

Advice to the Department of Health and NHS England on the integrated personal commissioning (IPC) programme independent evaluation

Nicholas Mays

Julien Forder

Karen Jones

Sandra Mounier-Jack

Mary Alison Durand

Policy Innovation Research Unit, London School of Hygiene and Tropical Medicine

Economics of Health and Care Research Unit, University of Kent at Canterbury

12 May 2016

Funding acknowledgement and disclaimer

This work was funded by the Policy Research Programme of the Department of Health for England, via its core support for the Policy Innovation Research Unit and the Economics of Health and Care Research Unit. This is an independent report commissioned and funded by the Department of Health. The views expressed are not necessarily those of the Department and its partners.

Contents

1	Introduction	2
2	What is integrated personal commissioning (IPC)?	3
3	What we can learn from evaluation of similar policy initiatives	6
3.1	The evaluation of the individual budget pilot programme	6
3.2	The evaluation of the personal health budget pilot programme	6
3.3	The evaluation of the Direct Payments in Residential Care trailblazers.....	7
4	Issues to consider for an evaluation of IPC.....	8
4.1	Purpose and scope	10
4.2	Complexity	11
4.3	Scale, variation and implementation	11
4.4	Timing.....	13
	Annex. Scoping study and early evaluation of the integrated personal commissioning (IPC) programme	15
	References	17

1 Introduction

Integrated personal commissioning (IPC) is intended to be a new way of working to better coordinate health, social care and other services around individuals with long-term conditions. It aims to bring together health and social care funding to be used to meet the objectives of a person-centred care plan, designed with the full involvement of individuals and their families. The IPC programme will also work to develop community and voluntary sector resources that seek to support people to self-manage their condition.

This report to the Department of Health (DH) and NHS England provides expert advice relevant to the independent longer-term evaluation of the Integrated Personal Commissioning (IPC) programme. It is designed to assist research teams applying to undertake the summative evaluation, which is being commissioned by the Department of Health (DH) Policy Research Programme (PRP), in the development of their proposals. It discusses what is likely to be feasible in terms of evaluation approaches.

The research draws mainly on the evaluation methodology literature and the authors' previous experience with evaluating similar policy initiatives – in particular, the individual (social care) budgets pilot projects' evaluation (IBSEN) (Glendinning et al, 2008), the evaluation of the personal health budget (PHB) pilot programme (Forder et al, 2012) and the evaluation of direct payments in the residential care trailblazers (Wittenberg et al, 2015).

This report is an early output from a scoping study and early evaluation of the IPC programme being undertaken jointly by the Policy Innovation Research Unit (PIRU) and the Economics of Health and Care Research Unit (ESHCRU). See the Annex for the aims and objectives of the early evaluation. The final report of the early evaluation will be produced in autumn 2016 and will be made available to the team successful in winning the contract for the longer-term evaluation of the IPC programme along with any other outputs from the work.

This advisory note, which represents an interim output from the early evaluation, is structured as follows. First, we re-cap the policy aims and description of the IPC programme and their implications for the longer-term evaluation. Second, we consider recent experience of evaluating similar policies in England to identify some of the factors that are likely to shape the choice of evaluation approach. Third, we draw out some implications for the longer-term evaluation of the IPC programme, drawing on the experience of the previous evaluations and on the general lessons from the literature about policy evaluation.

2 What is integrated personal commissioning (IPC)?

According to NHSE's website:

'Integrated Personal Commissioning consolidates a shift in power to people who use these services [NHS, local authority and voluntary sector] to help them shape care that is effective and meaningful to them in their lives. It builds on and brings together work that has already started to explore new funding models and places that have taken the lead in implementing personal budgets in the NHS. It aims to bring health and social care together, identifying the totality of expenditure at the level of the individual, giving people more control over how this is used and enabling money to be spent in new ways to achieve the three programme goals.'

(<https://www.england.nhs.uk/commissioning/ipc/> accessed 8 February 2016)

It appears that the underlying purpose of introducing IPC into the system in the shape of a 'demonstrator programme' rather than as a (time-limited) 'pilot' is to signal that IPC is intended to become part of the mainstream of the system.

The stated desired outcomes of the IPC programme are:

- to improve the quality of life of people with complex needs and their carers;
- to enable people with complex needs, their families and their carers to achieve important goals through greater involvement in their own care so that they are able to design support around their own 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.

To achieve this, the programme comprises the following inter-related elements of an individualised commissioning process involving a range of organisations and individuals:

- a single integrated, individual needs assessment across health, social care, education and other areas of personal need;
- a single individual care plan designed with the full involvement of individuals and their families;
- a system for information sharing across sectors at the individual level developed by local commissioning organisations;
- a financial model for budget setting across sectors;
- allocation of a single integrated budget (excluding emergency health care) based on the above;
- review, shaping and development of the local provider market by commissioning organisations;
- development of community and voluntary sector assets to enlarge the choice set available to individuals; and
- the ability for individuals to choose providers from the developed local market including incumbents and new entrants.

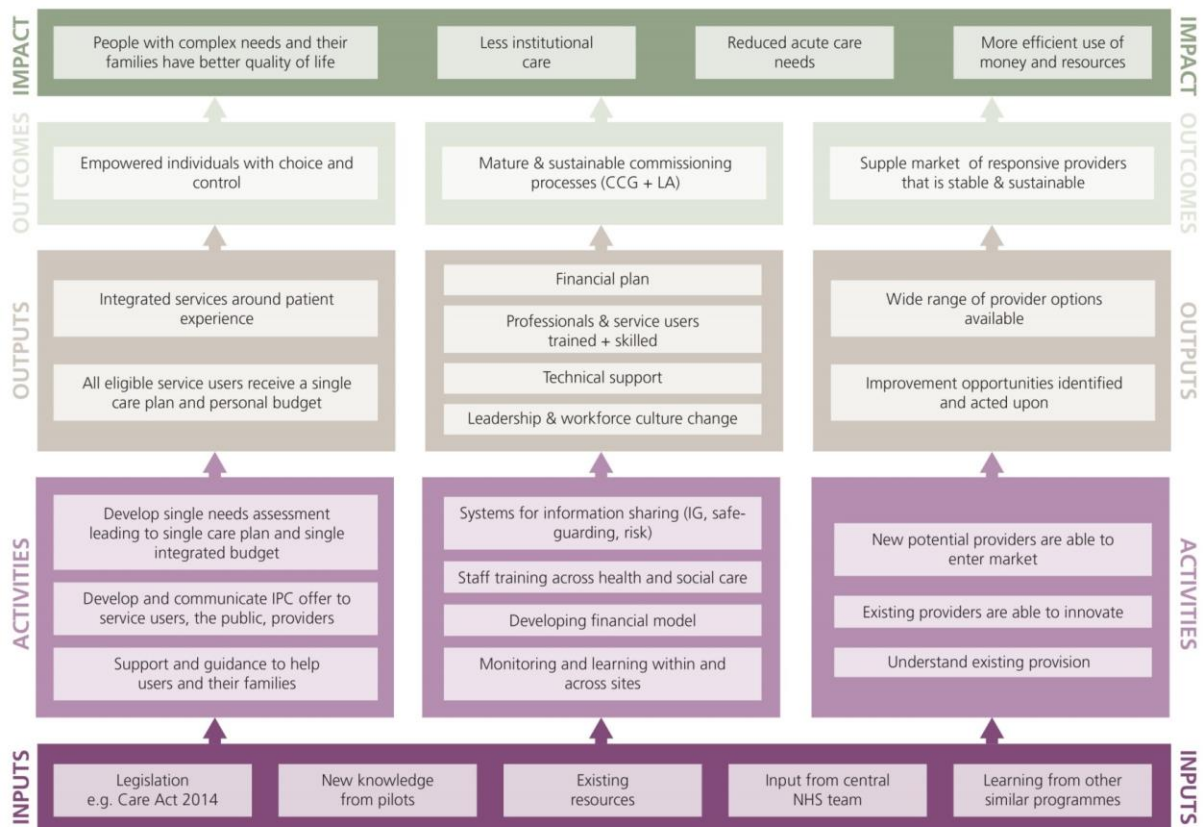
From this description, it is apparent that while the calculation and allocation of integrated personal budgets are important to the IPC model, activities relating to the budget are only part of a wider programme of activity.

RAND Europe, in its work with the IPC 'demonstrator sites' commissioned by NHS England during the summer of 2015, identified from the logic models of each site the following three main groups of activities related to IPC that should be considered within the scope of the evaluation (see also Figure below):

1. Activities around the service user, including a single care plan and personal budget, designed to lead 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 designed to 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 is more likely to be driven by patients' needs and choices (Sim et al, 2015, 17).

Thus, IPC is a programme that combines a number of elements of previous policies and initiatives (e.g. 'personalisation', integrated care, personal budgets, self-directed care, market development and testing, user choice, etc.). The relative novelty lies in the way that these are to be combined in a commissioning process. Logically, the second and third activities would most likely precede the first set of activities relating to care planning and offering an integrated personal budget to service users. They are also potentially capable of being observed by a research team before large numbers of users are enabled to hold personal budgets. Thus these activities could form an early focus of any longer term evaluation ahead of being able to undertake a summative assessment of the IPC programme as required by DH PRP's invitation to tender.

Figure: Integrated Personal Commissioning Programme logic model



Source: Sim M, Saunders K, Winpenny E, Ling T. (2015) *Future evaluation of the Integrated Personal Commissioning Programme: mapping the logic and assessing evaluability*. Confidential draft report to NHS England. Cambridge: RAND Europe

3 What we can learn from evaluation of similar policy initiatives

Three Department of Health-funded evaluations provide insights relevant to designing the longer-term summative evaluation of the IPC programme currently being commissioned by DH PRP.

3.1 The evaluation of the individual budget pilot programme

The evaluation of the individual budget pilot programme (IBSEN) was a two-year study (2006-2008) that compared the experiences of people who were offered an individual budget with the experiences of those receiving conventional services (Glendinning et al. 2008). Overall, the evaluation found some evidence to suggest that individual budgets had an impact on outcomes and costs. However, the impact of individual budgets varied between client groups.

The evaluation adopted a mixed methods approach involving in-depth qualitative interviews among service users and organisational representatives to help interpret the statistical analysis of data on outcomes. An RCT was included in the design to explore whether individual budgets offered a better way of supporting older people and other adults with social care needs, compared to conventional methods of funding.

While an RCT can help ensure a robust evaluation, reservations among the local authority staff involved in the programme potentially delayed the implementation of the RCT. Also, policy announcements supporting individual budgets as the future model undermined the case for a summative evaluation from the perspective of local sites.

A total of 2,521 service users were approached by the 12 pilot sites to request their participation in the research; of these, 1,594 agreed to participate. However, there were significant delays in returning baseline data to the evaluation team, so that 1,356 baseline forms were available in sufficient time to complete the follow-up 6-month outcome interviews. In total, 959 service users were included in the quantitative sample; 510 in the individual budget group and 449 in the comparison group. The requirement that pilot sites recruited a target number of people offered an individual budget within a short timescale was a challenge for the pilot sites and the evaluation.

The study did show, nonetheless, that an RCT was possible in a practice area where such methods had previously not been much used.

3.2 The evaluation of the personal health budget pilot programme

To explore the impact that personal health budgets have on health and social care outcomes, the DH commissioned a three-year evaluation (2009-2012) that ran alongside the pilot programme (Forder et al. 2012; Jones et al. 2013). The aim of the evaluation (PHBE) was to identify whether personal health budgets deliver better health and care outcomes when compared to conventional service delivery; and, if so, to identify the most effective implementation process (Forder et al., 2012). Twenty sites from across the pilots were selected to be *in-depth* evaluation sites.

Recognising the significant research demand placed on sites, each received funding of £100,000 per year (for three years). It was important to engage closely with sites and local PHB project managers to ensure a good level of support.

An intervention – control design was used with selection either by randomisation or more pragmatically by selecting similar patients from PHB pilot-participating and non-participating health professionals. In this case, data were collected at baseline and follow-up to allow a difference-in-difference analysis. Overall, there were 1,171 people in the personal health budget group and 1,064 in the comparison group.

Similar to the evaluation of the individual budget pilot programme (IBSEN), the quantitative data collection was complemented by data from in-depth interviews among budget holders, informal carers and organisational representatives.

Part of the evaluation was designed to inform the national roll-out of PHBs by providing information about how the initiative *should* be implemented if shown to be sufficiently successful. Five implementation models (i.e. different configurations of PHB arrangement) were identified by the evaluation team and compared. The sub-group analyses between implementation models were a little under-powered but still produced important significant results.

As with IBSEN, policy announcements about rolling-out PHBs during the evaluation shifted attention to the implementation analysis results, rather than the overall impact of PHBs.

Both evaluations highlighted the importance of adopting a mixed-methods approach to be able to fully explore the impact of a new initiative. Furthermore, despite highly complex interventions, the methods appeared to be able to produce robust results in both cases, in part because there were sufficient numbers of service users participating in the studies and because of the methods used.

3.3 The evaluation of the direct payments in residential care trailblazers

The evaluation of the Direct Payments (DPs) in residential care trailblazers began in January 2014 and will present its final report in June 2016.

The evaluation aimed to: understand the different ways in which DPs are being offered to residents of care homes and to examine the challenges arising from implementing DPs (process evaluation); assess the impacts of DPs in residential care (impact evaluation); and examine, as far as possible, the relative costs and cost-effectiveness of different approaches (economic evaluation).

This evaluation also used mixed methods. Data collection included surveys (baseline, with some including two rounds of follow up questionnaires) of: services users with and without DPs; family members of users with and without DPs; easy read versions of questionnaires for service users; and a survey of managers and owners of care homes with and without service users with a DP. Semi-structured interviews were conducted with service users, family members, brokerage personnel, frontline staff and project leads in councils, and managers and owners of care homes. Finally, administrative data were collected from sites including data on the costs of facilitating DPs. The main limitation of the evaluation was the small number of people offered a DP compared with original plans. This particularly affected the ability to study the outcomes of DP receipt quantitatively and to estimate cost-effectiveness.

4 Issues to consider for an evaluation of IPC

Based on the policy evaluation literature, the suitability of different evaluation options depends on a range of factors, including:

- Desired purpose and scope of evaluation
- Degree of complexity of the policy or programme
- Scale of implementation, variation between pilot sites in how the programme is being implemented and readiness of sites to participate in new data collection
- Site organisation
- Ethics and governance processes required
- Routine data availability at sites

The main types of evaluation are:

- experimental studies with random allocation of study participants – randomised controlled trials (RCTs).
- specific trials or studies with non-random allocation – non-randomised studies (NRS), where the new intervention is being specifically tested but the control group is drawn from people outside the study population (e.g. a neighbouring area).
- non-experimental observational studies where the intervention is in routine operation and the control group are people not receiving the intervention for reasons outside the evaluation.
- process evaluations, which usually involve the development of classification frameworks and the subsequent population of those frameworks using descriptive accounts of the programme, the aims of the programme, the resources used, etc. (Moore et al. 2015).
- case studies involving qualitative methods (e.g. realist evaluation), usually with in-depth analysis of the intervention/subject within a specific context.

Depending on how the above factors apply to the subject of the evaluation, there are a number of design options for policy evaluations, summarised in Table 1 below.

This framework and the experience from previous evaluations help identify a set of issues that need to be considered when developing the evaluation of the IPC programme.

Table 1. Potential options for policy evaluation

Factor	Type	Suggested evaluation option(s):
Purpose and scope of evaluation	Implementation issues	Process
	Overall impact (e.g. cost-effectiveness)	RCT ^a or possibly: NRS ^b or observational (e.g. controlled before and after)
	Context-specific potential impact	Case study/qualitative, NRS or possibly observational, realist evaluation
	Best configuration (of IPC model) in practice	NRS or RCT ^c , mixed methods, realist evaluation
	Wider implications	Case study/qualitative, observational, mixed methods
Degree of complexity	Complicated	Impact – any
	Complex	Impact with sub-groups (multi-arm RCT, NRS, observational, realist) plus process evaluation
	Highly Complex (system)	Case study/qualitative with use of routine data/indicators
Scale/variation	Whole population	Observational, case study/qualitative
	Pilot with significant participation	RCT, NRS
	Small scale pilot	Case study/qualitative
<i>Ex ante</i> expectations about benefit/effectiveness (ethics)	Little <i>a priori</i> knowledge/expectations of benefits, high level of uncertainty	Experimental/quasi-experimental (RCT or NRS)
	Strong <i>a priori</i> expectations of benefits	Observational, case study/qualitative
Budget and required timeframe for the evaluation	Limited e.g. 2 years or less	Observational, Case study/qualitative
	Open-ended: 3 years or more	Experimental plus process
Existing data	Little data on outcomes and costs	Any but not observational
	Substantial existing (e.g. admin) data	Any, and especially observational

Notes: ^a RCT, randomised controlled trial; ^b NRS, non-randomised study

^c but only if eligibility is not being tested (because people are allocated randomly to model configurations in an RCT);

4.1 Purpose and scope

NHS England introduced IPC as a ‘demonstrator’ programme rather than a ‘pilot’. This indicates that an important purpose of the independent evaluation is to find out which parts of the programme, in combination (i.e. which approaches to IPC), should be retained, removed or changed, and/or which configurations of IPC work better than other IPC configurations (Ettelt et al, 2015). In particular, if the evaluation were *purely* for demonstration, then an option would be only to compare different IPC models, and forgo the need to a ‘usual care’ control group, rather than to also establish whether the IPC policy works overall (i.e. IPC as compared to ‘usual’ care). A perhaps more widely relevant option would be for the evaluation to be able to compare different IPC models with the status quo (usual care). Awareness of the ‘demonstrator’ nature of the programme might help the evaluators avoid the need to alter the focus of the evaluation after it has started. It would certainly highlight the need for the evaluation to be able to identify and define a number of ‘implementation models’ or ‘configurations’.

The changing policy context regarding both the IB and PHB evaluations led to a shift in attention from the overall impact of the policies to a focus on which implementation model was most (cost-) effective.

Given the multi-faceted nature of the IPC programme, the precise scope of, and emphasis within, the evaluation also merits careful thought. At one end, there could be a comprehensive evaluation of the entire programme with all its components. Alternatively, the evaluation could focus on the key components, especially the commissioning mechanism, or it could look particularly at those areas where knowledge relevant to IPC is most limited, in particular:

- Market development/stimulating the supply side. For some policy makers this may be a high priority because it has not been a feature of previous personal budget initiatives and there is virtually no research in England on such activities.
- Integrating funding streams across organisations and sectors. The financial model and especially the pooling of budgets between health and social care remains a challenge despite permissive legislation. Accounting, professional and cultural barriers are likely and should be investigated.
- Proper costing of fair personal budgets. Personal budget calculation was particularly difficult in the PHB pilots and sites developed a number of different approaches.
- The choice of a direct payment rather than a managed budget.

The PHB evaluation experience is also relevant in regard to potential outcome measures used in the analysis. It was clear from the study that the impact of policies like PHBs – and indeed IPC – was broader than just its health effects (especially just physical health effects), at least over the 12-month follow-up period that was feasible within the study. The implication is that a range of outcome measures should be used, including care-related quality of life measures such as ASCOT, as well as health-related quality of life (e.g. EQ-5D) in addition to process indicators such as changes in service use (e.g. emergency admission rates). Research being conducted by the Policy Research Unit on the Quality and Outcomes of Person-Centred Care (QORU), would be relevant (<http://www.qoru.ac.uk/>).

In addition, the PHB evaluation collected information on carer outcomes. The ASCOT carer quality of life measure could be used for this purpose.

The PHB study also indicated that the use of PHBs led budget-holders to change the services they used and possibly the providers commissioned to deliver those services. A current PHB evaluation

follow-on study is exploring the difficult issue of assessing 'market impact' of PHBs. The findings will be available to inform the IPC evaluation.

The cost implications of IPC will also be wide-ranging, including not just the cost of the individuals' budgets but also the indirect effects on, for example, secondary care utilisation. Moreover, it is important to measure comparable costs for the control group, including social care, community health care, primary care, and also the indirect effects on secondary care utilisation. The scope of the study will determine whether costs such as housing, informal care support, community resources etc., should be included in the analysis.

4.2 Complexity

IPC is a highly complex system intervention, more so even than PHBs, residential care DPs or PBs. It has a range of components (the budget, integrated funding, a financial model, a focus on commissioning (new) services from the community, etc.). It is also likely to have highly context-specific effects. IPC can thus be regarded as more a system change than a discrete intervention. As a consequence, any finding about the average impact (e.g. cost-effectiveness) of IPC (across the nine 'demonstrator sites') may not be representative of the impact of IPC if it were adopted in other areas. In other words, it will be important to explore why effects are occurring. Impact evaluations can contribute to this but such studies require large samples.

At present local IPC sites have a degree of autonomy over how their IPC programme is configured locally. Sites vary according to how budgets are to be determined and offered, client groups covered, governance arrangements, etc. though so called 'standard replicable models' of some of the components in the process are under development, led by NHS England which should reduce some of the variation between sites in how IPC is being implemented.

This complexity poses particular challenges for conventional experimental evaluation approaches such as RCTs (Bonell et al. 2012; Shiell, Hawe, and Gold 2008). In theory, multi-arm trials testing a range of IPC models could be used, but they would require large samples and would be high cost. Process and context-specific (case study) designs might be well suited in this case, but there would be limits on how far case studies could be used to generalise about the overall impacts of IPC. A combination, hybrid or mixed methods approach is the most likely solution, as was used for the studies described above. In particular, qualitative research would be highly relevant in developing 'implementation models' or configurations of IPC that could then be used to develop categories or types of IPC implementation for comparative quantitative analysis.

4.3 Scale, variation and implementation

Any sample-based estimation of IPC effects will be subject to statistical error. In regard to IPC, effect sizes can be expected to be highly variable, which means that large sample sizes would be needed.

Evaluations of impact would need a sufficiently large population of patients from which to draw a control group. It is important not to draw controls from people who refuse to be part of IPC, as this is, by definition, a 'selected' group. Staggered implementation is one option, with controls derived from a waiting list of people identified before they become eligible to join the IPC programme. Ethical issues might be raised by 'delaying' access for some people, but as there is uncertainty about whether the intervention is likely to be beneficial, this potential ethical issue should be avoided. In addition, sites may also be unable to cope with all those potentially eligible for IPC, thereby creating an implicit waiting list which could be used to assist the evaluation.

People in neighbouring (or similarly matched) areas that are not IPC demonstrators would be another option for creating a control group. Resources would be needed to engage those other areas and potential research participants.

The importance of having a counterfactual as a benchmark to compare the progress of people in the IPC programme was clearly demonstrated by the PHB evaluation. Before-and-after studies, although appealing from a practical standpoint, are likely to produce misleading results – and indeed to under-estimate the impact of IPC. People with frailty and long-term conditions are likely to suffer deteriorating health and wellbeing through time – the effect of IPC might be to slow this deterioration. If so, a before-and-after study would not detect this benefit. The dynamic nature of effects also strongly indicates a need for a follow-up study with outcome data collected at baseline and then again at (at least one) follow-up point.

In the nature of any innovative programme, there is inevitable uncertainty about how IPC models will develop, how many people will participate, how quickly implementation can proceed and so on. One response to uncertainty is to pursue an adaptive and flexible approach to the evaluation with room to scale up or down, alter methods and so on, over time. While this is challenging in practice, it may well be a sensible approach in relation to the evaluation of IPC given the limitations of current knowledge.

Unsurprisingly, for such complex interventions, the implementation of IBs, PHBs and DPs was slower than initially planned. An important question for the IPC evaluation is when to collect data - relatively early as the IPC model develops or after the policy has had time to become embedded within local working practices. Different things can be learned at different points in the evolution of the programme.

The implementation delays in the IBs, PHBs and DPs pilots, in turn, made the recruitment of participants slower than anticipated and led to much of the recruitment occurring towards the end of the planned evaluation period. These delays affected how quickly participants could be recruited and also the length of the follow-up period. Building in some flexibility and realistic follow-up periods is essential. In the PHB and IBSEN studies, this issue was managed successfully, with target sample numbers achieved, and the projects delivered on time. However, this required highly-motivated efforts by sites, the research team and support from policy leads. Foreknowledge and contingency planning for such delays would be highly beneficial.

In the DPs in residential care trailblazers' evaluation, data collection was severely hindered by the slow uptake of DPs by service users and/or the family members supporting them. At interview, conducted as part of the feasibility study, project leads had indicated their plans to recruit a total of 400-500 people with a DP by the end of the programme. However, these plans did not come to fruition, with sites deploying a total of 70 direct payments, of which 31 were operational (i.e. money had been transferred to the user), by the end of the programme.

In total, the evaluation team conducted over 120 interviews allowing an in-depth process evaluation and a thorough analysis of the many barriers to implementing direct payments in residential care. However, data collection for the outcome evaluation has remained limited, with a total of 59 baseline questionnaires collected from service users and families to date.

The planned number of sites in the IPC evaluation is lower than for IBs and PHBs, which will mean that the evaluation will need to be especially effective at maximising the potential level of recruitment of participants. Early indications from the IPC sites are that recruitment rates will have to increase to allow a sufficient sample for analysis of impact, depending on the purpose of the evaluation. Using PHB evaluation results as a benchmark, main overall effects (i.e. PHB group versus controls) with regard to measuring impact on care-related quality of life (ASCOT) required a smaller sample size than an analysis of the effect of PHBs in client sub-groups. Moreover, regarding cost-effectiveness analyses, costs were highly variable in the PHB pilot (and difficult to measure), which again put pressure on sample sizes. Finding relevant data on means and standard deviations for power calculation is often difficult in policy evaluations, but in this case, PHB evaluation and IBSEN would be useful and relevant sources.

This position highlights the need to ensure and sustain a substantial degree of commitment and support for the evaluation within the 'demonstrator sites' (both to help recruitment and to supply data to the evaluation). Such engagement will place demands on local sites. During the three evaluations discussed above, a significant demand was placed on local staff and professionals working in the pilot sites, both in implementing the scheme and helping the evaluations (e.g. including in particular, recruitment, data collection, follow-up of participants, etc.). In part, this is unavoidable because there are certain tasks that only local staff can do, given current research ethics constraints, such as introducing users to the evaluation and securing their consent, in principle, to participate. Researchers are not permitted to approach users and carers directly. Also local staff have opportunities to help with data collection (e.g. when delivering services) that are not open to researchers. This workload needs to be fully recognised by the IPC evaluation, ideally with resources committed to help and motivate local sites to achieve agreed commitments.

In the case of the IPC programme, there are resources potentially available though they are in the form of support from NHS England to help sites undertake their own local evaluations rather than directly to become part of an independent cross-programme evaluation.

4.4 Timing

Finally, in terms of study timing it is helpful to reflect on the experience of recent similar evaluations:

- Gaining research ethics, NHS R&D (CCGs, NHS trusts) and research governance (LA) approvals, required before any data can be collected: currently taking 6 months or more (especially given the time needed for LA research governance)
- Process evaluation: these can be generally achieved within 18 months, including initial IPC implementation period currently being studied (6 – 9 months). Lead times in setting up interviews with busy local professionals should not be under-estimated.
- Observational analysis of early impact indicators (with process evaluation and case studies): generally 24 months, with some overlap with the period of IPC early implementation (i.e. 3 months), although clearly contingent on data availability and access.
- Summative (experimental) study (e.g. RCT or NRS). With a baseline and follow-up time period of a year or more (e.g. as in the PHB evaluation) then three years is really required (e.g. with a 9 month or more recruitment period, and considering the initial implementation phase, but also acknowledging potential delays, and allowing for analysis, reporting and dissemination etc.). Some overlap would be possible with an initial implementation phase.

With a shorter follow-up period, and good recruitment rates, a shorter study period would be possible – for example, IBSEN was a two-year study with a 6-month follow-up period.

- Case studies of implementation and mechanisms of change: 18 months, again factoring in time to set up and conduct stakeholder interviews among busy staff.

Annex. Scoping study and early evaluation of the integrated personal commissioning (IPC) programme

The aims of the PIRU-ESHCRU study are to undertake an early, largely process-focused evaluation of the IPC programme between January and September 2016. The study began in December 2015 and focuses on the following:

- Implementation of the key mechanisms of change:
 - Progress with implementation of the key activities (or mechanisms) identified in the IPC programme logic model developed by RAND in its confidential draft report to NHS England (Sim et al, 2015) and the related description of IPC processes set out recently by NHS England (NHSE, 2015). For example, what processes have been put in place for single assessment, for a viable financial model, to promote market entry by new providers and to involve the community in service design and developing new community assets?;
 - The nature of the financial model; in particular, how sites calculate the size of individual integrated budgets, including how private contributions are taken into account and the scope of costs that can be met by the budgets offered;
 - The skills required and progress with appropriate training of staff.
- Challenges and adaptation:
 - The implementation challenges facing each pilot as they pursue the key activities identified by RAND and how they are being addressed;
 - Whether the differences in approach between the pilots identified by RAND (see above) persist and what implications, if any, they have for implementation of the programme;
 - Whether the overall logic model identified by RAND and its variants are plausible and realistic in light of experience over the first year of the pilots. This activity will include views about the development of the logic model. For example, there is likely to be an additional strand within the overall logic framework concerning community empowerment activities
- A stocktake of key outputs (where the data are available):
 - the numbers and types of people being offered personalised care plans, budgets, direct payments or budgets managed by a third party;
 - the services being purchased with budgets (e.g. in terms of the range and types of services, and the nature of the providers selected) and any evidence about unmet needs;
 - the effect of IPC on overall service use and costs per person and across the pilot programme;
 - whether new providers have been commissioned to provide services or existing providers replaced as a result of the use of IPC budgets, as well as the steps taken by the pilots to shape and develop the local provider market.
- Potential impacts:
 - which, if any, of the different approaches to IPC implementation appears to be more promising than others, and the identity of the change mechanism underlying each approach (e.g. what is the contribution of having a budget in the 'personalisation' of services and user choice versus care planning?);
 - insights into the added value of IPCs in commissioning better integrated services.
- Evaluation:
 - Implications for the longer term evaluation of the progress made by pilot sites by autumn 2016, and their ability to collect data routinely and validly on those aspects of IPC activity relevant to a robust evaluation without excessive burden, plus some insight into how to study the contracting and market shaping activities of the sites. Advice on the choice of

appropriate outcome measures to use for evaluation, building on the assessment of relevance, strengths and weaknesses of a range of outcome measures reported by RAND.

The main methods are to combine a review of the literature, documentary analysis and interviews with key stakeholders. Final ethical approval for the fieldwork was granted on 1 April 2016. The team is currently gaining the relevant R&D governance approvals.

References

- Bonell, C., A. Fletcher, M. Morton, T. Lorenc, and L. Moore. 2012. Realist randomised controlled trials: a new approach to evaluating complex public health interventions. *Social Science & Medicine* 75 (12):2299-306.
- Braun, V., and V. Clarke. 2014. What can "thematic analysis" offer health and wellbeing researchers? *International Journal of Qualitative Studies on Health and Well-Being* 9.
- Ettelt S, Mays N, Allen P. The multiple purposes of policy piloting and their consequences: three examples from national health and social care policy in England. *Journal of Social Policy* 2015; 44: 319-37.
- 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.
- Glendinning, Caroline, David Challis, José-Luis Fernández, Sally Jacobs, Karen Jones, Martin Knapp, Jill Manthorpe, Nicola Moran, Ann Netten, Martin Stevens, and Mark Wilberforce. 2008. Evaluation of the Individual Budgets Pilot Programme. University of York , London School of Economics, University of Kent.
- 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? *Journal of Health Services Research & Policy* 18:59-67.
- Marchal, Bruno, Sara van Belle, Josefien van Olmen, Tom Hoérée, and Guy Kegels. 2012. Is realist evaluation keeping its promise? A review of published empirical studies in the field of health systems research. *Evaluation* 18 (2):192-212.
- Moore, Graham F, Suzanne Audrey, Mary Barker, Lyndal Bond, Chris Bonell, Wendy Hardeman, Laurence Moore, Alicia O’Cathain, Tannaze Tinati, Daniel Wight, and Janis Baird. 2015. Process evaluation of complex interventions: Medical Research Council guidance. *BMJ* 350.
- NHS England. Integrated Personal Commissioning developing standard replicable models: planning and process. NHSE, November 2015.
- Pawson, R, and N Tilley. 1997. *Realistic Evaluation*. London: Sage.
- Sim M Saunders K, Winpenny E, Ling T. (2015) *Future evaluation of the Integrated Personal Commissioning programme: mapping the logic and assessing evaluability*. Confidential draft report to NHS England. Cambridge: RAND Europe.
- Shiell, Alan, Penelope Hawe, and Lisa Gold. 2008. Complex interventions or complex systems? Implications for health economic evaluation. *BMJ* 336 (7656):1281-1283.
- Wittenberg R, Ettelt S, Williams L, Damant J, Lombard D, Perkins M, Mays N. *Evaluation of direct payments in residential care trailblazers: second interim report*. PIRU Publication 2015-16. London: Policy Innovation Research Unit, London School of Hygiene and Tropical Medicine, December 2015 <http://piru.lshtm.ac.uk/assets/files/DP%20Trailblazer%20Second%20Interim%20Report.pdf>