent. This may be due to a number of factors, for example the barriers to implementation have not been fully understood, implementation and project management are poor, the intervention itself is ill defined, or the wider context has changed. Paying attention to, and evaluating, the process of the intervention is important to understanding why the desired changes are not occurring, and in order to modify the intervention as appropriate. Hence, it is necessary to consider a mixed-method study that tracks the process of implementation, as well as the outcomes.

One important element of the process of implementation is developing the wider ownership for a service change among the people and organisations that may be affected. This was seen in Inner North West London, where the evaluation process itself helped to identify key lesson for the future stages of the pilot (Nuffield Trust and Imperial College London, 2013).

Carefully consider the best models for evaluation

As we have indicated there are many different ways to do evaluations of complex interventions such as integrated care and the most successful recipes use a mix of methods (Goodwin and others, 2012). However, we also note that many complex interventions we have evaluated took time to develop. Doing a full blown evaluation from Day 1 may not be appropriate. Pilots may take longer to develop their plans, to overcome the barriers needed to progress the project and enable it to function smoothly. Evaluating some of these pilots from Day 1 thus risks: wasting resources on projects that are not ready (and may never be functioning as intended); demonstrating a negative end result (purporting to show the intervention did not work, but in fact the intervention was never at any strength or working as intended); demotivating staff who may have tried their best; and dampening down similar efforts at innovation elsewhere.

Thus, we suggest that pilots of complex interventions should have a bedding-in period, during which there may be some light monitoring and assessment of the process of implementation, followed by an assessment as to whether the pilot merits full blown evaluation (Leviton and others, 2010). For many of the initiatives we have been asked to address a question of ‘has it worked?’ The assumption is that complex service innovations

have a binary outcome and that future investment (or disinvestment) hinges on a positive result. While there is a clear place for well-structured summative evaluations of this type, they risk missing some of the more subtle effects. A pilot may work partially and learning from that partial success is important. Or it may be, of course, that the beneficial outcomes were not measured or were measured too early.

As well as considering summative evaluation (a method of judging the worth of a pilot at its end, or at a fixed point in time) there is value in considering formative evaluation (a method using feedback of evaluation findings to the pilot to modify the pilot activities as they develop). Formative evaluation can help pilot activities to evolve appropriately and give early and ongoing feedback on progress, which may help motivate staff. In realistic approaches to evaluation (Pawson and Tilley, 2004) even the description of programme objectives may change and evolve during the course of a formative evaluation.

Work with what you have: organisation and structural change may not deliver outcomes

Developing service models to integrate care for patients does not necessarily mean developing integrated organisations: in many cases such structural change can absorb a significant amount of senior management time. It is important to be clear about what changes are intended to happen on the ground and when (Curry and Ham, 2010). The concept of integrated care take many different forms, and our understanding of models of integration has evolved over time, for example, from care coordination and care planning in the 1980s, to shared decision-making or inter-professional working more recently (Shaw and others, 2011).

The way ahead

To date, the apparent limited success of the various community-based service interventions we have studied is disappointing, despite the good intentions of those driving change and the energy expended. We are also aware of the impatience in the system to show changes on what may be an unrealistically short timescale.

The above nine points may increase the chances of success of future service innovations. But we have been struck by the approach used by the Center for Medicare and Medicaid Services Innovation Center in the USA to trialling pioneer accountable care organisations. Key features are that:

- Payment and objectives for service delivery changes are co-produced by the centre and sites over a considerable period of time (at least 12 months) prior to implementation. During this period, a change model is developed which specifies how the interventions developed could plausibly result in the desired outcomes.
- The evaluation approach is crafted in parallel to this development of objectives, payment and other features. It includes both quantitative and qualitative elements.
- The formative aspect of the evaluation (including regular feedback to sites about site-specific and overall programme progress) is combined with a high degree of technical assistance, designed to monitor fidelity to the intervention, increase performance and share learning between sites.
- A summative evaluation, conducted at the end of the pilot, feeds directly into future policy decisions about the national roll-out of changes across Medicare.

For future work we believe the key will be in developing formative evaluation methods that:

- exploit the potential of linked data sets, including greater use of GP data to develop cohort-based techniques for tracking the care of individuals with long-term conditions that include analysis of the quality of care as well as estimated cost and service use
- enable the provision of technical and organisational development advice to sites, as part of the evaluation approach
- include regular and intensive tracking of the work of clinicians and managers leading local change, and include qualitative techniques that enable close integration with quantitative tracking
- develop robust methods to provide interim reporting of service changes and feedback on observations about process – in a way that informs decision-making, enables learning, informs the next stage of service change and can itself be tracked within the evaluation
- link evaluative methods with organisational development to provide a ‘rapid development cycle’ offer to innovative organisations in the NHS
- enable learning to be shared between sites within a series of action learning sets or similar activities designed to tackle specific issues, informed by consistent data on performance.

References

- Bardsley M, Blunt I, Davies S and Dixon J (2013) 'Is secondary preventive care improving? Observational study of 10 year trends in emergency admissions for conditions amenable to ambulatory care', *BMJ Open* 3 (1). pii: e002007. doi:10.1136/bmjopen-2012-002007.
- Bardsley M, Georghiou T, Chassin L, Lewis G, Steventon A and Dixon J (2012) 'Overlap of hospital use and social care in older people in England', *J Health Serv Res Policy* jhsrp.2011.010171.
- Berwick (2008) 'The triple aim: care, health and cost', *Health Aff* vol. 27 no. 3, 759–69.
- Best A, Greenhalgh T, Lewis S, Saul J, Carroll S and Bitz J (2012) 'Large-system transformation in health care: a realist review', *Millbank Quarterly* 90 (3), 421–56.
- Blunt I, Bardsley M and Dixon J (2010) *Trends in Emergency Admissions in England 2004–2009: Is greater efficiency breeding inefficiency?* Nuffield Trust.
- Bower P, Cartwright M, Hirani SP, Barlow J, Hendy J, Knapp M, Henderson C, Rogers A, Sanders C, Bardsley M, Steventon A, Fitzpatrick R, Doll H and Newman S (2011) 'A comprehensive evaluation of the impact of telemonitoring in patients with long-term conditions and social care needs: protocol for the Whole Systems Demonstrator cluster randomised trial', *BMC Health Services Research* 11, 184. eScholarID:128699 DOI:10.1186/1472-6963-11-184.
- Cartwright M, Hirani SP, Rixon L, Beynon M, Doll H, Bower P, Bardsley M, Steventon A, Knapp M, Henderson C, Rogers A, Sanders C, Fitzpatrick R, Barlow J and Newman SP (2013) 'Whole systems demonstrator evaluation team. Effect of telehealth on quality of life and psychological outcomes over 12 months (Whole Systems Demonstrator telehealth questionnaire study): nested study of patient reported outcomes in a pragmatic, cluster randomised controlled trial', *BMJ* 26;346:f653. doi: 10.1136/bmj.f653.
- Curry N, Harris M, Gunn L, Pappas Y, Blunt I, Soljak M, Mastellos N, Holder H, Smith J, Majeed A, Ignatowicz A, Greaves F, Belsi A, Costin-Davis N, Jones-Nielson J, Greenfield G, Cecil L, Patterson S, Car J and Bardsley M (2013) 'Integrated care pilot in North West London: a mixed methods evaluation', accepted for publication *Int Journal of Integrated Care*.
- Curry N and Ham C (2010) *Clinical and Service Integration: The route to improved outcomes*. The King's Fund.
- Chitnis X, Georghiou T, Steventon A and Bardsley M (2012) *The Impact of the Marie Curie Nursing Service on Place of Death and Hospital Use at the End of Life*. Nuffield Trust.
- Chitnis X, Georghiou T, Steventon A and Bardsley M (2013) 'Effect of a home-based end-of-life nursing service on hospital use at the end of life and place of death: a study using administrative data and matched controls', *BMJ Supportive & Palliative Care* 1–9. doi:10.1136/bmjspcare-2012-000424.
- D'Souza S and Guptha S (2013) 'Preventing admission of older people to hospital', *BMJ* 2013;346:f3186.
- Department of Health (2009) *Partnerships for Older People Projects*. Available at: <http://webarchive.nationalarchives.gov.uk/+/www.dh.gov.uk/en/SocialCare/Deliveringadultsocialcare/Olderpeople/PartnershipsforOlderPeopleProjects/index.htm> . Accessed 14 May 2013.
- Department of Health (2012) *Reference Costs 2011–12*.
- Department of Health (2013) Letter inviting expressions of interest for health and social care integration 'pioneers'. Gateway Reference Number: 00079. Available at: <https://www.gov.uk/government/publications/social-care-integration-pioneers>
- European Observatory on Health Systems and Policies (2012) *Health Systems and the Financial Crisis* vol 18 no 1.
- Forder J (2009) 'Long term care and hospital utilisation by older people: an analysis of substitution rates', *Health Economics* 18:1322–38.
- Goodwin N, Smith J, Davies A, Perry C, Rosen R, Dixon A, Dixon J and Ham C (2012) *Integrated Care for Patients and Populations: Improving outcomes by working together*. A report to the Department of Health and the NHS Future Forum. The King's Fund. Available at: www.kingsfund.org.uk/sites/files/kf/integrated-care-patients-populations-paper-nuffield-trust-kings-fund-january-2012.pdf . Accessed 16 April 2013.

- Henderson C, Knapp M, Fernández JL, Beecham J, Hirani SP, Cartwright M, Rixon L, Beynon M, Rogers A, Bower P, Doll H, Fitzpatrick R, Steventon A, Bardsley M, Hendy J and Newman SP (2013) 'Cost effectiveness of telehealth for patients with long term conditions (Whole Systems Demonstrator telehealth questionnaire study): nested economic evaluation in a pragmatic, cluster randomised controlled trial,' *BMJ* 346:f1035. doi: 10.1136/bmj.f1035.
- Leviton LC, Kettel Khan L, Rog D, Dawkins N and Cotton D (2010) 'Evaluability assessment to improve public health policies, programs, and practices,' *Annu. Rev. Public Health* 31:213–33.
- Lewis G, Bardsley M, Vaithianathan R, Georghiou T, Billings J and Dixon J (2011a) 'Do 'virtual wards' reduce rates of unplanned hospital admissions, and at what cost? A research protocol using propensity matched controls,' *International Journal of Integrated Care* Vol 11, URN:NBN:NL:UI:10-1-101516 / ijic2011-79.
- Lewis G, Curry N and Bardsley M (2011b) *Choosing a Predictive Risk Model: A guide for commissioners in England*. Nuffield Trust.
- Lewis G, Georghiou T, Steventon A, Vaithianathan R, Chitnis X, Billings J, Blunt I, Wright L, Roberts A and Bardsley M (2012) *Analysis of Virtual Wards: A multidisciplinary form of case management that integrates social and health care*. NIHR Service Delivery and Organisation programme. In press.
- Ling T, Bardsley M, Adams J, Lewis R and Roland M (2010) 'Evaluation of UK Integrated Care Pilots: research protocol,' *International Journal of Integrated Care* [serial online] Available at: www.ijic.org/. URN:NBN:NL:UI:10-1-100969.
- McNulty T and Ferlie E (2002) *Reengineering Health Care: The complexities of organisational transformation*. Oxford University Press.
- NHS England (2013) *Personal Health Budgets* www.personalhealthbudgets.england.nhs.uk
- Nuffield Trust and Imperial College London (2013) *Evaluation of the First Year of the Inner North West London Integrated Care Pilot*. Nuffield Trust in partnership with Imperial College London.
- Pawson R and Tilley N (2004) *Realistic Evaluation*. Sage.
- Powell AE, Rushmer RK and Davies HT (2009) *A Systematic Narrative Review of Quality Improvement Models in Health Care*. NHS Quality Improvement Scotland.
- Purdy S (2010) *Avoiding Hospital Admissions: What does the research evidence say?* The King's Fund.
- Purdy S, Paranjothy S, Huntley A, Thomas R, Mann M, Huws D, Brindle P and Elwyn G (2012) *Interventions to Reduce Unplanned Hospital Admission: A series of systematic reviews*. Bristol University Final Report.
- RAND Europe and Ernst & Young LLP (2012) *National Evaluation of the Department of Health's Integrated Care Pilots*. Available at: www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_133126.pdf
- Roland M, Dusheiko M, Gravelle H and Parker S (2005) 'Natural history of emergency admission in older people: analysis of routine admission data,' *British Medical Journal* 330:289–92.
- Roland M, Lewis R, Steventon A, Adams J, Bardsley M, Brereton L, Chitnis X, Staetsky L, Tunkel S and Ling T (2012) 'Case management for at-risk elderly patients in the English Integrated Care Pilots: observational study of staff and patient experience and secondary care utilisation,' *International Journal of Integrated Care* vol 12, URN:NBN:NL:UI:10-1-113731 / ijic2012-130 .
- Rossi PH, Freeman HE and Lipsey MW (2004) *Evaluation: A systematic approach*, 7th ed. California: Sage.
- Rubin DB (2010) 'On the limitations of comparative effectiveness research,' *Statistics in medicine* 29(19):1991–5; discussion 1996–97. Available at: www.ncbi.nlm.nih.gov/pubmed/20683890 . Accessed 14 March 2013.
- Shaw S and Levenson R (2011) *Towards Integrated Care in Trafford*. Research report. Nuffield Trust.
- Shaw S, Rosen R and Rumbold B (2011) *What is Integrated Care? An overview of integrated care in the NHS*. Nuffield Trust.
- Smith J, Porter A, Shaw S, Rosen R, Blunt I and Mays N (2013) *Commissioning High-quality Care for People with Long-term Conditions*. Nuffield Trust.

- Steventon A, Bardsley M, Billings J, Georghiou T and Lewis G (2012a) 'The role of matched controls in building an evidence base for hospital-avoidance schemes: a retrospective evaluation', *Health Serv Res* 47(4):1679-98. doi: 10.1111/j.1475-6773.2011.01367.x.
- Steventon A, Bardsley M, Billings J, Dixon J, Doll H, Hirani S, Cartwright M, Rixon L, Knapp M, Henderson C, Rogers A, Fitzpatrick R, Hendy J and Newman S (2012b) 'Effect of telehealth on use of secondary care and mortality: findings from the Whole System Demonstrator cluster randomised trial', *BMJ* 344:e3874.
- Steventon A, Bardsley M, Billings J, Dixon J, Doll H, Beynon M and others (2013) 'Effect of telecare on use of health and social care services: findings from the Whole Systems Demonstrator cluster randomised trial', *Age and Ageing* 0:1-8.
- Wagner EH, Austin BT and Von Korff M (1996) 'Organizing care for patients with chronic illness', *Millbank Quarterly* 74, no. 4: 511-44.
- Windle K, Wagland R, Forder J, D'Amico F, Janssen D and Wistow G (2009) *National Evaluation of Partnerships for Older People Projects*. The Department of Health.

Our work on evaluation

Recent developments in data linkage mean that it is now possible to provide evidence on the impact of innovations on secondary care utilisation in a timely and less resource-intensive manner. The Nuffield Trust has developed evaluation methodologies that exploit the large amount of administrative information on individual patients that is available in the NHS and social care. We have used these methods to evaluate a range of community-based service innovations in the NHS over the past four years. Our focus has usually been on whether the service model being implemented and evaluated has had an impact on service use and costs, using quantitative methods and summative evaluation.

To find out more, and to access all our publications and other resources on this topic, please visit the evaluation section on our website at:

www.nuffieldtrust.org.uk/our-work/evaluation

More detailed information on our methods is available from our project page:

www.nuffieldtrust.org.uk/our-work/projects/developing-evaluation-methods-help-nhs-provide-care-more-effectively

For more information about the Nuffield Trust,
including details of our latest research and analysis,
please visit www.nuffieldtrust.org.uk

 Download further copies of this summary from
www.nuffieldtrust.org.uk/publications

 Subscribe to our newsletter:
www.nuffieldtrust.org.uk/newsletter

 Follow us on Twitter: [Twitter.com/NuffieldTrust](https://twitter.com/NuffieldTrust)

Nuffield Trust is an authoritative
and independent source of
evidence-based research and
policy analysis for improving
health care in the UK

59 New Cavendish Street
London W1G 7LP
Telephone: 020 7631 8450
Facsimile: 020 7631 8451
Email: info@nuffieldtrust.org.uk

 www.nuffieldtrust.org.uk

Published by the Nuffield Trust.
© Nuffield Trust 2013. Not to be reproduced
without permission.