Realism and Rhetoric in the Evaluation of a New Care Model

New Care Models and Evaluation

Between October 2016 and October 2018, Health and Social Care Evaluations (HASCE), based at the University of Cumbria, were commissioned to evaluate the Morecambe Bay Primary and Acute Care System (PACS) vanguard. This was one of 50 vanguard sites of the New Care Models (NCM) programme for integrated care in England, and took place within an existing local improvement strategy in Morecambe Bay, Better Care Together (BCT). NCMs were proposed as one of the three main pillars of the NHS *Five Year Forward View* (2014), the wide-ranging strategy plan addressing the challenges of changes in patient health needs and treatment preferences, in treatment technologies, and in health services and funding structures. NCMs were developed in particular to act as a blueprint for working beyond traditional divisions between primary, secondary, mental health and social services, which the *Forward* strategy cited as often precluding the provision of personalised and coordinated care. The PACS care model used population-based analyses in order to design services tailored to individual needs, and identify groups that need specialised care; addressing these through redesigned patient care pathways, collaborative networks with community assets, and redefined workforce roles where appropriate to improve resource deployment and respond to the needs of local communities. HASCE were commissioned to deliver an independent qualitative evaluation of the care model, synthesising their findings with quarterly quantitative reports delivered by the University Hospitals Morecambe Bay Trust Business Intelligence Unit. The evaluation answered questions specified by the national New Care Models Team, around the context of delivery, the changes introduced, the resources used and the intended and unintended outcomes and impacts of the NCM for staff,

stakeholders and patients across Morecambe Bay. The evaluation was delivered to Bay Health and Care Partners, the strategic deliverers of the NCM, with regular interim reports to inform ongoing delivery of the vanguard.[1]

Evaluating NCMs raised a range of distinctive challenges, both practically and methodologically. In this paper, I want to reflect on the character of three of these challenges for the Morecambe Bay evaluation: the challenges posed by the complexity of the NCM and its implementation; the challenge of discerning 'top down' strategic approaches to integrated care from 'bottom up' tactical responses; and the challenge of the fundamentally rhetorical dimensions underpinning the NCM programme. Discussing the realist methodology which HASCE used to address these, I will also suggest that evaluation needs to consider how a new model of integrated care may differ from traditional programmes – and in particular the role of rhetoric within them – when capturing a realistic and nuanced account of new models of care.

The Challenge of NCMs (1): Complexity

The first challenge of NCMs to consider is the type of complexity they introduce for evaluators. The content and structure of the PACS vanguard moved beyond conventional models of health intervention, by attempting to deliver a shift in meaning and behaviour in social, cultural and infrastructural terms; working from interventions addressing contextual, population-based needs and drawing together existing assets to better link acute, primary and

[1] This paper is a reflective consideration of the methods of evaluation used. For the outcomes of the evaluation itself, see Grimwood, T., Bell, L., Palazsczuk, A., Robinson, C., Skyrme, S., Maden-Weinberger, U. (2017). *Local Evaluation of Morecambe Bay PACS Vanguard: 12 Month Report* and Grimwood, T., Bell, L., Skyrme, S., and Weinberger-Maden, U. (2018). *Local Evaluation of Morecambe Bay PACS Vanguard: Final Report*, commissioned by NHS England via Bay Health and Care Partners. A full copy of the evaluation reports are available on request from the author.

community care. In this sense, the main task for delivery was how to translate an ambitious schematic, conceived at a strategic and political level, into tangible changes that were meaningful and sustainable within localised practice. Correspondingly, the main task for evaluation was how to evidence these tangible changes across a wide and varied geographical footprint.[2] Not only was the programme of work transformative, requiring collaboration across a range of sectors and wider communities (between NHS organisations, local authorities, the third sector and other local partners, as well as patients and the public), it was also evolving, and subject to multiple changes and adjustments as the programme developed. Delivery of BCT, and the vanguard funding within it, was organised into four overarching workstreams,[3] with some elements of each workstream at an operational stage, but others developing at a slower pace; others still were dependent on distinct elements (such as relevant Information Governance structures) to fully realise their design. The commissioners requested that the evaluation did not assess each workstream separately, but rather identified cross-programme cause and effect; a request which befitted the ethos of an integrated model of care, but also allowed the vanguard itself the potential for change during the course of the programme (such as reorganisation of workstreams).

In this sense, the Morecambe Bay vanguard bore the hallmarks of Patricia Rogers' definition of a complex intervention (2008: 38-9): unlike 'complicated' interventions, which involve multiple strands (different sites, stakeholders, outcomes and so on), in complex interventions, the interactions between these strands results in recursive causality (where high level outcomes are produced by a number of interactions between lower-level outcomes), and

[2] Morecambe Bay covers a geographical area of 1,800 km^2 (twice that of the average Trust nationally), but with a population of only 365,000. While there is a higher than average aging populations across the footprint, there are contrasting rates of life expectancy, as well as employment levels and average income. See https://www.uhmb.nhs.uk/files/bct-publications/Better-Care-Together-Plan.pdf

[3] At the beginning of the project these workstreams were Planned Care, Out of Hospital (North Lancs), Out of Hospital (South Cumbria), and Women and Children. Details on the main elements of the programme can be found at http://www.bettercaretogether.co.uk/.

emergence (where the shape of a successful outcome emerges only during the implementation, or through negotiations between different organisations). In order to evaluate robustly the causes and effects of the vanguard activity, a methodology was needed which engaged constructively with the tension between the linear expectations of an NCM, and their likely non-linear outcomes, rather than seeking to reduce or obscure it. As such, applying evaluation methods which were, in Ray Pawson's words, 'missing […] any notion of emergence, of internal, adaptive, self-generated, or historical change,'[4] (2013: 49) were likely to fall short of capturing the causal and impactful features of the NCM.

To this end, the evaluators utilised a realist approach (see Pawson and Tilley 1997; Pawson 2013). This approach assumes that physical and social systems are ordered, yet infinitely complex: they constitute a reality which is interdependent with our interpretations of them. As such, no amount of observation or measurement will allow a complete understanding of their organisation, because 'actors and programmes are rooted in a stratified social reality, which results from an interplay between individuals and institutions, each with their own interest and objectives.' (Marchal et al. 2012: 195) Instead, causality needs to be understood as a form of contextualised decision-making; or, as Marchal et al. summarise:

> Actors and society have potential mechanisms of causation by their very nature. Change occurs when interventions, combined with the right contextual factors, release the generative mechanisms. (Marchal et al. 2012: 202)

Identifying these generative mechanisms establishes the well-known mantra of realist evaluation: what *mechanisms*, in which *contexts*, allow change to happen and produce clear *outcomes*. This conceptualising of causation allows evaluation to demonstrate why a

[4] Pawson's target in this quote is the Medical Research Council's guidance on evaluating complex interventions (https://mrc.ukri.org/documents/pdf/complex-interventions-guidance/). Quite rightly, Pawson notes that this guidance defines complexity only through frequencies; this would align itself more with a description of complicated interventions on Rogers' typology.

programme may produce positive outcomes in one context, and negative in another. It does this by identifying a programme theory, describing the way the intervention proposes to link certain contexts, certain mechanisms and certain outcomes. This allows hypotheses to be derived, which identify where such configuration/s will be found; and testing of these hypotheses to take place through data collection. Or: 'given a generative mechanism (M) and a conducive context (C) for its triggering, we can expect to see and measure specific observations (O) and events.' (Connelly 2007: 936; see Pawson and Tilley 1997: 84; Astbury and Leeuw 2010: 365) In doing so, the realist evaluator can identify that certain mechanisms are likely to be more effective than others in certain contexts. In other words: the evaluation aimed to show what works, for who, and why.

Contexts were here defined as elements that were 'external to the intervention, present or occurring even if the intervention does not lead to an outcome,' (Marchal et al. 2012: 207) but which may have an influence on the firing of a mechanism, leading to an outcome. Where appropriate, existing quantitative data was collated and configured with the qualitative themes: thus, in Morecambe Bay, the organising contexts were geography and demographics; skill supply; the availability of resources; previous interventions and organisational cultures. Mechanisms were defined as particular things which had the power to initiate an event which would not have otherwise taken place, constituted by 'a combination of resources offered by the social programme under study and stakeholders' reasoning in response' (Dalkin et al. 2015: 3). The theme of 'integration of services,' for example, was interpreted as a strategic mechanism for the vanguard, because it initiated several of the outcomes specified by the vanguard logic model. At the same time, this overarching mechanism could be broken down into smaller enabling or disabling mechanisms derived from the interaction between the resources of, and responses to, the production of the NCM in practice. In this way, resources and responses were seen as 'mutually constitutive of a mechanism.' (Dalkin et al. 2015: 4)

Outcomes, meanwhile, described events which were produced directly from the application of the mechanism to certain contexts. Outcomes, in this sense, were not simply visible proof that the programme 'worked', but were used to test whether the hypothesised connection between contexts and mechanisms were reliable. This meant that while outcomes from the Morecambe Bay vanguard included both quantitative and qualitative measures (e.g. improvements in the quality of care, increased awareness of and response to population needs), they also included 'preventers': clearly articulated blocks to *any* outcome becoming visible, such as the insufficiency of reporting metrics within the system, or disengagement at key points within the NCM programme by staff and stakeholders.

This approach was particularly appropriate because, despite their nomenclature, no New Care Model is ever really *new*. As with many large-scale transformation programmes (see the Health Foundation and King's Fund 2015: 6), the Morecambe Bay workstreams typically sat alongside pre-existing systems and pathways, contemporaneous national transformation initiatives, as well as the (often-overlooked) relics of previous unsuccessful interventions aimed at improving health outcomes in the local area. This produced both systematic strains, and historical and cultural contexts which affected the extent to which individuals engaged with the vanguard activities. These contextual factors turned out to be key for shaping how participants negotiated the idea of an NCM. In this sense, the complexity of delivery was as much an enabler of change as it was a disabler.

The Challenge of NCMs (2): Strategies and Tactics for Transforming Care

If the complexity of the vanguard suggested a realist approach to its evaluation, the design and delivery of the new model of care posed some more specific challenges for implementing this methodology once the evaluation project had begun. The most prominent

of these was to identify what the programme theory being evaluated *was* – which contexts, mechanisms and outcomes should be observed and tested – in order to ensure the most meaningful data was being collected and analysed; an issue which numerous realist evaluators have also found with health transformation programmes (for example, Rycroft-Malone et al. 2010; Greenhalgh et al. 2009; Barnes et al. 2003). As noted above, the vanguard was part of a longer-running transformation programme and, due to both the history of the strategy and its ambition to bring together numerous organisations across the locality, there were differing understandings and little consensus as to the nature and purpose of BCT. As a result, identifying 'what success looked like' was a continuous site of debate amongst participants.[5] Simultaneously, the pre-existing transformation activities meant that it was not always possible to specify what the vanguard funding had delivered, separate from existing interventions. Due in part to contextual disablers (such as the Information Governance processes which troubled the NCM programmes across England), workstreams often lacked specific outcome measures at local levels, and instead relied on high-level, hospital-based outcomes such as Emergency Department (ED) Attendance and Bed-days.

This raised a deeper problem regarding the prevalent *culture of sense-making* within the vanguard, which informed both the programme design and its existing reporting structures (as well as the quarterly quantitative reports to the National New Care Models Team). This might be termed the 'strategic approach' to transformation programmes: a small set of high-level quantitative outcome measures were analysed for improvement or otherwise, and then correlated with changes reported from individual localised interventions. The benefit of this approach for system design was that it allowed population-wide changes to be identified straightforwardly, by mapping general trends, savings and improvements across the

[5] While each vanguard site was required to have a logic model describing its model of care, it was of note that in Morecambe Bay, in interviews and focus groups with over 150 participants, only two made mention of the programme logic model, suggesting that the logic model did not have a prominent operational usage.

health economy. It had the added benefit of providing metrics that can be shared across pathways and workstreams to demonstrate success (for example, reducing hospital bed-days). It therefore lent itself to a level of analysis appropriate for strategic delivery groups, where the lens is focused on the investment made and outcomes reached, without losing itself in the potentially infinite variations of local delivery that may cloud the bigger picture. But the term 'strategic' also invokes Michel de Certeau's distinction between the 'strategies' and 'tactics' for negotiating the changing world. Strategies are technocratic: they are able to 'produce, tabulate, and impose these spaces, [and] when those operations take place.' (1984: 30) Tactics, meanwhile, are the ways in which stakeholders negotiate and transgress those same operations. Certeau argues that social science traditionally fails to see the tactical use of culture because 'what is counted is what is *used*, not the *ways of using*.' (1984: 35, my emphasis) While the use of larger metrics facilitated a quasi-experimental approach (allowing, for example, before/after comparisons, and the use of a synthetic population as a control group), they did not show the decision-making processes by individuals or groups which lead particular mechanisms to result in particular outcomes. In short, it did not show the *tactics* with which each individual engaged with the high-level schematic of the NCM at a local level.

This was particularly important in cases where the slow pace of infrastructural change meant that ground-level delivery had to traverse the strategic spaces allotted to them. The heavily contextualised act of 'having conversations' became an important enabler for change, often preceding any strategic agreement on what a particular intervention would be. But consequently, the qualitative data collected during the Morecambe Bay evaluation found participants voicing concerns and criticisms over the models of measurement being used to report their activities.

For example, the vanguard established 12 Integrated Care Communities (ICCs), which were given relative autonomy to develop according to local contexts. At the level of delivery some ICCs were consistently discussed as being more developed than others: they had brought together different organisations in dialogue, facilitated multi-disciplinary team working and receiving positive anecdotal feedback. Others were not as well-developed; but through the course of the evaluation period, higher-level outcome measures showed no significant differences between ICCs. In part, this was due to measures of success at ICC level often being reported anecdotally rather than systematically. Certeau (1984) is again useful for unpicking this: he suggests that strategizing involves a *spatializing* of knowledge, removing its temporally unfolding qualities and replacing them with categories of measurement. Thus, when qualitative data suggested that changes at local levels were rooted in aspects such as the character of individuals working on a project, the cultures they worked in, or the local history of similar interventions, these *temporal* variables – which are deeply embedded in particular moments of sense-making between specific audiences – are unable to be mapped on to a spatialised strategic approach.[6] The result of this tension was that participants often did not always trust the outcomes being produced, which introduced an unexpected negative feedback loop to emerge within the programme delivery: even when high-level outcomes showed improvement, this by itself did not necessarily convince participants that the *longer-term* cultural change underlying the principle of NCMs was taking place, which in some cases affected engagement with the programme.

[6] I prefer to use Certeau's terminology here rather than the standard discussion in evaluation theory around successionist models of causation (Maxwell 2012: 36-7), where two or more data sets are compared, and correlation inferred (for example: frequency of ED attendance pre-vanguard funding could be compared to the frequency post-funding). The tension at work here is not the underlying philosophical assumptions of causality, but the philosophical principles of transformative practice – that is, between the ethos of the NCM as place-based and population-focused, and the assumptions about the relationship between data and meaning within the vanguard design and delivery.

It should be emphasised that this is not simply restating the over-familiar process vs outcome, qualitative vs quantitative arguments within evaluation theory. After all, one can detect the same top-down 'strategic' approach at work in the King's Fund interim report on PACS vanguards (Naylor and Charles 2018), which utilises a purely qualitative approach to identify the enablers for system-wide change. The report is notable drawing together high level 'answers' at the expense of complexity and difficulty. Just one example is a specific conclusion it reaches:

> *Bold goals can be fuel for change*. In a stressed system, the tendency to lower sights is common. The vanguards evinced a very different psychological dynamic, seeking, without apparent trepidation, breakthroughs and unprecedented results. (Naylor and Charles 2018: 41, emphasis original)

While this is no doubt an encouraging, aspirational and policy-friendly point to make, the evidence provided to support this conclusion remains based on self-reported case studies from (often high-ranking) individuals. There is little critical appraisal, or synthesis of wider evidence, to support their reliability. In contrast, the more detailed examination of the application of the PACS logic model into practice within the Morecambe Bay vanguard evaluation suggested that positing ambitious goals had a *converse* effect: when goals were aspiration-heavy, they tended to inherently lack clear *operational* signs of success. 'Bold goals' such as 'improving patient self-care' might have fuelled change, but they also introduced a lack of clarity around what success looked like, as they circulated across the vanguard site; particularly when achieving such goals would always take longer to achieve than the lifespan of the non-recurrent vanguard funding. The result was that one consistent finding of the evaluation was, while participants could detail the reasons for the NCM being

delivered, they found it far more difficult to identify tangible evidence of that rationality in practice.

The Challenge of NCMs (3): Rhetoric and Realism

The problem at work in what I have termed the sense-making culture of the vanguard was not the method of data collection, but rather the gap between localised data and high-level data which the strategic approach caused to appear. This, in turn, reflects the wider tensions in the premise of the NCM, between locally-driven, population-based interventions, supported by wider changes around information governance and information technology, and the strategic leadership and management of these changes to ensure they are sustainable and replicable. This gap can be represented in a schematic representation of the strategic approach in Figure 1. The evaluators' task of bridging the micro and the macro levels of delivery; without which, there would be no constructive space for a dialogue around the frustrations and failures of the project, as much as its successes. Rather than begin from a pre-established programme theory, this involved using the first twelve months of the project to construct theories from the 'tactical' level of delivery, and develop an over-arching middle-range theory based on a 'level of abstraction that cannot capture every decision […] in the programme pathway but which is able to illuminate a common […] dynamic from which all interventions can learn.' (Pawson 2013: 156; see Pawson and Tilley 1997: 84; Pederson and Rieper 2008: 291)[7] This theoretical model of the NCM mapped these localised causes and effects on to the broader strategic aims of the vanguard (Figure 2), and provided the basis for the cross-programme theories which the commissioner had requested.

[7] See Grimwood, T., Bell, L., Palazsczuk, A., Robinson, C., Skyrme, S., Weinberger-Maden, U. (2017). *Local Evaluation of Morecambe Bay PACS Vanguard: 12 Month Report*. Bay Health and Care Partners, p.36

[FIGURE 1]

[FIGURE 2]

To do this, the evaluation team identified purposive sample routes through particular projects, based on scoping meetings with commissioners and work-stream leads. These routes were initiated by approaching named project and clinical leads, and from there following a chain-referral (or 'snowballing') path. Data was collected from interviews with staff and stakeholders, and focus groups held with patient groups, with interview schedules guided by the themes outlined in the BCT strategy documents. Data was coded via a Template Analysis. This began with the coding of data into basic, organising and global themes, which were then placed on a template of context, mechanism and outcome themes, allowing the causal interrelations between them to be identified. Findings were validated through a number of stakeholder workshops and dissemination events held during the course of the evaluation. From this, the evaluators produced a configuration of interlinking hypotheses around the relationship between contexts, enabling mechanisms, disabling mechanisms, preventers and outcomes across the vanguard. This analysis resulted in a C-M-O configuration of the NCM as a whole, which could then be used as the basis for testing and analysing on particular areas of the vanguard in the second year of the evaluation.

But one final challenge remains for discussion. By utilising the method suggested in Figure 2, the NCM risks appearing as a coherently structured system, with the evaluation simply joining together a number of component 'activity theories' (Cole 1999) which are fitted together in the middle level of the diagram (the 'structured pathway'). But this would be misleading. Firstly, this overlooks a key reason for basing the evaluation on generative causality. Realistic mechanisms are not simply those described in the linear logic models of a programme's design:

> When investigating mechanisms that might plausibly account for program outcomes, it is important not to be "mechanistic." There are no set procedures to be followed rigidly, no columns, rows, or logic model boxes to be filled with generic examples of program "inputs," "outputs," or "activities." […] [T]he evaluator is an applied theorist. (Astbury and Leeuw 2010: 374)

Secondly, these generative characteristics became particularly pertinent for the NCM, because the exploratory nature of its delivery – indeed, the entire principle of vanguard funding, pump priming locally-focused initiatives as a means to whole-scale system change – meant that a structured pathway could not be mapped into the system until it had already been running for some time, and new ways of working had borne fruit. However, high-level and ground-level data still interacted and responded to one another in *some* way, as demonstrated by the negative feedback loops which emerged when, for example, some participants felt that not enough time was provided for fruit to be borne.

The ambiguity of these interactions was not due to the measures used for either system management or evaluation, but was rather derived from the multiple ways of answering one question: 'what *is* a New Care Model, really?' Indeed, all of the challenges raised in this paper are rooted in the expectation that an NCM was, in itself, a 'thing', but with no consistent agreement as to what this 'thing' ultimately was: a policy decision, a set of population-focused activities, a cost-saving exercise, a pre-defined programme, or something else. NCMs were formulated as strategic and political responses to the future demands of health provision; the ethos of NCMs regarding integrated care, resonated with many existing transformation programmes (such as Better Care Together in Morecambe Bay), which enabled them to be realised in practice in some areas; conversely, this resonance could also introduce a disabling element if pre-existing transformation programmes were known by

brand, but not by tangible content. The 'strategic' approach framed the NCM as a system of inputs and outputs, which would result in changes to high-level outcomes: in this way, the NCM was assumed to be something *taking place*, and correlations then sought between investment and outcome. The realist evaluation approach suggested the NCM was a set of interactions between contextual factors and generative mechanisms which caused change to happen via individual decision-making (such as Integrated Care Communities yielding successes through improved communication and relationships).

The differences in definition can be ascribed to the fact that, in each case, they address particular audiences, and employ appropriate modes of reasoning to do so. As a result, the key to evaluating the Morecambe Bay vanguard was not to look for one, decisive definition of the NCM, but rather to engage with this ambiguity, and understand its role in the changes to care provision taking place. This involved understanding the NCM as a fundamentally *rhetorical* entity: in the sense that it emerged as a mobile and dynamic process of persuasion formed by its relationship to multiple audiences. It is here that the temporal aspects of the vanguard's success which the qualitative data collection found – that is, mechanisms which made sense in a particular time, to a particular audience – became fundamental to measuring the value of the NCM. Thus, vanguard funding depended upon addressing a political context; delivery addressed local contexts (and the different audiences within them, from public, private and third sectors); and so on. Correspondingly, the difficulties the vanguard faced with demonstrating success – the challenges of complexity, and the difficulties of managing strategic and tactical approaches to delivery – were often rooted in failing to address these differences between audiences (such as with the example of

ICC development). The changing perception of what an NCM 'was' at different levels of delivery actively shaped the interplay of evidence and delivery.[8]

At first, rhetoric's association with the slipperiness of words and their power to persuade may seem to sit uncomfortably with the scientific orientation of both realism, and evaluation more generally. But 'rhetorical' here refers more broadly to the persuasiveness by which an intervention of some kind brings about the 'adherence of minds' (Perelman and Olbrechts-Tyteca 2008: 8) between deliverers and stakeholders. Pawson's realist manifesto already stresses heavily the importance of an intervention as a persuasive entity, which he identifies as a form of context:

> Programmes seek to change minds. The likelihood of this happening depends not only on pre-existing mindsets […], but on the process whereby minds are changed. Programmes do not work through Pauline conversions and divine deliverance. On the whole they persuade people to reconsider their options and such cogitation tends to occur gradually. (Pawson 2013: 34)

The case of NCMs seem to suggest we widen this process of persuasion from a contextual feature to a more fundamental level. While evaluation has a particular tradition of rhetorical analysis at the level of speech – attention to the particular language, tropes and imagery used by participants, and so on – my suggestion here is that we understand NCMs as possessing a rhetorical character at more of an ontological level. In other words, they require

[8] This raises some interesting issues for Dalkin et al.'s distinction between resources and responses in realist evaluation. Dalkin et al. suggest that, in order to overcome the problem of discerning context from mechanism, the resources of a programme and the responses of actors within it are 'disaggregrated', with resource mechanisms identified as that which is introduced into pre-existing contexts, and responses deriving from these. The political dimensions of the NCM, and the extent to which the evaluation found its very identity to be rooted within the quality of its persuasiveness, tended to blur the operational effectiveness of such a differentiation, however. Identifying what resources were in play was, in this sense, often inseparable from stakeholder's responses.

extending this sense of persuasiveness beyond individual use, and to the systematic properties of the transformation process as well. In this sense, the mid-level 'structured pathways' of Figure 2 are separate from a logic model or system design which, in the case of NCMs, often remains schematic and, indeed, dependent upon tactical interpretation at the more local levels of delivery. Instead, they are produced by evidencing the ways in which particular elements of a programme come together 'so as to affect or transform' how particular working spaces are inhabited (Rickert 2013: 160), directed towards identifiable audiences. Just as the paper began by stating that the main task for delivery was how to translate an ambitious schematic into meaningful changes within localised practice, so the main task for evaluation is to assess the adaptability (Perelman and Olbrechts-Tyteca 2008: 508) of the NCM to its audiences, and the outcomes this adaptation produces. The evaluation question, in this sense, was not what worked, or even what worked for who, but rather *what made this new way of working persuasive to which audience*. Given the complexity of multi-disciplinary working within integrated care, this rhetorical dimension of new working practices is a fundamental, but often overlooked, part of successfully developing and implementing new models.

Conclusion

The reflections in this paper have suggested that the Morecambe Bay evaluation highlighted how translating ambitious strategic initiatives into tangible, localised outcomes demands nuanced and detailed approaches to evidence, as well as the representation of that evidence. In particular, the realist method of evaluation enabled a robust assessment of the complexity and non-linearity of the NCM's attempt to implement new models of integrated care. This non-linearity formed a crucial part of understanding both the success and shortcomings of the programme, and suggests a strong case for the use of realist methods in

future evaluations of integrated care initiatives. But more than this, the approach allowed the evaluators to raise questions around the *ontological* basis of the NCM – how we conceive of what it 'is' – and the corresponding requirements of evidence to support it, which remains a fundamental challenge to integrating health provision. In doing so, the fundamentally rhetorical dimensions of the programme became apparent. I have suggested, albeit briefly, that these principles of persuasiveness are not simply part of the communication of integrated care models, but rather form their very basis. Further work incorporating rhetorical theory into realist evaluation more prominently offers the potential for improving our insights into how practitioners and patients make sense of new models of integrated care, and in doing so bridge disparities between the strategic-level interests and ground-level delivery.

Acknowledgements

The vanguard evaluation was commissioned by Bay Health and Care Partners, on behalf of NHS England.

References

Astbury, B. and Leeuw, F. (2010). 'Unpacking Black Boxes: Mechanisms and Theory Building in Evaluation.' *American Journal of Evaluation*. 31:3, pp. 363-381

Barnes, M., E. Matka, and H. Sullivan. (2003). 'Evidence, Understanding and Complexity: Evaluation in Non-Linear Systems.' *Evaluation* 9:3, pp.265–84.

Connelly, J. (2007). 'Evaluating complex public health interventions: theory, methods and scope of realist enquiry.' *Journal of Evaluation in Clinical Practice*, 13, pp.935–941

Dalkin, S.M., Greenhalgh, J., Jones, D., Cunningha,, B. and Lhussier, M. (2015). 'What's in a mechanism? Development of a key concept in realist evaluation.' *Implementation Science* 10:49, pp.1-7

De Certeau, M. (1984). *The Practice of Everyday Life*. Trans. Rendall, S. Berkeley: University of California Press

Greenhalgh, T., Humphrey, C., Hughes, J., Macfarlane, F., Butler, C. And Pawson, R. (2009). 'How Do You Modernize a Health Service? A Realist Evaluation of Whole-Scale Transformation in London.' *The Milbank Quarterly*, 87:2, pp. 391–416

The Health Foundation and the King's Fund (2015). *Making change possible: a Transformation Fund for the NHS.* Available at https://www.kingsfund.org.uk/publications/making-change-possible

Marchal, B., van Belle, S., van Olmen, J., Hoerée, T. and Kegels, G. (2012). 'Is realist evaluation keeping its promise? A review of published empirical studies in the field of health systems research.' *Evaluation* 18:2, pp. 192-212

Maxwell, J. (2012). *A Realist Approach for Qualitative Research*. London: SAGE

Naylor, C. and Charles, A. (2018). *Developing new models of care in the PACS vanguards: A new national approach to large-scale change?* The Kings Fund, Available at https://www.kingsfund.org.uk/sites/default/files/2018-04/PACs-new-models-Kings-Fund-April-2018.pdf

Pawson, R. and Tilley, N. (1997). *Realistic Evaluation*. London: SAGE

Pawson, R. (2013). *The Science of Evaluation: A Realist Manifesto*. London: SAGE

Pederson, L. and Rieper, O. (2008). 'Is Realist Evaluation a Realistic Approach for Complex Reforms?' *Evaluation*.14:3: 271–293

Perelman, C., and Olbrechts-Tyteca, L. (2008 [1969]) *The New Rhetoric: A Treatise on Argumentation*. Translated by John Wilkinson and Purcell Weaver. Notre Dame, IN: University of Notre Dame Press

Rickert, T. (2013). *Ambient Rhetoric*. Pittsburgh: University of Pittsburgh Press

Rogers, P. (2008). 'Using Programme Theory to Evaluate Complicated and Complex Aspects of Interventions.' *Evaluation*, 14:1, pp.29-48

Rycroft-Malone, J., Fontenla, M., Bick, D. and Seers, K. (2010) A realistic evaluation: the case of protocol-based care. *Implementation Science* 5:1, 38.