From summative to developmental

Incorporating design-thinking into evaluations of complex interventions

Carolyn Steele Gray
Bridgepoint Collaboratory for Research and Innovation,
Lunenfeld-Tanenbaum Research Institute, Toronto, Canada, and

James Shaw
Institute for Health System Solutions and Virtual Care,
Women's College Research Institute,
Toronto, Canada

Abstract

Purpose – Models of integrated care are prime examples of complex interventions, incorporating multiple interacting components that work through varying mechanisms to impact numerous outcomes. The purpose of this paper is to explore summative, process and developmental approaches to evaluating complex interventions to determine how to best test this mess.

Design/methodology/approach – This viewpoint draws on the evaluation and complex intervention literatures to describe the advantages and disadvantages of different methods. The evaluation of the electronic patient reported outcomes (ePRO) mobile application and portal system is presented as an example of how to evaluate complex interventions with critical lessons learned from this ongoing study.

Findings – Although favored in the literature, summative and process evaluations rest on two problematic assumptions: it is possible to clearly identify stable mechanisms of action; and intervention fidelity can be maximized in order to control for contextual influences. Complex interventions continually adapt to local contexts, making stability and fidelity unlikely. Developmental evaluation, which is more conceptually aligned with service-design thinking, moves beyond these assumptions, emphasizing supportive adaptation to ensure meaningful adoption.

Research limitations/implications – Blended approaches that incorporate service-design thinking and rely more heavily on developmental strategies are essential for complex interventions. To maximize the benefit of this approach, three guiding principles are suggested: stress pragmatism over stringency; adopt an implementation lens; and use multi-disciplinary teams to run studies.

Originality/value – This viewpoint offers novel thinking on the debate around appropriate evaluation methodologies to be applied to complex interventions like models of integrated care.

Keywords Evaluation, Complex interventions

Paper type Viewpoint

Introduction

Integrated care is a prime example of a complex intervention. These models seek to integrate services within and across organizational and professional boundaries, through the use of governance, system structures and innovative technology. Not only are these models complex but often the clients they serve, such as older adults and those with multi-morbidity and social vulnerability, are complex as well. Complex interventions like integrated care include multiple interacting components which impact on variable outcomes through multiple mechanisms of

change (Craig *et al.*, 2008). With all these moving and interacting parts, there is a growing debate on how to best evaluate these interventions to assess value and impact.

This commentary explores summative, process and developmental approaches to evaluating complex interventions. Arguments offered draw on the literature, and the experiences of the co-authors who have studied implementation of integrated care, and evaluated multiple digital technologies to discuss the values and drawbacks of these approaches. Digital interventions are a useful comparator to integrated care models, as they are similarly complex interventions that require ongoing modification in their implementation, and often include an integration or coordination goal (Kvedar and Jethwani, 2014).

Summative and process approaches: the dominance of Medical Research Council Guidelines

A 2013 review of the complex intervention evaluation literature identified that most of these evaluation designs are based on Medical Research Council (MRC) 2000 and 2008 Guidelines on evaluating complex interventions (Datta and Petticrew, 2013). MRC guidelines advocate for a phased approach which allows for exploring and pre-testing components of interventions to help unpack the complexity, and support the design of stronger randomized clinical trials (RCTs) (Campbell *et al.*, 2000, 2007; Craig *et al.*, 2008). Given RCTs persist as the gold standard within the hierarchy of evidence (Akobeng, 2005; Rychetnik *et al.*, 2004), it is understandable why these evaluation approaches would be favored. These methods rest on summative evaluation principles aimed at evaluating the outcomes of an intervention after it has been adopted.

Phased approaches, depicted in Figure 1 (Campbell *et al.*, 2000), recognize that complex interventions require attention to influences at the context, process and outcome levels. This iterative approach allows each evaluation phase to build the next, marching toward the definitive clinical trial.

Preliminary phases, such as the observational and exploratory phases, are intended to assess feasibility, better define control and intervention arms of the explanatory trial, identify appropriate outcomes and test logistics of the study itself (e.g. power calculations and recruitment process). These stages are additionally intended to capture the mechanisms
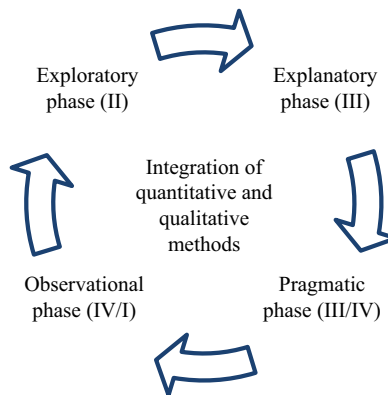

Figure 1.
Iterative evaluation approach

Source: Adapted from Campbell *et al.*'s (2000) figure "Iterative view of development of randomized controlled trials of complex interventions"

of action underlying interventions to clearly identify what about the intervention is leading to observed outcomes (Campbell *et al.*, 2000, 2007). Arguably, this assumes that mechanisms can be clearly articulated and quantified in these early phases, so they can be incorporated into models in later testing.

Although early phases leave room for the development of the complex intervention, there is an assumption that at some point the intervention becomes stable enough to be evaluated through trial methods. This notion of stability may be problematic as complex interventions often adapt and change throughout their implementation, leading some to question the applicability of RCTs for studying complex interventions (Hawe, 2004; Greenhalgh and Swinglehurst, 2011). The MRC attempted to fill this gap through developing guidance for process evaluations in 2015 which help to measure to what degree an intervention was delivered as intended (Moore *et al.*, 2015); this is known as program fidelity. When approached from certain perspectives, process evaluations can also help capture mechanisms and help us understand why interventions work in some contexts and do not in others (May *et al.*, 2007).

While incorporating process evaluations alongside trials may be useful to capture the many interacting components driving complex interventions, the underlying assumption of these approaches that we can clearly identify stable mechanisms of action; and get to a point, where we maximize fidelity in order to control for contextual influences, is problematic. Complex interventions often need to be continually adapted to local contexts and changing environments, and rely on mechanisms that are multiple and may shift over the life of the intervention (Plesk and Greenhalgh, 2001; Kirsh *et al.*, 2008).

Developmental approach: incorporating implementation and design thinking in each step

Developmental approaches are grounded in complexity theory, using evaluation findings, collected from ideation to implementation of a new model, to inform modifications and improvements to the intervention to address needs of a changing environment (Patton, 2008, 2011). The developmental approach follows a distinct logic from that of the MRC approach described above, in which the focus is on using insights generated by evaluation activities to promote the success of the program, as opposed to simply intending to understand whether and why the program worked. Developmental evaluation is based on the premise that actively achieving the outcome is more important than the perceived epistemic certainty arising from conventional evaluation approaches (e.g. RCTs), and that specific outcome measures can be developed and adapted throughout the intervention process (Patton, 2006, 2010). In this way, programs obtain real-time data on outcomes that may arise from naturally occurring data sets, offering more local insights that can inform local development and design activities.

Measures of success in developmental evaluations are very different than those for summative evaluations. Summative approaches seek to measure success based on pre-determined goals with an aim of generalizability to other contexts. With regard to developmental approaches, on the other hand, Patton (2006) suggests that:

> None of these traditional criteria [of attaining pre-determined goals, replicability, and generalizability] are appropriate or even meaningful for highly volatile environments, systems-change-oriented interventions and emergent social innovations (p. 31).

Patton goes on to elaborate that in ever changing environments we often cannot expect to conduct summative evaluations because interventions "can't hold still long enough for summative review" (p. 31). This does not preclude our collection of outcome measures to determine value; however, in complex interventions outcomes and the means of achieving them are often emergent and may be difficult to predict (Rogers, 2008). Instead of focusing on pre-determined outcomes, as is the case in summative evaluations, the emphasis moves toward building and rebuilding program theories, or logic models, which demonstrate the

connection between program activities and expected short- and long-term outcomes (Rogers, 2008).

Development evaluation resonates strongly with design-oriented approaches to implementation and evaluation. Design is fundamentally pragmatic, oriented toward making changes in a service process or physical artifacts that engage with peoples' experience in novel ways (Steen, 2013). Service design is holistic, human-centered, iterative and acknowledges the influence of the ecosystems in which services are nested (Saco and Gonclaves, 2010). Service design explicitly emphasizes the importance of trying new things out in small doses, gaining quick feedback and determining whether adaptations grow or detract from the energy of a new or altered program. Although there are a variety of service-design methods, they tend to be laser focused on the value that new services offer to users and the ways in which services can be configured to be easiest to use and to maximize benefit for all involved (Osterwalder *et al.*, 2014). The aim is for a service to be valuable and constantly updated to achieve intended goals, as opposed to assuming program stability and evaluating effectiveness *post hoc*.

A developmental evaluation approach, grounded in service-design thinking, may offer the best approach to evaluate complex interventions that need to adapt to local contexts over their life-course. Even in cases where a similar model has been followed, the specific individuals, relationships and practices vary simply as a result of being situated in different service ecosystems. These approaches have been used to evaluate complex health promotion interventions (Jolley, 2014), which like integrated care models are unique and require local adaptation.

The electronic patient reported outcomes (ePRO) example

An exploratory trial was conducted as part of the developmental approach adopted to evaluate of the ePRO tool, a mobile technology and web-based platform designed to support goal-oriented care delivery by inter-professional teams (a digitally enabled clinical integration model). The MRC recommended phased approach was used in combination with design thinking to inform the overall developmental approach to drive technology design and inform adoption process (Steele Gray *et al.*, 2016). The exploratory trial of the ePRO tool collected context, process and outcome measures over four months at two primary care practice sites (in press). Two findings from that study are particularly helpful in our understanding of the pros and cons of the summative and process approach vs a developmental one.

First, exploratory trial findings revealed dozens of context, process and outcome level variables that could play a role in the intervention; many of which are complex and not conducive to quantification to be built into models. For example, ePRO use was driven by individuals' complex motivations and nuanced personal relationships. Additionally, multiple mechanisms of action were identified, many of which changed over time. For example, provider motivation to engage with the technology (mechanism) was activated by different contexts at different stages of the study (e.g. an interest in reducing workload at the start vs patient-identified value at the end). As such, being able to ensure stability of the driving process of the intervention as critical to trials (Hawe, 2004) proved elusive.

Second, even though a user-centered development approach was used, the ePRO exploratory trial revealed that the tool and the implementation process had to be adapted to local contexts and participants to meet increasingly complex environments. The tool needed to become even more flexible, the user flow needed to adjust and we needed to allow primary care providers to identify their own processes for working between each other and with patients using the tool. A service-design lens helped to focus our developmental evaluation approach, capturing and analyzing data that could help us adapt the tool and process to better meet local needs. This moves beyond adaptation of simple components of the intervention, argued to be modifiable even for RCTs (Hawe, 2004), toward a need to alter the entire process.

Incorporating design thinking into evaluation methods
Lessons learned from designing and evaluating ePRO and evaluating other complex technologies, as well as our experiences studying the implementation of international models of integrated care (Breton *et al.*, 2017; Steele Gray, Wodchis, Baker, Carswell, Kenealy, McKillop, Breton, Parsons and Sheridan, 2018; Steele Gray, Barnsley, Gagnon, Belzile, Kenealy, Shaw, Sheridan, Wankah and Wodchis, 2018; Evans *et al.*, 2017; McKillop *et al.*, 2017; Kuluski *et al.*, 2017; Shaw *et al.*, 2017) suggest that the nuance and intricacy of leadership, personal relationships and different communication strategies is critical. These intangible qualities have been argued to be cornerstones of integrated care (Glasby, 2016). While there is still an important place for RCTs in evaluating simple interventions, they are not well suited to capture these nuances which are not easily quantified. Furthermore, the infinite number of complex permutations possible when implementing a model of integrated care makes any intention to make one model the same as a model elsewhere (needed to ensure fidelity in typical RCT designs) a fool's errand.

A service-design-oriented approach may also be more appropriate to support evidence-informed policy making. RCTs have been argued to be insufficient to be able to capture the complexity of the policy making process due to the inability to establish clear cause and effect pathways (challenges in identifying mechanisms related to the policy), and a lack of clarity of the appropriate timing of the RCT in relation to the implementation of a new policy (Ettelt and Mays, 2015). Additionally, the pressure to provide supportive evidence of previously adopted policies can result in questionable rigor of RCT reporting (Ettelt *et al.*, 2015). Adopting a service-design informed developmental approach can address these challenges, by providing quick feedback regarding the implementation of new policies, allowing for evidence to be embedded as part of the implementation process itself. Evaluators are no longer pressured to provide supportive evidence of policies that have already been adopted, and policy makers can benefit from actual evidence-informed policy at every step of the policy process.

Policy makers often look to evaluations to determine overall value or success of a program. Although summative approaches are often used as they provide policy makers with outcome measures linked to particular programs, the evaluation literature has consistently shown that outcome data alone are insufficient to provide decision makers with the information necessary to understand a programs value and why it did or did not work (Pawson *et al.*, 2005). While Patton suggests that we need to focus primarily on emergent goals, we argue that there is a role for both pre-determined program goals as well as emergent ones. The critical factor here is ensuring program, survival decisions are not based solely on pre-determined program goals interpreted in the absence of process, context and implementation information collected through developmental approaches. The additional data can further help policy makers seeking to spread complex interventions that have demonstrated value, as spreading requires attention to local contexts which influences program success (Lanham *et al.*, 2013).

As noted previously, service-informed developmental approaches will yield insights in the process of the program, building program theories and logic models that demonstrate how activities of stakeholders are linked to anticipated outcomes. Using program theory, rather than outcomes alone, can support generalizability and transferability of evaluation findings to other settings. Rogers (2008) provides an example of an agricultural development program in Africa being implemented in multiple communities, where effectiveness was greatly affected by local conditions. To address this gap, a program theory or logic model was developed which was sufficiently generic to apply to multiple contexts while capturing the key mechanisms of change needed to drive success. Similarly, in another example offered by Patton (2006), an evaluation of a complex housing intervention determined program success by focusing on higher level goals such as working

toward the program vision, observing positive change toward that vision and ability to adapt to local contexts that change over time.

Simply put, we cannot disentangle the outcomes of a complex intervention from the implementation process. As such, when it comes to evaluating complex models of integrated care, we are convinced that the best approach is to invest in the service-design process to augment and build upon the novel strengths of a program being implemented in a new place. Even though this might result in changes to the program logic, or to the perceived infidelity of the intervention itself, it is the only way to come to terms with the practical complexities of integrated care. Hence, blended approaches that rely more heavily on formative and developmental strategies of evaluation that resonate with service-design thinking are essential. While a short commentary cannot offer an in-depth set of guidelines, we do offer our top 3 guiding principles for adopting a blended approach:

(1) Commit to prioritizing the development over the evaluation; when faced with tensions between sticking to strict methods restrictions that may arise even in pragmatic trials, err on the side of practicality and feasibility. If the intervention cannot run, there will be nothing to evaluate.

(2) Adopt an implementation science lens (better yet an appropriate theory) to guide elements that will need to be paid attention to, and modified, as the intervention moves forward. This not only will help modify and adapt interventions, but also provides an analytic framework that can guide ongoing analysis and interpretation of findings from different data sources.

(3) Build a multi-disciplinary team that can offer expertise in the different methodologies used (e.g. trialists and analysts, ethnographers and social scientists, service designers), and that includes primary users of the intervention (e.g. patients, caregivers, providers, managers), who will guide each step of the evaluation from design to final reporting.

References

Akobeng, A.K. (2005), "Principles of evidence based medicine", *Archives of Disease in Childhood*, Vol. 90, pp. 837-840.

Breton, M., Steele Gray, C., Sheridan, N., Shaw, J., Parsons, J., Wankah, P., Kenealy, T., Baker, G.R., Belzile, L., Courturier, Y., Denis, J.-L. and Wodchis, W.P. (2017), "Implementing community based primary healthcare for older adults with complex needs in Quebec, Ontario and New-Zealand: describing nine cases", *International Journal for Integrated Care*, Vol. 17 No. 2, pp. 1-14, 12, doi: https://doi.org/10.5334/ijic.2506

Campbell, M., Fitzpatrick, R., Haines, A., Kinmoth, A.L., Sandercock, P., Spielgelhalter, D. and Tyrer, P. (2000), "Framework for design and evaluation of complex interventions to improve health", *British Medical Journal*, Vol. 321 No. 7262, pp. 694-696.

Campbell, N.C., Murray, E., Darbyshire, J., Emery, J., Farmer, A., Griffiths, F., Guthrie, B., Lester, H.E., Wison, P. and Kinmoth, A.L. (2007), "Designing and evaluating complex interventions to improve health care", *British Medical Journal*, Vol. 334 No. 7591, pp. 455-458.

Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I. and Petticrew, M. (2008), "Developing and evaluating complex interventions: the new Medical Research Council Guidance", *British Medical Journal*, Vol. 337, p. 19, available at: www.ncbi.nlm.nih.gov/pmc/articles/PMC2769032/

Datta, J. and Petticrew, M. (2013), "Challenges to evaluating complex interventions: a content analysis of published papers", *BMC Public Health*, Vol. 13, p. 18, available at: www.ncbi.nlm.nih.gov/pmc/articles/PMC3699389/

Ettelt, S. and Mays, N. (2015), "RCTs: how compatible are they with policy-making?", *British Journal of Healthcare Management*, Vol. 21 No. 8, pp. 379-382.

Ettelt, S., Mays, N. and Allen, P. (2015), "Policy experiments: investigating effectiveness or confirming direction?", *Evaluation*, Vol. 21 No. 3, pp. 292-307.

Evans, J.M., Grudniewcz, A., Steele Gray, C., Wodchi, W.P., Carwsell, P. and Baker, G.R. (2017), "Organizational context matters: a research toolkit for conducting standardized case studies of integrated care initiatives", *International Journal of Integrated Care*, Vol. 17 No. 2, pp. 1-10, 9, doi: https://doi.org/10.5334/ijic.2502

Glasby, J. (2016), "If Integration is the answer, what was the question? What next for English health and social care", *International Journal of Integrated Care*, Vol. 16 No. 4, pp. 1-3, 11.

Greenhalgh, T. and Swinglehurst, D. (2011), "Studying technology use as social practice: the untapped potential of ethnography", *BMC Medicine*, Vol. 9 No. 45, p. 7, available at: https://bmcmedicine.biomedcentral.com/track/pdf/10.1186/1741-7015-9-45

Hawe, P. (2004), "Complex interventions: how 'out of control' can a randomised controlled trial be?", *British Medical Journal*, Vol. 328 No. 7455, pp. 1561-1563.

Jolley, G. (2014), "Evaluating complex community-based health promotion: addressing the challenges", *Evaluation and Program Planning*, Vol. 45, Augsut, pp. 71-81.

Kirsh, S.R., Lawrence, R.H. and Aron, D.C. (2008), "Tailoring an intervention to the context and system redesign related to the intervention: a case study of implementing shared medical appointments for diabetes", *Implement Science*, Vol. 3 No. 1, p. 15.

Kuluski, K., Sheridan, N., Kenealy, T., Breton, M., McKillop, A., Shaw, J., Nie, J.X., Upshur, R.E., Baker, G.R. and Wodchis, W.P. (2017), " 'On the margins and not the mainstream:' case selection for the implementation of community based primary health care in Canada and New Zealand", *International Journal of Integrated Care*, Vol. 17 No. 2, pp. 1-4, 15.

Kvedar, J. and Jethwani, K. (2014), " 'Real-world' practical evaluation strategies: a review of telehealth evaluation", *JMIR Research Protocols*, Vol. 3 No. 4, p. 10.

Lanham, H.J., Leykum, L.K., Taylor, B.S., McCannon, J., Lindberg, C. and Lester, R.T. (2013), "How complexity science can inform scale-up and spread in health care: understanding the role of self-organization in variation across local contexts", *Social Science & Medicine*, Vol. 93, September, pp. 194-202.

McKillop, A.M., Shaw, J., Sheridan, N., Steele Gray, C., Carswell, P., Wodchis, W.P., Connolly, M., Denis, J.-L., Baker, G.R. and Kenealy, T. (2017), "Understanding the attributes of implementation frameworks to guide the implementation of a model of community-based integrated health care for older adults with complex chronic conditions: a metanarrative review", *International Journal of Integrated Care*, Vol. 17 No. 2, pp. 1-14, 10.

May, C., Finch, T., Mair, F.S., Ballini, L., Dowrick, C.F., Eccles, M., Gask, L., Macfarlane, A., Murray, E., Rapley, T., Rogers, A., Treweek, S., Wallace, P., Anderson, G., Burns, J. and Heaven, B. (2007), "Understanding the implementation of complex interventions in health care: the normalization process model", *BMC Health Services Research*, Vol. 7, p. 7, available at: https://bmchealthservres.biomedcentral.com/track/pdf/10.1186/1472-6963-7-148

Moore, G., Audrey, S., Barker, M., Bond, L., Bonell, C., Hardeman, W., Moore, L., O'Cathain, A., Tinati, T., Wight, D. and Baird, J. (2015), "Process evaluation of complex interventions: Medical Research Council Guidance", *British Medical Journal*, Vol. 350, March 19, p. 7.

Osterwalder, A., Pigneur, Y., Bernarda, G., Smith, A. and Papadakos, T. (2014), *Value-Proposition Design: How to Create Products and Services Customers Want*, Wiley, Hoboken, NJ.

Patton, M. (2008), *Utilization-Focused Evaluation*, Sage, Thousand Oaks, CA.

Patton, M. (2011), *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*, The Guilford Press, New York, NY.

Patton, M.Q. (2006), "Evaluation for the way we work", *The Nonprofit Quarterly*, Vol. 13, Spring, pp. 28-33.

Patton, M.Q. (2010), *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*, Guilford Press, New York, NY.

Pawson, R., Greenhalgh, T. and Harvey, G. (2005), "Realist review – a new method of systematic review designed for complex policy interventions", *Journal of Health Services Research & Policy*, Vol. 10 No. S1, pp. 21-34.

Plesk, P.E. and Greenhalgh, T. (2001), "Complexity science: the challenge of complexity in health care", *British Medical Journal*, Vol. 323, No. 323, pp. 625-628.

Rogers, P.J. (2008), "Using programme theory to evaluate complicated and complex aspects of interventions", *Evaluation*, Vol. 14 No. 1, pp. 29-48.

Rychetnik, L., Hawe, P., Waters, E., Barratt, A. and Frommer, M. (2004), "A glossary for evidence based public health", *Journal of Epidemiology and Community Health*, Vol. 58 No. 7, pp. 538-545.

Saco, R.M. and Gonclaves, A.P. (2010), "Service design: an appraisal", *Design Management Review*, Vol. 19 No. 1, pp. 10-19.

Shaw, J., Kontos, P., Martin, W. and Victor, C. (2017), "The institutional logic of integrated care: an ethnography of patient transitions", *Journal of Health Organization and Management*, Vol. 31 No. 1, pp. 82-95.

Steele Gray, C., Barnsley, J., Gagnon, D., Belzile, L., Kenealy, T., Shaw, J., Sheridan, N., Wankah, P. and Wodchis, W.P. (2018), "Using information communication technology in models of integrated community-based primary health care: learning from the iCOACH case studies", *Implement Science*, Vol. 13 No. 1, p. 14.

Steele Gray, C., Wodchis, W.P., Baker, G.R., Carswell, P., Kenealy, T., McKillop, A.M., Breton, M., Parsons, J. and Sheridan, N. (2018), "Mapping for conceptual clarity: exploring implementation of integrated community-based primary health care from a whole systems perspective", *International Journal of Integrated Care*, Vol. 18, No. 1, pp. 1-12, 14.

Steele Gray, C., Wodchis, W.P., Upshur, R., Cott, C., McKinstry, B., Mercer, S., Palen, T., Ramsay, T., Thavorn, K. and QoC Health (2016), "Supporting goal-oriented primary health care for seniors with complex care needs using mobile technology: evaluation and implementation of the health system performance research network, bridgepoint electronic patient reported outcome tool", *JMIR Research Protocols*, Vol. 5 No. 2, p. 17.

Steen, M. (2013), "Co-design as a process of joint inquiry and imagination", *Design Issues*, Vol. 29 No. 2, pp. 16-28.

Corresponding author

Carolyn Steele Gray can be contacted at: carolyn.steelegray@sinaihealthsystem.ca