

Future evaluation of the Integrated Personal Commissioning programme

Mapping the logic and assessing evaluability

Megan P.Y. Sim, Catherine L. Saunders, Eleanor Winpenny, Tom Ling

For more information on this publication, visit www.rand.org/t/RR1350

Published by the RAND Corporation, Santa Monica, Calif., and Cambridge, UK

© Copyright 2016 NHS England

RAND® is a registered trademark.

RAND Europe is an independent, not-for-profit policy research organisation that aims to improve policy and decision making in the public interest through research and analysis. RAND's publications do not necessarily reflect the opinions of its research clients and sponsors.

All rights reserved. No part of this book may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the client.

Support RAND

Make a tax-deductible charitable contribution at
www.rand.org/giving/contribute

www.rand.org
www.rand.org/randeeurope

Preface

Announced by Simon Stevens, Chief Executive of NHS England, in October 2014, the Integrated Personal Commissioning (IPC) Programme is a new programme that joins up health and social care funding for individuals with complex needs and gives them greater control over how their combined health and social care budget is used. The programme specifically aims to enable people and communities to take a more active role in their health and care needs. It is a new approach to joining up health, social care and education at the level of each individual for children and adults with complex needs. IPC incorporates a range of personalised approaches: support to build people's knowledge, skills and confidence to manage their own health; the offer of greater choice and control through personal budgets; and changes in the design and commissioning of services to put people in the driving seat of decisions around their own care.

It is an important part of a range of initiatives exploring how best to provide more integrated social and health care, in particular for those with complex and chronic conditions. Nine demonstrator sites have been selected for the first wave of roll-out.

The goals of the IPC programme are: to improve the quality of life of people with complex needs, and their carers; to enable them and their families to achieve important goals through greater involvement in their care, and be able to design support around their needs and circumstances; to prevent crises in people's lives that lead to unplanned hospital and institutional care by keeping them well and supporting self-management; and to improve integration and quality of care, including better user and family experience of care.

Simon Stevens undertook to evaluate the programme in order to share learning and identify whether and how implementing the IPC programme works to achieve its goals.

RAND Europe was commissioned by NHS England to work with the nine demonstrator sites to:

- Develop a logic model setting out the core components of the IPC programme
- Identify existing data sources and additional data requirements for evaluation purposes
- Provide advice to local sites on how they will collect the required data
- Understand the core risks and challenges to a future evaluation of the IPC programme.

This work will support NHS England in defining the plans for a future evaluation of the IPC programme. It also reflects a growing interest at RAND Europe in bringing in evaluation skills and insights earlier in the design and implementation stages of complex interventions.

This report has been peer reviewed in accordance with RAND's quality assurance standards. For more information about RAND Europe or this document, please contact Tom Ling (tling@rand.org).

RAND Europe

Westbrook Centre, Milton Road

Cambridge CB4 1YG

United Kingdom

Tel. +44 1223 353 329

Abstract

The Integrated Personal Commissioning (IPC) Programme is a new programme that joins up health and social care funding for individuals with complex needs and gives them greater control over how their combined health and social care budget is used. The programme specifically aims to enable people and communities to take a more active role in their health and care needs. It is a new approach to joining up health, social care and education at the level of each individual for children and adults with complex needs. IPC incorporates a range of personalised approaches: support to build people's knowledge, skills and confidence to manage their own health; the offer of greater choice and control through personal budgets; and changes in the design and commissioning of services to put people in the driving seat of decisions around their own care. Nine demonstrator sites have been selected for the first wave of roll-out, which began on 1 April 2015.

RAND Europe was invited by NHS England to provide support for the sites, and at the national level, in collaboration with the NHS England team at this early stage of roll-out. RAND Europe carried out a workshop with each site to support the development of the logic model behind their local plans, and to identify suitable metrics to measure local progress against these logic models. With the national team, RAND Europe used this learning from the workshops to help define how to evaluate the IPC programme with a common logic model, and provide advice for evaluation going forward.

It should be recognised that such early evaluation work is formative and is not describing a finalised IPC model. It is likely that as sites develop, further core streams will emerge.

We found that there are differences in how well prepared each site is for an evaluation. All are in a position to draw upon and adapt the generic theory of change model and therefore have a model that can be evaluated. However, each site has recognised that they are unlikely to get it right the first time and that they will want to adapt and improve their activities. Ongoing evaluation could help produce site-level data to support local decision-makers and also support future decision-making nationally. Both of these aims would be supported by common measurements collected across the sites, and by a comparative evaluation.

The IPC demonstrator sites do not exist in isolation; integrated personal commissioning is one of several complementary programmes and new models of care being introduced or piloted across the country. The learning from IPC and the future IPC evaluation should be outward looking and contribute to these programmes as well, and so contribute more broadly than to IPC alone.

Table of Contents

Preface	iii
Abstract.....	v
Table of Contents	vii
Figures	viii
Summary	ix
Acknowledgements.....	xii
Abbreviations	xiii
1. Context and introduction.....	1
1.1. The Integrated Personal Commissioning Programme.....	1
1.2. This study	1
1.3. Organisation of this Report	3
2. Rationale for logic models.....	5
3. Logic model development	7
3.1. A common logic model across IPC sites.....	7
3.2. Variations to the core logic model	9
4. Evaluation design.....	13
4.1. Three core streams of activity within IPC.....	14
4.2. Health and social care population outcomes.....	20
4.3. Further technical comments	21
5. Discussion	29
5.1. Are the demonstrator sites evaluation-ready?.....	29
5.2. Conclusions	29
References	31

Figures

Figure 1 Integrated Commissioning Programme logic model	8
Figure 2 Three core streams of activity leading to desired outcomes	13
Figure 3 Linear process.....	17
Figure 4 The three elements of market facilitation.....	19
Figure 5 Assessing Value for Money	19
Figure 6 Adult Social Care Outcomes Framework Indicators	20
Figure 7 Individual comparisons within those ever offered IPC.....	22

Summary

Introduction

The Integrated Personal Commissioning (IPC) Programme is a new programme that joins up health and social care funding for individuals with complex needs and gives them greater control over how their combined health and social care budget is used. The programme specifically aims to enable people and communities to take a more active role in their health and care needs. It is a new approach to joining up health, social care and education at the level of each individual for children and adults with complex needs. IPC incorporates a range of personalised approaches: support to build people's knowledge, skills and confidence to manage their own health; the offer of greater choice and control through personal budgets; and changes in the design and commissioning of services to put people in the driving seat of decisions around their own care.

It is part of a range of new initiatives from NHS England, and jointly governed with the Local Government Association, which explores how best to provide more integrated social and health care, in particular for those with complex and chronic conditions.

RAND Europe is an independent not-for-profit research institute whose mission is to help improve policy and decision-making through research and analysis. We were asked by NHS England to provide 'research and evaluation support' for the initial phases of the IPC programme. Specifically, we aimed to work with demonstrator sites to elucidate the logic model behind their local plans for IPC; to help them identify suitable metrics to measure local progress against these logic models; to develop a core logic model across sites; and to provide advice on the design of the national evaluation.

Methodological approach

Good practice in both designing complex programmes and evaluating them includes development of a theoretical understanding of the strategy (a Theory of Change) and of how the intervention causes change, so that weak links in the causal chain can be identified and strengthened.

By working with each demonstrator site to make a logic model explicit, we aimed to provide an opportunity for those working within each site to decide for themselves how best to measure their success. In turn, this would provide each site with a stream of evidence to support its learning and adaptation.

We reviewed the IPC documentation from each of the demonstrator sites and developed a draft logic model which we took to workshops at each site where we refined, reviewed and explored the local sites' logic models. From this work we developed a common logic model.

Findings from this research

We found that within IPC there were three core streams of activity leading to desired outcomes. These three streams are around empowered individuals, sustainable commissioning and a vibrant market of providers. The main impact agreed across sites was the improved quality of life for patients and their carers. Reduced institutional and acute care, and more efficient use of money and resources, were also identified across sites as overall impacts for IPC.

To measure individual empowerment activities, outputs and outcomes, we recommend individual process measures, measures of individual empowerment and measures of individual quality of life. For sustainable commissioning, we recommend health and social care process measures, financial model and cost data, and measures of workforce culture and skills. For the vibrant market of providers, we recommend measures of market development and value for money.

A successful national evaluation of the IPC programme will need to capture elements across all three core strands of the programme, and elements across all levels (activities, outputs, outcomes and impacts), to be able to understand the causal process towards changes in impacts of the programme. The evaluation will also need to be able to compare IPC results to a counterfactual to answer the question: what impact has IPC had compared to what would have happened if IPC had not been implemented?

Conclusions

Integrated personal commissioning is one of several programmes and new models of care being introduced or piloted across the country and builds on previous programmes, including personal health budgets and the Better Care Fund. The learning within and across sites from early IPC work can contribute to these programmes as well, and so more broadly than to the IPC programme alone.

The implementation, evaluation and learning from IPC are iterative processes – sites are not expected to set up IPC in isolation. We recommend that this iterative process of local learning, maintaining communication between sites and with the NHS England programme team, form part of the evaluation process.

There are differences in how well prepared each site is for an evaluation. All are in a position to draw upon and adapt the generic theory of change model and therefore have a model that can be evaluated. However, each site has recognised that they are unlikely to get it right the first time and that they will want to adapt and improve their activities. To this extent, an ongoing evaluation could help produce site-level data to support local decision-makers. In addition, a scheme-wide evaluation could draw together the lessons from the first cohort to support future decision-making nationally. Both of these aims would be supported by common measurements collected across the sites and a comparative evaluation.

We also recommend that the methodological approaches and issues identified and described in this report feed into the design of a future evaluation of the demonstrator sites as well as a national evaluation. Although there are substantial methodological challenges to a national quantitative impact evaluation, acknowledging and addressing these from the outset will strengthen the learning from future work.

We hope the information above provides NHS England with a foundation on which they and the sites can build in order to continue to define IPC and develop standard models for delivery, which in turn will inform the implementation of IPC over the coming years.

Acknowledgements

The authors would like to thank the team at NHS England for support throughout this work, and workshop participants at the nine demonstrator sites. We would also like to thank our quality assurance reviewers at RAND Europe, Jon Sussex and Teresa Bienkowska-Gibbs. The views presented in this report are the authors'. Any remaining errors are our own.

Abbreviations

CCG	Clinical Commissioning Group
EQ-5D	Euro-QoL 5-Dimensions
HSCIC	Health and Social Care Information Centre
IPC	Integrated Personal Commissioning
NPT	Normalisation Process Theory
PAM	Patient Activation Measure
POET	Patient Outcome and Experience Tool

1. Context and introduction

1.1. The Integrated Personal Commissioning Programme

The Integrated Personal Commissioning (IPC) Programme targets four groups of high-need individuals (older people with long-term conditions, children with disabilities and their families, people with learning disabilities, and people living with serious mental illness) with the aim of giving them greater control over how their own combined health and social care budget is spent. Nine demonstrator sites have been selected for the first wave of roll-out, which began on 1 April 2015. These are:

- Barnsley
- Cheshire West and Chester
- Hampshire
- Lincolnshire
- Luton
- Portsmouth
- South West Consortium
- Stockton-on-Tees
- Tower Hamlets.

1.2. This study

RAND Europe was invited to provide support for the sites and at the national level, in collaboration with the team from NHS England. Specific responsibilities were as follows.

With the sites:

- Support the demonstrator sites in articulating the logic model behind their local plans for IPC.
- Help sites to identify suitable metrics to measure local progress against these logic models.
- Provide advice for setting up the baseline data collections.

With the national team:

- Use the work with the local demonstrator sites to help NHS England define how to evaluate the IPC programme and define the IPC models being used.
- Understand common elements of IPC across the nine different demonstrator sites.

- Think about how the existing online portal for the IPC sites can best be used locally and nationally.
- Identify early areas of good practice that may be helpful in the national roll-out of the programme, and identify whether they could be captured in any future evaluation.
- Understand how local demonstrator sites are developing and using local data collection mechanisms, and assess how these can feed into a single data collection process for the evaluation of the national programme.
- Understand how the evaluation of IPC can work in parallel with, or learn from, the evaluations of other initiatives, for example personal health budgets, personal social care budgets, and integration pioneer sites.¹ Understand what will need measuring that is unique (or not) to this programme. Acknowledge the need not to over-burden sites with duplicative evaluations.
- Ensure that service user voices are heard at this stage, and are central to any future strategy.

To these ends, RAND Europe undertook to deliver:

- A clear logic model setting out the core components of the IPC programme (as seen in the IPC sites) and identify the core hypotheses that will be tested through evaluation after the national roll-out.
- Identification of relevant existing data sources as well as additional data requirements to allow baseline data collection from September 2015 onwards; including an understanding of local difficulties in obtaining the data and how to overcome these difficulties.
- Advice to local sites to develop plans on how they will collect the required data.
- An understanding of the core risks and challenges (e.g. in terms of methodology, data, sample sizes) to the evaluation of the national programme.
- A final workshop reporting findings to the sites and inviting further feedback and comment.

¹There are several related programmes and initiatives with cross-over or synergy with IPC, either at the national level, or at the local level in the demonstrator sites themselves. They were mentioned in workshops across most sites. A brief introduction is provided here for the reader for those that are mentioned in this report; this list is not meant to be exhaustive:

- The personal health budget programme has recently run a pilot programme, and has now been rolled out nationally. They work in a similar way to personal budgets in social care and give individuals with long-term conditions and disabilities greater choice and control over the healthcare and support they receive.
- The Better Care Fund (formerly the Integration Transformation Fund) creates a local single pooled budget to incentivise the NHS and local government to work more closely together, integrating health and social care with well-being as the focus of health and care services.
- The aim of the Integrated Care Pioneers is to make health and social care services work together to provide better support at home and earlier treatment in the community to prevent people needing emergency care in hospital or care homes.
- Year of Care aims to provide personalised care planning for people with long term conditions by working in partnership with patients and care professionals.
- PAM pilots are working with NHS England to help clinicians assess the care and support needed by people with long term conditions using a new assessment tool, known as the patient activation measure (PAM).
- Vanguard sites are early adopters for the new care models programme introduced in the [Five Year Forward View](#) to support improvement and integration of services. To date these vanguard programmes include integrated primary and acute care systems; enhanced health in care homes; multispecialty community provider, urgent and emergency care, acute care collaborations.

1.3. Organisation of this Report

In this report we discuss the rationale for logic models in Chapter 2, discuss the common logic model development in Chapter 3, and evaluation design in Chapter 4. We end in Chapter 5 with some concluding remarks.

2. Rationale for logic models

The IPC programme is an example of a complex intervention, one which contains many interacting components, and in this case may include personalised care and support planning, independent advocacy, peer support and brokerage, service change, and integrated personal budgets (Moore et al. 2015, Ling 2012). The different components may interact such that outputs are non-linear and difficult to predict. In addition, implementation of the programme differs across intervention sites, both as a result of planned differences (typically building on local strengths and experiences), but also because of unexpected responses from carers and service users or from individuals and groups implementing the changes. Although the intervention is complex, it is not chaotic and each demonstrator site not only builds on existing relationships and practices but also has a plan to change these to achieve more or less clearly defined outcomes and impacts.

In such circumstances, we anticipated that each demonstrator site would have at least an implicit, tacit logic model describing the intended causal pathways connecting how changes to, for example, care planning, commissioning or incentives to providers might lead to more empowered service users and families, more responsive provision, and more efficient care systems. By working with each demonstrator site to make such a logic model explicit, we aimed to provide an opportunity for those working within each site to decide for themselves how best to measure their success. In turn, this would provide each site with a stream of evidence to support its learning and adaptation. Good practice in both designing complex programmes and evaluating them includes development of a theoretical understanding of the strategy (a Theory of Change) and of how the intervention causes change, so that weak links in the causal chain can be identified and strengthened. The benefits include creating greater clarity among contributing agencies and providing a tool to demonstrate to funders and other stakeholders the strategy and measurable value of the demonstrator programme.

The IPC programme aimed to create opportunities for improved measurement and learning for each site and also to strengthen the evaluability of each site. Evaluability is the extent to which a programme or entity can demonstrate in a reliable and credible fashion what is being achieved and how (Davies and Payne 2015). In the context of a pilot programme it is important that each pilot should be 'evaluation ready' to maximise the learning across the sites and to inform future policy in this area. Therefore, we aimed to understand the extent to which there was a shared, generic logic model, as well as site-specific models. To this end, we took the pragmatic step of proposing the same basic logic model to each site and allowing them to adapt the model. The initial template was based on the various plans put forward by each site. Methodologically this suffers from a danger that we 'led' the sites, but practically it allowed us to

identify shared things to measure, and preferred measurement tools, that were common to the whole programme as well as those that were site-specific. The time spent with each site persuaded us that the sites were unlikely to be unduly led towards a version of the logic model that they could not own.

In this way, and in line with best evaluation practice, we aimed to use the logic modelling activity to:

1. Strengthen the demonstrator sites' theoretical understanding of how the intervention will cause the intended change, so that weak links in the causal chain can be identified and strengthened.
2. Undertake preparatory work to identify an appropriate evaluation design to assess effectiveness (and this will be taken forward as part of the national evaluation).
3. Support the careful choice of appropriate process and outcome measures.

3. Logic model development

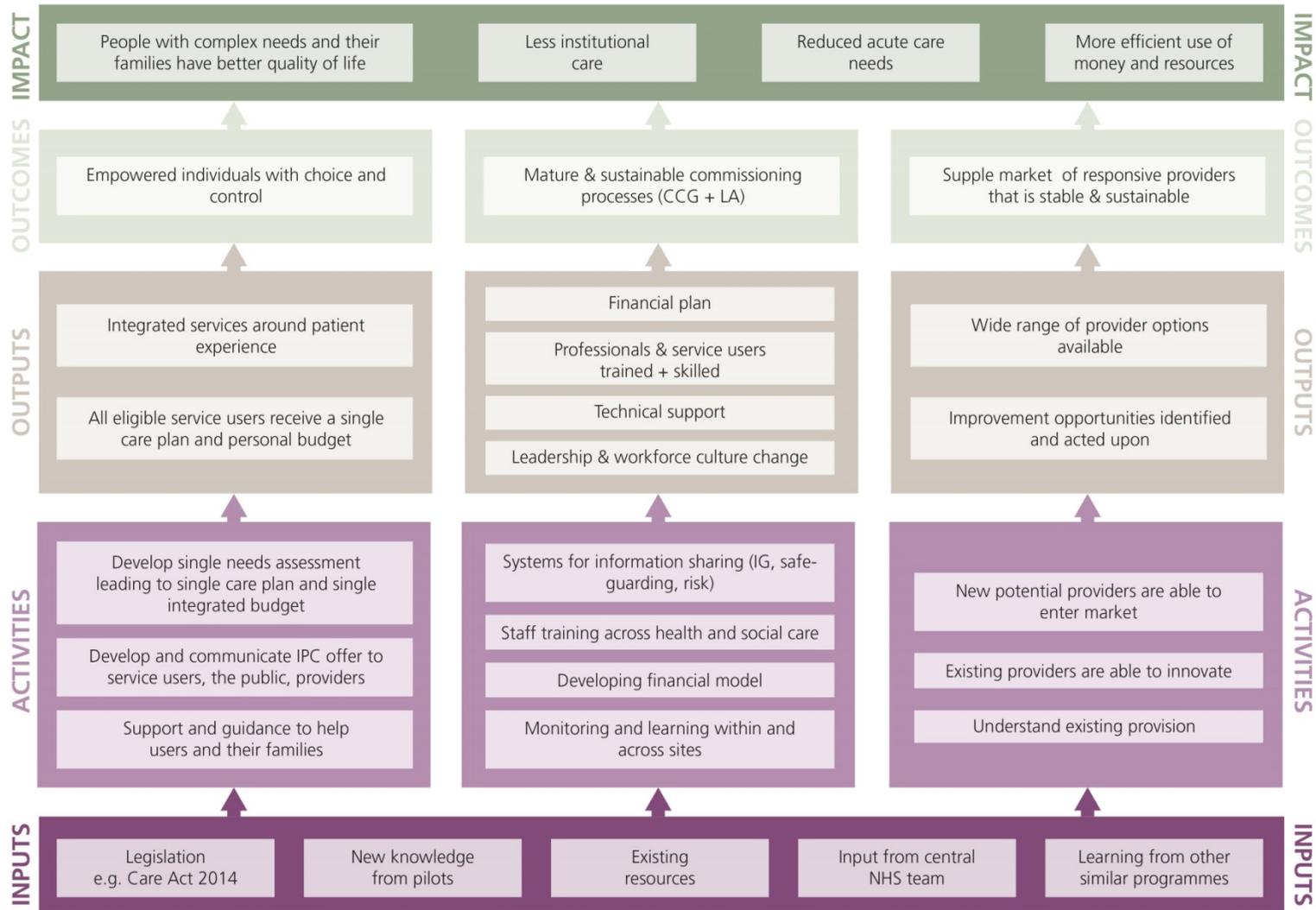
3.1. A common logic model across IPC sites

As discussed in Chapter 2, logic models help us to describe and understand an intervention programme, setting out the core activities that form part of the programme, the expectations of outcomes for the programme, and the causal assumptions which underpin these expectations. In order to describe and understand the IPC programme, as a first step towards evaluation, we worked with each of the IPC sites to develop a logic model for each individual site, with the aim of synthesising these logic models into an overall programme logic model which can then underlie a future programme evaluation.

We first developed a draft logic model, based on a review of the proposals which each IPC site submitted to NHS England as part of their application to be an IPC demonstrator site, plus other documents shared with us by NHS England. The RAND Europe team then visited each of the sites to conduct a workshop with key members of each project team. These workshops aimed to revise and develop the logic model in line with the specific plans in place in each site. We incorporated detail on the activities which were planned in each site, inputs into these activities, outputs that were expected from these activities, outcomes and intended impacts. We also discussed in detail the causal pathways that would be expected to flow through the logic model, leading from inputs and activities through to impacts.

Having gained an in-depth understanding of the site-specific logic models, we then drew these logic models together to form an overall logic model for the programme. This specifies the core elements of the programme, although there are a number of details that will vary for the different sites. Below we set out first the core logic model, and then describe any important variations that were discussed by sites.

Figure 1 Integrated Commissioning Programme logic model



There was considerable agreement as to the main items of the logic model, demonstrating a good understanding and consensus among IPC sites as to the key aims of the IPC programme. All sites agreed with the common aims and perceived impact of the IPC programme, and discussed three main streams of activity that would lead to these impacts. These three main streams were:

1. Activities around the service user and their families, including a single care plan and personal budget, leading to empowered individuals with greater choice and control and improved individual and carer health/quality of life outcomes.
2. Changes in commissioning activities and the health and social care workforce which will lead to mature and sustainable commissioning processes across health and social care, as well as trained employees able to work across boundaries and support and assist individuals to take control of their own care.
3. Development of the provider market, to allow greater flexibility so that service provision can be driven by patient needs and choice.

These three broad themes are depicted as different sections within the logic model shown in Figure 1.

3.2. Variations to the core logic model

We found some variations between site-specific logic models, in part due to the variation in context between sites and the differences between the target populations for IPC in each site. Each of the areas where there were some differences in emphasis between sites is discussed further below.

3.2.1. *The IPC offer to individuals*

All sites were clear that IPC needed to offer each individual a single, personalised care plan. In some sites, this aim was extended by the suggestion that all individuals would require a single joint needs assessment across all their health and care needs which would lead not only to the care plan, but also to an indicative budget offer. The extent to which service users within the IPC programme would expect to receive a personal budget or direct payment differed between sites and was not always clear.

In some sites, participants considered that it may be necessary to define clearly what was included in the IPC offer and on what the personal budget could be spent. Participants had concerns that a personal budget might be spent on things that were not considered acceptable. This process might involve development of parameters that set out the boundaries for what individual budgets could be used for, or the development of an IPC policy.

Participants from some sites spoke of the need for additional services for those on the edge of eligibility for IPC. For example, phone-in or drop-in services might be provided for those needing additional support but who were not eligible for the full IPC offer.

3.2.2. *Offer of support for those taking up IPC*

In order to allow individuals to become empowered to participate fully in the management of their own care and their budget, several sites had plans for additional support for service users and their families.

Increasing the uptake of direct payments could be a first step, and support for service users to build their confidence and experience might assist with this. Brokerage support services could also assist service users and their families in making choices for their care needs. A step methodology whereby individuals would receive different levels of support dependent upon their needs might also work. Participants also suggested that it may be useful to outsource care and support planning to voluntary organisations who can offer peer support. Participants suggested additional support as one method to ensure equality of access to services, as support services might increase those who might otherwise not take up IPC.

3.2.3. Communication, education and culture change

Participants understood the communication of IPC to be a key part of the IPC programme. Workshop participants discussed the need to communicate IPC to different groups, including service users and their families, in order to get uptake and engagement, to communicate to providers to allow market development, and also to communicate the aims of the programme to the general public. However, sites had different emphases and different suggestions for engaging the population. Some sites were building on existing programmes, working with third sector providers, or holding specific events to engage people in the programme. Many sites discussed a need for service user empowerment as well as broader culture change among commissioners and providers, to allow staff to support service users, and in some cases they were considering specific training for staff.

3.2.4. Other ongoing integration or personal budget programmes

A number of sites mentioned ongoing programmes which overlap to some extent with the IPC programme, including 'Year of Care', 'PAM pilots', 'Integrated care pioneers', and being 'Vanguard' sites for new models of care. Potential positive benefits could be seen from introducing learning from these programmes into IPC, however in some cases there were thought to be conflicts arising from a number of programmes trying to do similar things but in different ways, and differences in timing.

3.2.5. Shaping the market

Some sites had made more progress than others when thinking about how to work with providers to ensure a responsive and flexible market. Many sites cited the need to move from block contracts to tariff-based contracts with providers, but also the need to avoid destabilising the market too much. Several sites mentioned aims to develop new commissioning models or for providers' service models to change. In some cases there was thought to be a need to develop the personal assistant market, but in others this market was already well developed. Several sites mentioned the need to get a better baseline understanding of the current provider landscape, as well as an understanding of the needs of service users and their families across health and social care. Workshop participants cited work by Age Concern that included interviews of service users to understand from their point of view what is working well or less well, which could be used to improve the choice of services available. Sites could also collect input from early IPC service users to inform development of the market.

3.2.6. Outcomes

There was a high level of agreement on the outcomes. The main outcome mentioned was improved quality of life for service users and their carers. Participants emphasised that this was overall quality of life, and could include education and employment opportunities, not only health- and care-related quality of life. Additional local aims included to have people living in the community longer, with fewer people needing residential care. In some sites there was also an aim to ensure that the services required could be provided within the local area. Improved life expectancy and reduced health inequalities were possible outcomes. Another aim mentioned was to enhance prevention and decrease the number of people requiring complex care.

An additional outcome participants discussed in many sites was the possibility of better value for money or efficiency of resource use. There was some disagreement as to whether this was an aim of the programme, or an expected secondary outcome from better care processes. Some sites suggested that money would be saved through back office efficiency savings, while others proposed that better preventative care and fewer acute admissions would lead to cost reductions.

3.2.7. Equality

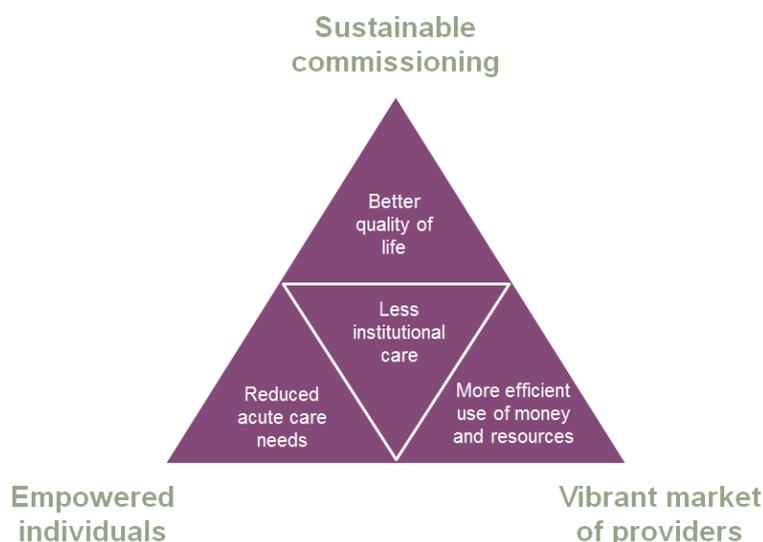
We included issues of equality of provision in our draft logic model, including ‘rights of all service users are equally protected’ and ‘ensuring equality of eligible cohort’. These issues however were not discussed extensively in the site workshops, and have thus been removed from the final logic model. In some site workshops, participants discussed the difficulty for some service users to actively participate in their own care planning and budgeting, and suggested ways to support service users, as described above. In some sites, there was also an aim to encourage a wide range of people to access the IPC programme, for example among various ethnic groups. Across most sites there was less awareness of the impact that the IPC programme might have on those not eligible to participate in the programme, and no discussion of the potential risk of the emphasis on IPC detracting from resources available for other services.

4. Evaluation design

We propose that the evaluation design should focus on the core elements of the programme as captured in the logic model. The common logic model developed through the workshops includes the following common or key elements, as illustrated in Figure 2:

1. Activities around the service user and their families, including a single care plan and personal budget, leading to empowered individuals with greater choice and control and improved individual and carer health/quality of life outcomes.
2. Changes in commissioning activities and the health and social care workforce which will lead to mature and sustainable commissioning processes across health and social care, as well as trained employees able to work across boundaries and support and assist individuals to take control of their own care.
3. Development of the provider market, to allow greater flexibility so that service provision can be driven by patient needs and choice.

Figure 2 Three core streams of activity leading to desired outcomes



Measures are needed across these three streams of work, including measures to capture activities, outputs, outcomes and impacts. While we are interested in the outcomes and impacts of the programme, it is important to capture activities and outputs that lead to these outcomes and impacts to understand the elements of the programme that have made the programme successful or unsuccessful.

Against this aim of capturing key data across the logic model, there is a need for data collection to be feasible and not overly burdensome for sites (e.g. neither too many surveys, nor individual surveys which are too long or time consuming to collect). In previous work, the data collection for the personal health budget evaluation included several survey tools designed to understand individual outcomes among budget recipients and controls including: Health-related quality of life (Euro-QoL 5-Dimensions or EQ-5D); Care-related quality of life (Adult Social Care Outcomes Toolkit or ASCOT); Psychological well-being (General Health Questionnaire-12); Subjective well-being; perceived quality of life (a seven-point scale); and Perceived health (a five-point scale) (Jones et al. 2013). The data collection burden from the use of so many tools was anecdotally reported during the site workshops to be high. The need to embed data collection that has value for local learning as well as for a national evaluation, and that does not just end at the end of the evaluation process, was also highlighted. The Patient Outcome and Experience tool (POET, discussed below) has also been used in this cohort (Hatton & Waters 2013).

4.1. Three core streams of activity within IPC

Below we discuss measures that could capture different elements of IPC progress within each of the three streams.

4.1.1. *Activities around the service user and their families, and individual level process and outcome measures*

Individual process measures

It will be important to collect data on the number of individuals who are taking part in the IPC programme. Data to collect should include:

- Number of eligible services users identified
- Number of needs assessments conducted
- Number of individuals with a single care plan
- Number of individuals with a personal budget.

Care plan goals

For an individual, meeting their own care plan goals was identified as a key outcome. These (as well as the PAM and POET tools) could be collected in coordination with care plan reviews.

Workshop participants identified understanding when goals were not met as a key step for understanding unmet needs in the IPC cohort.

The tension between statutory requirements for care plan reviews and the need for care plan reviews to be individually rather than institutionally determined nonetheless leaves the care plan review as a key moment where data may be captured for the IPC evaluation.

Measures of individual empowerment

The POET and PAM tools have been identified as key metrics of interest for NHS England, with POET used in previous evaluations (Hatton & Waters 2013), and a recent review of PAM identifying some early

successes (Armstrong et al. 2015). These also came up repeatedly in our site workshops as metrics for measurement of patient empowerment and patient outcomes.

PAM

NHS England has recently been involved with work testing the Patient Activation Measure (PAM).

PAM measures patient activation, a measure of someone's skills and confidence to manage one's own health, and is a well validated, short, tool with a strong evidence base (Hibbard et al. 2005, Hibbard et al. 2004). Better activation is identified as being associated with better patient outcomes, even after accounting for obvious confounding factors such as income.

Initial NHS England work on the implementation of PAM found that it is being used both at a person level, i.e. during consultations to allow health professionals to identify how activated a patient is and to tailor their care appropriately; and also at a population level for needs assessment planning. Anecdotally, however, the CCG that took this approach found it difficult to apply the findings in practice, and further analysis is ongoing by the Health Foundation (Armstrong et al. 2015).

A second key limitation is that PAM has not been evaluated in cohorts of people with learning disabilities or dementia, where proxy completion may be more important.

Doctors are reported anecdotally by sites and in conversations with staff from NHS England to like PAM because there is strong research evidence; however, it is also reported that it is less popular in social care as it is quite a 'clinical' tool.

There is also a version appropriate for carers, and one for clinicians (although the version for doctors is developed as a self-reflection tool, and is not appropriate as an evaluation tool or performance metric).

POET

The Personal Outcomes Evaluation Tool (POET), developed by InControl and Lancaster University, is designed to assess the quality of the process of obtaining a care plan or personal budget, and the outcomes of having a plan.

Two key limitations for using the POET tool were identified during the workshops:

1. The lack of a baseline measure.
2. The fact that if sites do not send sufficient returns to InControl then they will only contribute to national learning, and site-level findings will not be fed back locally.

Given that POET assesses change over time, we would not consider the lack of a baseline measure as a major concern for evaluation, since absolute values will remain important for this metric.

Measures of individual quality of life

In addition to the more person-centred outcome measures such as the POET, we would suggest that sites are asked to collect data using at least one overall quality of life assessment tool. Measurement of quality of life is difficult, and as far as we are aware, there is not a single well-accepted tool that covers all aspects of a service user's quality of life, across health, care, employment and many other aspects of life which may be important to the individual. We therefore suggest that EQ-5D may be an appropriate tool to use as it is well validated and well trusted, together with a second tool that assesses quality of life over a wider range

of dimensions (Dronavalli & Thompson 2015). Tools worth investigating further for this purpose include the Quality of Life Scale (Burckhardt & Anderson 2003) and the Personal Wellbeing Index (International Wellbeing Group 2013). Measures of health-related quality of life such as EQ-5D may also give important insight into the outcomes of carers (Thomas et al 2015).

Specific outcome measures

There are specific tools to capture quality of life in different groups, for example the Family Experiences with Coordination of Care tool, identified to measure the experiences of children with chronic health problems (COE4CCN 2015), or the tools identified in the Kent Integration Pioneers Evaluation Framework (Billings & de Weger n.d.), in the Evaluation Frameworks already in development by some IPC sites, or in previous research (Reeves et al. 2014, Burt et al. 2012). However, it is difficult to recommend a single quantitative measure that will appropriately capture quality of life or individual outcomes across IPC cohorts, and while such specific tools may be important for local evaluation, they are unlikely to be relevant for a national programme. Specific tools will also work differently in different cohorts – although the psychometric properties may vary – for example between different ethnic groups (Grant & Bowling 2011).

4.1.2. Outcomes from changes in commissioning activities and the changes in the health and social care workforce

Health and social care process measures

Key process measures for health and social care will be the development of key systems to facilitate the IPC programme. These include:

- Systems for information sharing
- Systems for pooling budgets for the individual
- Policies in place for IPC
- Information governance procedures.

Financial model and cost data

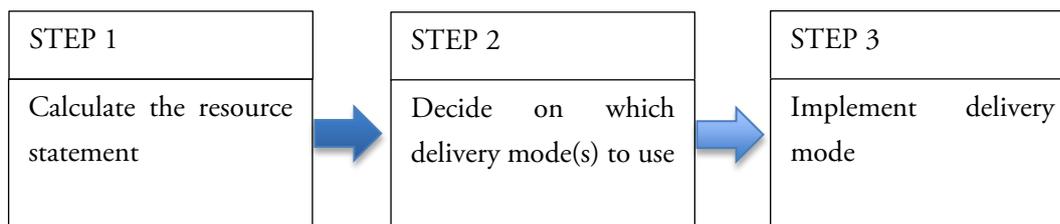
A number of common building blocks that different sites might be expected to draw upon in developing a financial model to support the care model were identified. It was rightly stressed that the ‘financial model’ should be thought of as an ‘enabling model for a model of care’ rather than a separate entity:

- A **resource statement** setting out the total annual care cost for an individual service user (across health and social care); AND
- **Care planning**, support and advocacy: the service user will be involved in determining what care is purchased for them and they will receive appropriate support to do so; AND
- **An integrated notional personal budget** covering the user’s health and social care expenditure, but managed by the local authority and CCG; OR
- **A direct payment** covering (part of) the user’s expenditure, managed by the user themselves; OR

- An integrated personal budget managed by a **third-party coordinator** whose role it is to advise the service user, signpost and broker services – the coordinator is given outcomes-based incentives; OR
- An integrated personal budget managed by a **third-party provider** who is doing the above, but also responsible for the provision of services. They typically will hold a capitated budget covering all of or part of the services needed by the user.

None of the demonstrator sites have fully developed their financial models. Some sites are well developed towards capitation, and others are still working towards being able to do shadow budgeting, ideally by end of financial year 2015, conduct shadow budgeting during 2016, before roll out in FY2017. However, this is not compulsory. A linear process for the financial model is described in Figure 3 below.

Figure 3 Linear process



However, from meeting with the sites it is improbable that these steps will be successfully concluded at the first attempt. Therefore, one approach would be to not get fixated on perfection, to acknowledge there will inevitably be mistakes, and to concentrate on iterations and improvements. In addition, when deciding on which delivery modes to use (and how to adapt them), use the knowledge of services of both participating and non-participating eligible individuals and carers.

As financial models mature, it will be important that credible estimates of costs are available. We suggest there are at least three viable routes to this. The first is to generate a model of standard costs for the 'typical' service user with the attributes of those opting in, and compare this with the actual costs of the typical bundle. A second is to compare the actual costs of those who were eligible but did not participate with those who did. A 'difference-in-difference' element in the second approach would be useful as it would allow for underlying changes in practice, and also perhaps reveal whether resources were being switched away from the non-intervention group rather than just staying in steady state (which is the default assumption). A third approach would be 'before-and-after' cost comparison for people newly enrolled in an IPC scheme: comparing costs over 12 months before and 12 months after enrolment. It is of course understood that people's health generally declines over time and so costs associated with any individual may generally rise, which is a problem for a before-and-after design. But if the comparison is kept to no more than 12 months 'after', it might provide useful evidence of cost changes. Finally, for most sites, numbers of hospital attendances/admissions would be the easiest proxy for costs (but not including those focused on learning disability).

Whichever approach is adopted to meet local circumstances, certain actions would be needed to support strengthening the financial model and understanding what is happening to costs:

1. Create a linked dataset across the cohort.
2. Understand what users have received historically and today.
3. Support care planners' decisionmaking.
4. Ensure that resources were allocated based on real and most recent needs rather than reverting to the mean.
5. Allow commissioners to look at the needs of whole cohorts.
6. Allow commissioners to achieve synergies and savings across settings (e.g. patients with diabetes also receiving counselling for depression).
7. It should enable both market development and personalisation.

Workforce culture and skills

The personal health budget evaluation found that the management of culture change, particularly acknowledging concerns held by frontline staff, was identified as an important element of implementation (Forder et al. 2012). During workshops, many sites identified that culture change and new ways of working may be difficult for staff.

As mentioned before, the PAM tool for staff may be a helpful metric here for understanding the current perspectives of key staff stakeholders, but as a metric it is only appropriate for individual and local learning, rather than for national evaluation or assessment. It may be necessary to develop a bespoke measure of staff skills and attitudes to track any changes over time.

Frequent opportunities for feedback at all levels (from domestic carers to board members) may support local learning. Most sites reported annual staff surveys, and IPC evaluation questions could be embedded within those surveys.

4.1.3. Development of the market

In evaluating the development of the market, a key step will be understanding current service provision, and the extent to which this provision addresses the needs and wishes of the cohort. While the need for understanding was discussed in a number of site visits, there was variation in the approaches taken by sites to answering this question.

It is clear that this is a very important step for IPC across all sites, but the evaluation metric should probably be qualitative, and at the site level – this would encourage the sharing of ideas and good practice across sites, and also allow sites to evaluate their progress by addressing this issue in the early stages of implementation of IPC.

We propose development of a market facilitation role that can be applied across all sites, which will map the services available and the way that such services are contracted. Such a mapping should include voluntary sector organisations. Briefly, the best test of the development of a market is whether the recipients of care feel they have sufficient, or indeed any, choice of good quality providers open to them. Asking the care recipients is the best way to judge that. The Developing Care Markets for Quality and Choice Briefing Paper (IPC Market Analysis Centre 2014) on market facilitation identifies three elements of market facilitation, which are outlined in Figure 4.

Figure 4 The three elements of market facilitation

Market intelligence – the development of a common and shared perspective of supply and demand, leading to an evidenced, published market position statement for a given market.

Market structuring – making explicit to the sector how the commissioner intends to behave in influencing the market. For example, this might include communications with providers and service users and their carers, ongoing planning, quality assurance, or performance management arrangements designed to encourage desired services and discourage those that are not needed.

Market intervention – the interventions commissioners make in order to deliver the kind of market believed to be necessary for any given community. For example, this might include financial incentives, offering specialist training, support to providers with business planning, setting up not-for-profit ventures, grants, or other forms of support for providers to encourage the development of particular services.

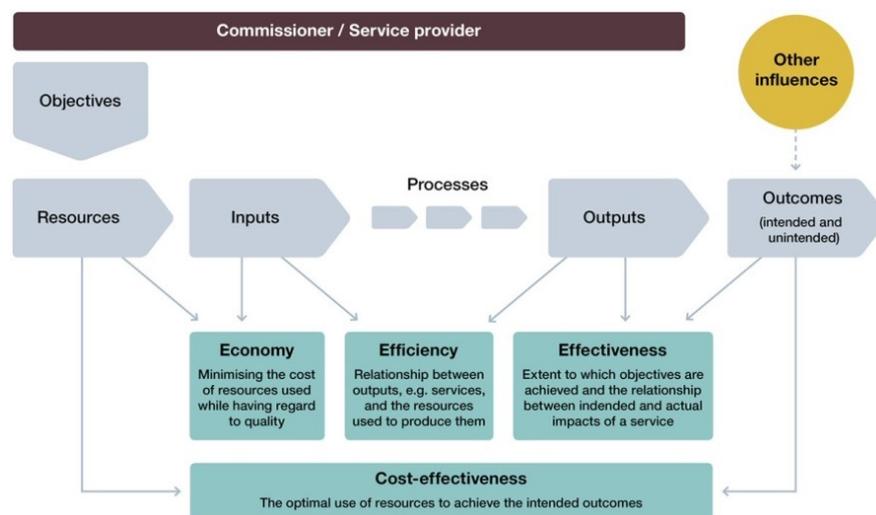
A comparison between market availability and whether individuals’ needs within the care plan are being met would provide an indication of whether the services available are appropriate for the needs.

Value for money

For several sites, improving the value for money of the services provided to the IPC cohort was identified as a goal. For some sites, participants identified the need for explicit cost savings, although they generally expected or planned for IPC to be cost-neutral.

The National Audit Office uses three criteria to assess the value for money of government spending, i.e. the optimal use of resources to achieve the intended outcomes: economy, efficiency and effectiveness (The National Audit Office 2015). The following diagram (Figure 5), taken from the National Audit Office website (The National Audit Office 2015), and used with permission, maps these onto the steps of the logic model for IPC.

Figure 5 Assessing Value for Money



On a practical level, the costs of IPC for individuals can be captured around service use from both health and social care data.

4.2. Health and social care population outcomes

There are a number of health outcomes that are expected to emerge as impacts of the IPC programme. These include a reduction in acute care needs, and a reduction in need for institutional care. At an economic level, such outcomes might be expected to lead to reduced cost of services. Understanding of these outcomes could be achieved by linking the individual identifiers for the population offered IPC to data from the Health and Social Care Information Centre (HSCIC), which collects data on health service use, as well as to indicators collected for reporting through the Adult Social Care Outcomes Framework (Department of Health 2014) (Figure 6).

Figure 6 Adult Social Care Outcomes Framework Indicators

- I am happy with the quality of my care and support and I know that the person giving me care and support will treat me with dignity and respect.
- I am supported to maintain my independence for as long as possible.
- I understand how care and support works, and what my entitlements are.
- I am in control of my care and support.
- I feel safe and secure.
- I have as much social contact as I want with people I like.

4.2.1. Availability of data and data governance

Although the issues for data and information governance are beyond the scope of this review, and at a national level there is much ongoing work to support this for IPC, it is worth highlighting some of the informal learning from the workshops as the topic came up frequently and is relevant to the implementation of IPC, the assessment of current needs and costs, as well as for both local and national evaluation.

In terms of IPC **implementation**, participants highlighted shared access to records for professionals as an issue across sites. If a care plan is maintained electronically it is important that professionals from both health and social care are able to access it and record notes if appropriate. Current systems for health and social care are usually separate.

In understanding the **current needs** of the IPC cohorts, participants identified access to electronic records from both health and social care in some sites as important.

Examples were given where electronic health service data was also used to understand the **current costs** of some cohorts.

In some areas, participants identified routine outcomes from electronic records as part of the local IPC **evaluation** plans.

During the workshops, participants described different local solutions for integrating health and social care systems, and noted that they were at different points in this process. It is unlikely that a single unified

system is going to develop across the country. Lonsdale et al. (2015) previously highlighted issues around joint working, particularly when different local and national information systems are incompatible.

Previous work in Cornwall found that difficulties in sharing information to support an integrated health and social care approach led to service users being asked for their explicit consent to share information, and an information sharing protocol being established to overcome the problem (Lonsdale et al. 2015). In one workshop, however, participants highlighted that the prior consent obtained for joint health and social care records was potentially not broad enough to allow their use across IPC.

In some sites, only patients who consented to the sharing of their health records for research purposes would be eligible for IPC. In practice, participants reported that all patients are happy to give individual consent, but acknowledged that it may be an issue if IPC is rolled out more broadly without explicit consent being sought.

Regarding individual preferences for the sharing of health information, recent RAND Europe research identified that across Europe there is little aversion to allowing nurses providing home-care being able to view health records, indicating that future policy directions to facilitate wide access and sharing could be acceptable (Patil et al. 2015).

4.3. Further technical comments

In this section we discuss some of the technical and methodological questions relevant for future evaluation, and the strengths and limitations of some possible methodological approaches.

4.3.1. *Personalised vs population-level outcomes*

As IPC is firmly embedded in the personalisation agenda – and away from standardised delivery of services – a natural tension arises between this very personalised care and its goals with the need to understand on a standardised national or population level whether IPC is working. The PAM tool goes some way to resolving this tension because it has the potential for both national measurement and relevance to the individual.

4.3.2. *Counterfactuals*

As well as collecting IPC data on activities, outputs and outcomes, a successful evaluation will also need to be able to compare IPC results to a counterfactual to determine what **impact** IPC has had compared to what would have happened had IPC had not taken place.

The measures of the ultimate impact of IPC are:

- Quality of life indicators
- Data on healthcare utilisation
- Data on institutional care utilisation
- Economic data on costs of health and social care usage.

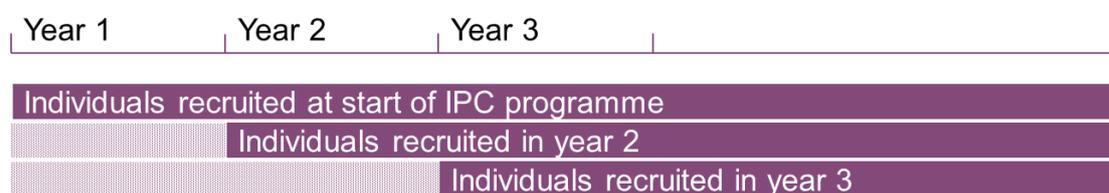
As IPC is a personalised intervention, person-level matching of controls will be the most relevant. Propensity score matching using routine health and social care data is one approach, as used in the

Integrated Care Pilots evaluation (Roland et al. 2013). A stratified propensity score matching approach, which essentially results in a quasi-randomised control trial within strata, is a second approach. However, past experience suggests that it may be difficult to generate a matched comparison group using routine data, given that those offered the intervention are selected locally based on a number of less tangible criteria which may not be captured by the routine data available.

For quality of life indicators such as EQ-5D there are no routine sources of data collected nationally at the person-level, although comparisons with population-level data may be possible because population level EQ-5D data is collected through the annual General Practice Patient Survey.

A third approach could be to restrict individual comparisons to only those ever offered IPC, thus limiting selection bias (discussed in more detail below in section 4.3.8). Those receiving the IPC programme at a given time point would be compared to those not receiving the IPC programme at that time point, but who were later recruited as the programme expanded (see Figure 7). The assumptions underlying this are that those recruited in later phases of the IPC programme are equivalent in need to those recruited at the beginning of the IPC programme and that there are no spill-over effects.

Figure 7 Individual comparisons within those ever offered IPC



4.3.3. Time scale

Integrated care does not happen quickly, and often depends on the development of new working relationships between groups of people, individuals, and organisations who have not worked together previously (Roland 2015a). The expected positive outcomes, particularly system level changes, will also take time to develop. At a practical level, immediate changes to existing block contracts for health and social care will not be possible, and at the individual level, the positive benefits of IPC would also be expected to accrue gradually. In personal health budget POET pilot sites, more positive outcomes were reported using the POET tool among people who had held budgets for over a year (Hatton & Waters 2013). A timescale of about three years is needed in order to assess whether integrated care interventions improve outcomes and are cost-effective (Roland 2015a).

4.3.4. Effect sizes

In order to be able to measure change, the expected effect sizes need to be meaningful and large enough to be measurable. Previous work on care planning and personal budgets suggest that effects may be detectable, but magnitudes of changes will not be large. This will have an impact on sample size requirements for future evaluation.

Care planning as an intervention

A recent Cochrane review of care planning found that it leads to improvements in certain physical and psychological health indicators and self-efficacy, but that the effects are not large. Care planning was found to be most effective when the intervention was more comprehensive, more intensive, and better integrated into routine care (Coulter et al. 2015).

Care planning was not found to be effective when it was simply a 'box-ticking' exercise. Most studies were focused on care plans and outcomes for people with single conditions or health problems. Doctors identified difficulties in knowing how to tailor care plans to people with multiple long-term conditions (Roland 2015b).

In a study of patients' experiences of care planning, no participants reported experiencing care planning discussions or that they had an explicit care plan (Newbould et al. 2012).

Personal budgets

The personal health budget evaluation found that for some outcome measures, personal health budgets were cost effective, but effect sizes were not consistently large across all outcomes (Jones et al. 2013).

4.3.5. Regression to the mean

In statistics, regression to the mean is the name for the fact that if something is an outlier on its first measurement, it will tend to be closer to the average on its second.

In the past, individuals who have been selected for case-management programmes, such as in the integrated care pilots (Roland et al. 2013), have been included particularly because they are ill or at high risk, or have large numbers of hospital admissions, for example. When monitored over time these individuals may subsequently have decreased service utilisation, simply because of this phenomenon of regression to the mean, rather than because of the effect of any intervention. This has previously been shown to have a potentially large effect (Roland et al. 2005). The issue is particularly a problem with small sample sizes. When samples are large, one can be more confident of who are genuine outliers as opposed to just part of the noise.

However, the ongoing health trajectory of the individuals recruited to IPC will be working in the opposite direction to regression to the mean. In several sites the patients in the cohorts identified have multiple severe health conditions, which may deteriorate over time. For example, in the RAND Europe evaluation of the Integrated Care Pilot programme, 8.4 per cent of patients selected for case management died in the six months following recruitment (RAND Europe & Ernst & Young LLP 2012).

4.3.6. Sample size

In some sites the numbers of individuals identified for IPC are very low, with target numbers for care planning and a personal budget lower still. In others a much larger cohort is being identified. In addition to the issues of regression to the mean and expected effect sizes discussed above, there are further issues that need to be considered.

First, if sites contribute very different numbers of individuals into a national evaluation, it is important also to measure and acknowledge outcomes at the site level, otherwise any national findings will simply reflect findings of the sample at the largest site, rather than provide a truly national picture.

Integrated care interventions to reduce hospital admissions are unlikely to be effective in the most high-risk groups because there may be little that can be done to reduce the need for admissions among the sickest individuals, but population-based approaches are unlikely to be cost-effective as the intervention would need to be offered to large numbers of people where the benefit is very small. High risk, but not the highest risk groups are likely to be those where integrated care is most cost-effective (Roland 2015a).

Similar thinking is probably required when considering the sample size requirements to assess IPC at different sites. Where only small numbers of people are participating then the population impact of any changes in health outcomes will be hard to detect, although changes in individual outcomes may be large.

In general, demonstrator sites would not have a sufficiently large cohort for some kind of site-level quantitative evaluation to be meaningful.

4.3.7. Heterogeneity across sites

Not only is IPC by design heterogeneous for the individual, the implementation of IPC is heterogeneous across sites. In general, for a research evaluation of a particular programme, the population, intervention, control group and outcome need to be specified. In IPC these all vary across sites. The ‘common’ aspects of IPC will need to be identified for the evaluation.

4.3.8. Selection bias

If selection of individuals for IPC is purposive instead of randomised – identifying people who would benefit most – then any selection of controls will be difficult, and the unique effect of IPC will be difficult to identify. Methodological approaches such as those suggested in section 4.3.2 are one possible approach to quantitative analysis.

4.3.9. Equality

This issue of selection bias for individuals included in IPC also highlights the issue of equality for all service users which should be captured through any evaluation. The following research questions will be important to consider:

- Are the characteristics or outcomes of eligible patients who opt not to join IPC different from those who choose to join?
- Are service users with strong support or advocacy networks (family, etc.) better able to benefit from IPC than those without?
- Does the focus on IPC distract resources from other services?
- Does IPC lead to ‘cherry-picking’ of profitable work by providers?
- Does empowerment of individuals and good communication help ensure that the rights of all service users are equally protected?

4.3.10. *Getting at the unique effect of IPC in the context of other changes*

The ongoing pace of change in the NHS over the coming years is unlikely to diminish, for example the recently announced programme to transform care for people with learning disabilities (NHS England 2015) will have an impact on several IPC cohorts when it is rolled out nationally. National programmes running alongside IPC will mean that identifying the unique effect of IPC may be difficult. In addition, several of the IPC sites are participating in other pilot projects or demonstrator programmes to improve care (for example Integrated Care Pioneers, Vanguard sites) and so locally it may also be difficult to separate the unique effect of IPC from other changes in the health economy.

4.3.11. *IPC successfully addresses previously unmet needs - leading to unchanged or increasing levels of service use*

At several sites one key aim for IPC was to identify and meet previously unmet needs among the target population. It is important to acknowledge that if IPC programmes are successful in this aim, this could lead to overall increases rather than decreases in service use. For example, a review of case-management interventions (the collaborative practice of care coordination by health professionals across community, primary and secondary care) found that in 9 of the 11 included trials there was no reduction in unplanned hospital admissions compared with usual care, despite strong anecdotal evidence of this approach being effective at avoiding admissions (Huntley et al. 2013, Roland 2013).

In a related commentary, the fact that new care models sometimes mean that previously unmet needs are addressed is highlighted as a possible issue (Roland 2013). Although IPC for some individuals may be extremely successful in improving health outcomes, and therefore reducing the need for hospital admissions or other services, for others, IPC may in fact lead to the new provision of appropriate services to address previously unmet needs, which may in turn result in an increase in (appropriate) activity, for example an increase in admissions to hospital.

Overall there may be increases, no changes, or decreases in population health or social care services use, despite IPC working appropriately.

4.3.12. *Survey tools*

Several of the proposed IPC evaluation tools involve survey data collection. Survey nonresponse and other methodological concerns are frequently cited as reasons why people do not believe or act on survey findings (Asprey et al. 2013, Iacobucci 2013). This section considers five common challenges to the validity of survey findings, and suggests a few simple approaches to minimise concerns about the validity of findings from survey research.

Challenge 1: You didn't ask enough people/let's do some more research.

Surveys need an adequate sample size to ensure that findings are unlikely to be due merely to chance.

- Research should be designed to allow a sufficiently large sample size of respondents to be collected.

- Use appropriate statistical methods to describe the precision of the results, for example 95 per cent confidence intervals where appropriate to allow the uncertainty in any findings to be quantified.
- If sample sizes are very small, consider collecting data over a bigger geographical area or longer time period.

Challenge 2: The wrong people responded.

Nonresponse bias is a big challenge to the validity of survey research – that responses from people who do respond to the survey are systematically different from those who do not respond.

- Ensure best practice to minimise nonresponse bias (Edwards et al. 2002).
- Send a personalised invitation from a named senior researcher and two follow-up invitations to initial nonresponders.
- Ensure that the survey is clearly worded, well presented and not too long.
- Ensure that the survey is only sent to relevant participants, and that questions are relevant for the individuals to whom it is sent. Relevance and interest in the survey topic areas are the strongest predictors of survey response.
- A low response rate alone does not necessarily mean that findings are biased (Groves 2006, Groves & Peytcheva 2008).
- Acknowledge that some groups may be under- or over-represented among responders. Men, people from ethnic minorities, and those living in more deprived areas are less likely to respond to surveys (Elliott et al. 2009).
- Describe the characteristics of both respondents and nonrespondents to allow the magnitude of any possible bias to be understood (Halbesleben & Whitman 2013), and adjust for this in the analysis if appropriate.

Challenge 3: You asked the wrong people/ I don't think the findings are relevant to me.

This is a concern about the generalisability of the findings from survey research.

- Ensure that all relevant participants are invited to respond to the survey.
- Clearly characterise the respondents to the survey and, if appropriate, compare these to external sources to understand to whom the findings may be most appropriately applied.

Challenge 4: Your survey is asking the wrong questions.

This is a concern about measurement bias, that the survey tool or the way that the data was collected means that some people respond in a different way to others.

- Use a consistent form of data collection across all groups.
- Use validated tools where appropriate.
- Pilot the survey among relevant participants to identify if there are any systematic problems in how responses are received.

Challenge 5: Your survey says that A causes B, I think that B causes A. Or does C cause both A and B?

Surveys are typically carried out at a single point in time, and reverse causality and the impact of confounding factors are challenges to analyses of associations from survey research.

- Ensure that there are no inappropriate inferences from cross-sectional data, and triangulate survey findings with other sources of evidence where appropriate. Consider more than one time point.
- Confounding is a challenge to the findings from observational research, including surveys, which looks at the relationship between interventions and outcomes. If something is related to both the intervention and the outcome, but not on the causal pathway between the two, it is a possible confounding factor, and needs to be considered in the analysis.

4.3.13. Normalisation Process Theory

A further evaluation approach, particularly relevant for the ‘commissioning activities’ strand of the logic model, could draw on Normalisation Process Theory (NPT) (May 2006), to specifically understand whether and how new ways of thinking, acting and organising are implemented and become integrated (‘routine embedding’) into usual practice. NPT proposes that (i) innovations become embedded in practice as a result of people working, individually and collectively, to implement them; (ii) the work of implementation is operationalised through four generative mechanisms identified as: coherence, cognitive participation, collective action and reflexive monitoring; and that (iii) organising structures and social norms specify the rules and roles that frame action. NPT can provide good theoretical understanding of whether and how change is caused, while identifying weak links in the causal chain.

4.3.14. Collections of case studies

Collections of case studies of good practice, or of early implementation of a new programme, can be important for more than just evaluation. One site highlighted that a collection of case studies where IPC was working well locally would be helpful to raise awareness and increase local support and buy-in for IPC. This is something that IPC has done successfully in initial documentation (NHS England 2014b), highlighting successful case studies from personal health budgets, and could certainly be built upon with stories from early participants in IPC. Workshop attendees at several sites expressed a desire to improve communication and sharing across the demonstrator programmes. Case studies might also represent a semi-formal mechanism to allow this to occur.

Case studies could also be considered as part of the evaluation of IPC nationally. Borrowing an example from universities, case studies were used successfully to evaluate research impact across institutions (Manville et al. 2015), and Martin et al. (2015) recently highlighted ‘soft intelligence’ as more important than metrics in understanding healthcare quality. This approach could be considered as an alternative or complement to more traditional evaluation methods, acknowledging, however, the major weakness of case studies: it is problematic to extrapolate and generalise lessons from individual stories, unless a large number of studies are carried out, which itself has cost implications.

5. Discussion

5.1. Are the demonstrator sites evaluation-ready?

There are differences in how well prepared each site is for an evaluation. All are in a position to draw upon and adapt the generic theory of change model and therefore have a model that can be evaluated. However, sites have recognised that they are unlikely to get it right first time and that they will want to adapt and improve their activities. To this extent, an ongoing evaluation could help produce site-level data to support local decision-makers. In addition, a scheme-wide evaluation could draw together the lessons from the first cohort to support future decision-making nationally. Both of these aims would be supported by common measurements collected across the sites and a comparative evaluation.

Preparation of sites for evaluation from the start of a programme (such as the work presented in this report) is also a very useful approach and is likely to contribute to a successful evaluation in the future.

5.2. Conclusions

The IPC demonstrator sites do not exist in isolation; integrated personal commissioning is one of several complementary programmes and new models of care being introduced or piloted across the country. IPC builds on previous programmes, including personal health budgets and the Better Care Fund. The learning from IPC and the future IPC evaluation should be outward looking, and may contribute learning to other programmes as well, and so more broadly than to the IPC programme alone.

The implementation, evaluation and learning from IPC are iterative processes – sites are not expected to simply set up IPC in isolation. A common theme across workshops was a hope that demonstrator sites could work together and share learning with each other as issues emerge. We recommend that this iterative process of local learning, maintaining communication between sites, and with the NHS England programme team, form part of the evaluation process. Clear practical communication channels are important here.

We also recommend that the methodological approaches and issues identified and described in this report feed into the design of the future evaluation for IPC, both the evaluation of the demonstrator sites and a national evaluation. Although there are substantial methodological challenges to a national quantitative impact evaluation, acknowledging and addressing these from the outset will also strengthen the learning from future work.

We hope the information above provides NHS England with a foundation on which they and the sites can build in order to continue to define IPC and develop standard models for delivery, which in turn will inform the implementation of IPC over the coming years.

References

- Armstrong, N., C. Tarrant, G. Martin, B. Manktelow, and L. Brewster. 2015. *Independent Evaluation of the Feasibility of Using the Patient Activation Measure in the NHS in England: Early Learning Report*. NHS England: the University of Leicester: the Health Foundation.
- Asprey, A., J. L. Campbell, J. Newbould, S. Cohn, M. Carter, A. Davey, and M. Roland. 2013. 'Challenges to the Credibility of Patient Feedback in Primary Healthcare Settings: A Qualitative Study.' *Br J Gen Pract* 63 (608):e200-8. doi: 10.3399/bjgp13X664252
- Billings, J., and E. de Weger. n.d. *Evaluation framework Kent Integration Pioneers: Draft 2*. CHSS University of Kent.
- Burckhardt, C. S., and K. L. Anderson. 2003. 'The Quality of Life Scale (QOLS): Reliability, Validity, and Utilization.' *Health Qual Life Outcomes* 1:60. doi: 10.1186/1477-7525-1-60
- Burt, J., M. Roland, C. Paddison, D. Reeves, J. Campbell, G. Abel, and P. Bower. 2012. 'Prevalence and Benefits of Care Plans and Care Planning for People with Long-Term Conditions in England.' *J Health Serv Res Policy* 17 Suppl 1:64-71. doi: 10.1258/jhsrp.2011.010172
- COE4CCN. 2015. 'Center of Excellence on Quality of Care Measures for Children with Complex Needs: Family Experiences with Coordination of Care (FECC) measurement set.' As of 9 October 2015: <http://www.seattlechildrens.org/research/child-health-behavior-and-development/mangione-smith-lab/measurement-tools/>
- Coulter, A., V. A. Entwistle, A. Eccles, S. Ryan, S. Shepperd, and R. Perera. 2015. 'Personalised Care Planning for Adults with Chronic or Long-Term Health Conditions.' *Cochrane Database Syst Rev* 3:CD010523. doi: 10.1002/14651858.CD010523.pub2
- Davies, R., and L. Payne. 2015. 'Evaluability Assessments: Reflections on a Review of the Literature.' *Evaluation* 21 (2):216-231. doi: 10.1177/1356389015577465
- Department of Health. 2013. 'Press Release: Integration Pioneers Leading the Way for Health and Care Reform.' As of 9 October 2015: <https://www.gov.uk/government/news/integration-pioneers-leading-the-way-for-health-and-care-reform--2>
- Department of Health. 2014. 'Adult Social Care Outcomes Framework (ASCOF) 2015 to 2016.' As of 9 October 2015: <https://www.gov.uk/government/publications/adult-social-care-outcomes-framework-ascof-2015-to-2016>
- Dronavalli, Mithilesh, and Sandra C Thompson. 2015. 'A Systematic Review of Measurement Tools of Health and Well-Being for Evaluating Community-Based Interventions.' *Journal of Epidemiology and Community Health*. doi: 10.1136/jech-2015-205491
- Edwards, P., I. Roberts, M. Clarke, C. DiGiuseppi, S. Pratap, R. Wentz, and I. Kwan. 2002. 'Increasing Response Rates to Postal Questionnaires: Systematic Review.' *BMJ* 324 (7347):1183.
- Elliott, M. N., A. M. Zaslavsky, E. Goldstein, W. Lehrman, K. Hambarsoomians, M. K. Beckett, and L. Giordano. 2009. 'Effects of Survey Mode, Patient Mix, and Nonresponse on CAHPS Hospital Survey Scores.' *Health Serv Res* 44 (2 Pt 1):501-18. doi: 10.1111/j.1475-6773.2008.00914.x
- Forder, J., K. Jones, C. Glendinning, J. Caiels, E. Welch, K. Baxter, J. Davidson, K. Windle, A. Irvine, D. King, and P. Dolan. 2012. *Evaluation of the Personal Health Budget Pilot Programme*.

- Grant, R. L., and A. Bowling. 2011. 'Challenges in Comparing the Quality of Life of Older People between Ethnic Groups, and the Implications for National Well-Being Indicators: A Secondary Analysis of Two Cross-Sectional Surveys.' *Health Qual Life Outcomes* 9:109. doi: 10.1186/1477-7525-9-109
- Groves, R. M. 2006. 'Nonresponse Rates and Nonresponse Bias in Household Surveys.' *Public Opinion Quarterly* 70 (5):646-675
- Groves, R. M., and E. Peytcheva. 2008. 'The Impact of Nonresponse Rates on Nonresponse Bias - A Meta-Analysis.' *Public Opinion Quarterly* 72 (2):167-189. doi:10.1093/Poq/Nfn011
- Halbesleben, J. R. B., and M. V. Whitman. 2013. 'Evaluating Survey Quality in Health Services Research: A Decision Framework for Assessing Nonresponse Bias.' *Health Services Research* 48 (3):913-930. doi:10.1111/1475-6773.12002
- Hatton, C., and J. Waters. 2013. *The POET Surveys of Personal Health Budget Holders and Carers*.
- Hibbard, J. H., E. R. Mahoney, J. Stockard, and M. Tusler. 2005. 'Development and Testing of a Short Form of the Patient Activation Measure.' *Health Serv Res* 40 (6 Pt 1):1918-30. doi: 10.1111/j.1475-6773.2005.00438.x
- Hibbard, J. H., J. Stockard, E. R. Mahoney, and M. Tusler. 2004. 'Development of the Patient Activation Measure (PAM): Conceptualizing and Measuring Activation in Patients and Consumers.' *Health Serv Res* 39 (4 Pt 1):1005-26. doi: 10.1111/j.1475-6773.2004.00269.x
- Huntley, A. L., R. Thomas, M. Mann, D. Huws, G. Elwyn, S. Paranjothy, and S. Purdy. 2013. 'Is Case Management Effective in Reducing the Risk of Unplanned Hospital Admissions for Older People? A Systematic Review and Meta-Analysis.' *Fam Pract* 30 (3):266-75. doi: 10.1093/fampra/cms081
- Iacobucci, G. 2013. 'Caution Urged Amid Wide Variation in Response Rates to Friends and Family Test.' *BMJ* 347:f4839. doi: 10.1136/bmj.f4839
- International Wellbeing Group. 2013. *Personal Wellbeing Index: 5th Edition*. Melbourne: Australian Centre on Quality of Life, Deakin University.
- IPC Market Analysis Centre. 2014. *DCMQC Briefing Paper 1: The Basics of Market Facilitation*. As of 9 October 2015:
http://ipc.brookes.ac.uk/services/documents/The_Basics_of_market_facilitation.pdf
- Jones, K., J. Forder, J. Caiels, E. Welch, C. Glendinning, and K. Windle. 2013. 'Personalization in the Health Care System: Do Personal Health Budgets have an Impact on Outcomes and Cost?' *J Health Serv Res Policy* 18 (2 Suppl):59-67. doi: 10.1177/1355819613503152
- Ling, T. 2012. 'Evaluating Complex and Unfolding Interventions in Real Time.' *Evaluation* 18 (1):79-91. doi: 10.1177/1356389011429629
- Lonsdale, J., D. Schweppenstedde, C. van Stolk, B. Guerin and M. Hafner. 2015. *One Place, One Budget? Approaches to pooling resources for public service transformation*. Santa Monica, Calif.: RAND Corporation, RR-1017-LGA. As of April 14, 2016:
http://www.rand.org/pubs/research_reports/RR1017.html
- Manville, C., S. Guthrie, M.-L. Henham, B. Garrod, S. Sousa, A. Kirtley, S. Castle-Clarke and T. Ling. 2015. *Assessing impact submissions for REF2014: An evaluation*. Santa Monica, CA: RAND Corporation, RR-1032-HEFCE. As of 25 March 2015:
http://www.rand.org/pubs/research_reports/RR1032.html
- Martin, G. P., L. McKee, and M. Dixon-Woods. 2015. 'Beyond Metrics? Utilizing 'Soft Intelligence' for Healthcare Quality and Safety.' *Soc Sci Med* 142:19-26. doi: 10.1016/j.socscimed.2015.07.027.
- May, C. 2006. 'A Rational Model for Assessing and Evaluating Complex Interventions in Health Care.' *BMC Health Serv Res* 6:86. doi: 10.1186/1472-6963-6-86
- Moore, G. F., S. Audrey, M. Barker, L. Bond, C. Bonell, W. Hardeman, L. Moore, A. O'Cathain, T. Tinati, D. Wight, and J. Baird. 2015. 'Process Evaluation of Complex Interventions: Medical Research Council Guidance.' *BMJ* 350:h1258. doi: 10.1136/bmj.h1258

- Newbould, J., J. Burt, P. Bower, T. Blakeman, A. Kennedy, A. Rogers, and M. Roland. 2012. 'Experiences of Care Planning in England: Interviews with Patients with Long Term Conditions.' *BMC Fam Pract* 13:71. doi: 10.1186/1471-2296-13-71
- NHS England. 2014a. *The Five Year Forward View*. As of 9 October 2015: <http://www.england.nhs.uk/wp-content/uploads/2014/10/5yfv-web.pdf>
- NHS England. 2014b. *Integrated Personal Commissioning Prospectus: Making a Reality of Health & Social Care Integration for Individuals*.
- NHS England. 2015. *Supporting People with a Learning Disability and/or Autism who have a Mental Health Condition or Display Behaviour that Challenges*. As of 9 October 2015: <http://www.england.nhs.uk/wp-content/uploads/2015/07/ld-draft-serv-mod.pdf>
- Patil, S., B. Patrui, H. Lu, F. Dunkerley, J. Fox, D. Potoglou and N. Robinson. 2015. *Public perception of security and privacy: results of the comprehensive analysis of PACT's pan-European survey*. Santa Monica, Calif: RAND Corporation. RR-704-EC. As of 27 January 2015: http://www.rand.org/pubs/research_reports/RR704.html
- RAND Europe and Ernst & Young LLP. 2012. *National evaluation of the DH Integrated Care Pilots*. Santa Monica, Calif: RAND Corporation. TR-1164-DH. As of 22 March: http://www.rand.org/pubs/technical_reports/TR1164.html
- Reeves, D., M. Hann, J. Rick, K. Rowe, N. Small, J. Burt, M. Roland, J. Protheroe, T. Blakeman, G. Richardson, A. Kennedy, and P. Bower. 2014. 'Care Plans and Care Planning in the Management of Long-Term Conditions in the UK: A Controlled Prospective Cohort Study.' *Br J Gen Pract* 64 (626):e568-75. doi: 10.3399/bjgp14X681385
- Roland, M. 2013. *Case Management Doesn't Reduce Emergency Admissions*. As of 9 October 2015: <http://www.cchsr.iph.cam.ac.uk/123>
- Roland, M. 2015a. *Why Don't Care Coordination Interventions Work?* As of 9 October 2015: <http://www.cchsr.iph.cam.ac.uk/2493>
- Roland, M. 2015b. *Wrong Kind of Care Plans?* As of 9 October 2015: <http://www.cchsr.iph.cam.ac.uk/2487>
- Roland, M., M. Dusheiko, H. Gravelle, and S. Parker. 2005. 'Follow Up of People Aged 65 and Over with a History of Emergency Admissions: Analysis of Routine Admission Data.' *BMJ* 330 (7486):289-92. doi: 10.1136/bmj.330.7486.289
- Roland, M., R. Lewis, A. Steventon, G. A. Abel, J. L. Adams, M. Bardsley, L. Brereton, X. Chitnis, A. Conklin, L. Staetsky, S. Tunkel, J. Newbould and Ernst & Young LLP. 2013. *Does integrated care deliver the benefits expected? Findings from 16 integrated care pilot initiatives in England*. Santa Monica, Calif: RAND Corporation. RB-9703-DH. As of 14 April: http://www.rand.org/pubs/research_briefs/RB9703.html
- The National Audit Office. 2015. *Assessing value for money*. As of 9 October 2015: <http://www.nao.org.uk/successful-commissioning/general-principles/value-for-money/assessing-value-for-money/>
- Thomas G.P., C.L. Saunders, M.O. Roland, C.A. Paddison. 2015. 'Informal Carers' Health-Related Quality of Life and Patient Experience in Primary Care: Evidence from 195,364 Carers in England Responding to a National Survey'. *BMC Fam Pract* 16: 62.