



Open Research Online

Citation

Calo, Francesca; Roy, Michael; Donaldson, Cam; Teasdale, Simon and Baglioni, Simone (2021). Evidencing the contribution of social enterprise to health and social care: approaches and considerations. *Social Enterprise Journal*, 17(1) pp. 140–155.

URL

<https://oro.open.ac.uk/74083/>

License

(CC-BY-NC-ND 4.0) Creative Commons: Attribution-Noncommercial-No Derivative Works 4.0

<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Policy

This document has been downloaded from Open Research Online, The Open University's repository of research publications. This version is being made available in accordance with Open Research Online policies available from [Open Research Online \(ORO\) Policies](#)

Versions

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding

Exploring the contribution of social enterprise to health and social care: approaches and considerations

Authors:

Francesca Caló – Corresponding Author^a - Francesca.caló@open.ac.uk

Michael James Roy^a - Michael.roy@gcu.ac.uk

Cam Donaldson^a- Cam.Donaldson@gcu.ac.uk

Simon Teasdale^a - Simon.Teasdale@gcu.ac.uk

Simone Baglioni^a- Simone.Baglioni @unipr.it

^a Pulse, Faculty of Business and Law, the Open University,
Walton Hall, Milton Keynes, MK7 6AA,
United Kingdom

^b Yunus Centre for Social Business and Health, Glasgow Caledonian University

^c Dipartimento di Scienze Economiche e Aziendali, Università di Parma

Abstract

Purpose

As the provision of public services in many advanced welfare states has increasingly come to be marked by competition, social enterprises have actively been encouraged by governments to become involved in the delivery of public services. While the evaluation of complex public health interventions has arguably become increasingly more sophisticated, this has not been the

case where social enterprise is concerned: evaluation of the actual impacts of social enterprises remains significantly underdeveloped by comparison.

Methodology

We assess the potential of three methodological approaches common in the evaluation of complex public health interventions and apply them to the complex realm of community-led social enterprise.

Findings

Only through the involvement of different comparator groups, based on the research questions addressed, would it be possible to disentangle the embedded characteristics of organisations such as social enterprises. Each of the methods adopted in this research is time-consuming and resource intensive and requires the researcher to possess advanced skills. Public officials should recognise the complexity and resource-intensive nature of such evaluation, and resource it accordingly

If the aim of policymakers is to understand the added value of social enterprise organisations, an integrative research approach combining different research methods and design should be implemented to improve generalisability.

Originality

We apply a range of favoured approaches to evaluate ‘complex’ public health interventions include systematic reviews, realist evaluation, and quasi-experimental investigation. However, such evaluation approaches have rarely been applied before in the context of social enterprise.

Keywords: social enterprise, evaluation, health and social care, evidence-based policy

Introduction

Over the last few decades, the ethos of markets and competition has increasingly been introduced into health and social care provision in many advanced welfare states, with the UK at the forefront of this development internationally (Hall et al., 2012). Social enterprises – broadly speaking, businesses that serve a social and/or environmental mission (Borzaga and Defourny, 2001; Defourny and Nyssens, 2006; Kerlin, 2006) – have been tasked with competing for, and delivering, health and social care contracts on behalf of the state (Alcock et al., 2012; Hall et al., 2012) on the perception that they provide higher levels of innovation, cost-effectiveness and responsiveness (Bovaird, 2014). Despite this rhetoric, the evidence that provision by social enterprise is ‘better’ than available alternatives is notoriously weak (Calò et al., 2018; Vo et al., 2016). This is particularly ironic given the supposed centrality of ‘evidence-based policy’ to health and social care (Kemmer, 2006; Smith, 2013). While a plethora of robust methodological frameworks has emerged to evaluate interventions in health and social care, with a view to informing policymakers on their effectiveness (Byford et al., 2010; Kah and Akenroye, 2020), such methods have rarely been tested in the realm of social enterprise (Lyon et al., 2015) and there have been calls for more peer-reviewed empirical studies utilising a broader range of study designs (Hervieux and Voltan, 2019; Milley et al., 2018).

In this paper we assess the potential of three methodological approaches common in the evaluation of complex public health interventions, with a view to supporting the establishment of a robust evidence base for the use of social enterprise as a policy instrument. In summary, our research question was to explore: *how, and in what ways, can methods from public health be used to assess social enterprise?* We provide lessons drawn from three different approaches: a systematic review, a realist evaluation, and a quasi-experimental investigation. Our intention

in this paper is to focus attention on the lessons we learned while applying each of the methods *in situ*, and on the appropriateness of employing methods commonly used in public health into the complex environment in which community-led social enterprises operate.

Our paper is organised as follows: after providing the background to the current position on social enterprise evaluation, we provide the context in which our three studies were undertaken, before examining the challenges and limitations we faced. We conclude by considering what these limitations potentially mean for the development of the social enterprise evidence base, and, indeed, for evaluation of social policy generally.

Evidencing the impact of social enterprise

In the process of actively encouraging social enterprises and non-profit organisations to deliver public services (Buckingham, 2009) a veritable ‘impact evaluation industry’ has emerged (Arvidson and Lyon, 2014; Hall and Arvidson, 2014). A variety of impact evaluation instruments have been utilised in non-profit management, social entrepreneurship, and programme evaluation for some time (Bagnoli and Megali, 2011; Epstein and Klerman, 2012; Kah and Akenrove, 2020; Kroeger and Weber, 2014; Leeuw and Vaessen, 2009) and a specific instrument called Social Return on Investment (SROI) (Flockhart, 2005) was piloted and promoted for use in the evaluation of UK Department of Health-funded programmes (Alcock et al., 2012). SROI has, however, proved challenging for evaluation of social enterprise in a health context (Alcock et al., 2012), suffering from difficulties in attributing financial proxies to so-called ‘soft outcomes’ such as improved self-esteem, self-confidence, or in relation to addressing social isolation (Bagnoli et al., 2019), which are commonly found among the advantages that social enterprises are considered to bring (Elmes, 2019; Farmer et al., 2020, 2019; Kelly et al., 2019; Macaulay et al., 2017; Roy et al., 2014; Suchowerska et al., 2019).

SROI can also be expensive to self-administer (Arvidson et al., 2013; Millar and Hall, 2013) and does not establish causal relationships or attribute specific results to the activity of the organisations (Arvidson et al., 2013; Dey and Gibbon, 2017; Harlock, 2013). Although SROI can be a useful tool to help organisations to self-evaluate their activities, such as approach is not considered rigorous enough to inform policymakers or upon which to base policy decisions (Arvidson et al., 2013).

If the aim of evaluation is to inform better policy decisions, then more rigorous and appropriate approaches to understand the plurality of outcomes of social enterprise activities, and how they are generated in health and social care contexts, are therefore required (Hervieux and Voltan, 2019). This is especially the case if social enterprises are to potentially compete for resources, including against interventions in health and social care that are, or have been, evaluated using more ‘robust’ methods. A range of favoured approaches recommended for use by the UK’s Medical Research Council to evaluate ‘complex’ public health interventions (Craig et al., 2008) include systematic reviews, realist evaluation, and quasi-experimental investigation. However, such evaluation approaches have rarely been applied in the context of social enterprise (Calò et al., 2018; Lyon et al., 2015). Before turning to our analysis of these three approaches, a brief overview is provided of the context in which the research has been conducted.

Research context

The focus of our evaluation approaches is on a social enterprise that we call ‘Active Life’, which operates in a rural community in the west of Scotland. The organisation was purposively selected as a case study due to the combined presence of specific health challenges (i.e. the rate of chronic conditions in the area), the community-centred approach (South, 2015) based upon a social enterprise model, and the opportunity it presented to explore a previously unexplored

and unusual revelatory setting (Eisenhardt and Graebner, 2007; Yin, 1994). The main aim of Active Life is to increase the physical activity levels of people with chronic health conditions. The organisation was created through a partnership between the local medical centre, the physiotherapy and dietetics department of the local hospital, and another already well-established community-based social enterprise operating in the area. The community is well known for being particularly active in funding non-profit organisations and in fostering community-led solutions to local needs. Although remote and rural community based-health care services in Scotland have long been regarded as “bastion of quality service provision” (Farmer et al., 2010, p.275), local health services in this particular area have had to cope with a proportionately high number of people dealing with at least one chronic condition (as evidenced by local health care statistics, which are drawn upon in, for example, the organisation’s annual reports). Active Life has been developed as an additional service to those offered by health care providers, reinforcing the links between the local provision of NHS services and the community. Specifically, the services are run as a GP referral scheme, providing classes and one-to-one services which aim to increase the levels of physical activity among local people with chronic conditions. Starting with a consultation with a fitness instructor, a 12-week programme tailored towards individual needs is created.

With regard to the definition of a social enterprise, Active Life is an organisation aiming to maintain economic sustainability through trading. Active Life is registered as a company limited by guarantee, with charitable status. Funding for the first two (pilot) years was derived from charitable foundations and public grants augmented by a proportion of revenue derived from services paid for by groups of beneficiaries who could afford to pay, albeit recipients of certain state benefits accessed the services free of charge. In terms of financial sustainability,

at the time of research the organisation was looking to create stable funding agreements based upon contracts with the NHS.

Research Methods

Our research employed three different approaches widely used in evaluating interventions in health care: systematic review, realist evaluation and quasi-experimental investigation. Systematic reviews collate and assemble all the up-to-date empirical evidence to answer a specific research question (Shemilt et al., 2010). They have been considered a robust form of research because, if undertaken correctly, they are believed to increase the generalisability of results and assess consistency of evidence (Mulrow, 1994). As a result, they are widely used as a tool to inform policymakers, practitioners and civil society (Chalmers et al., 2002). In health care research, a systematic review is among the methods used in comparing competing ways to treat beneficiaries, and is a useful vehicle for exploring the contribution of an intervention in terms of its effectiveness and relative costs (Donaldson et al., 2002). The systematic review conducted in our research was performed to compare and analyse how provision by social enterprise differed to what we termed ‘usual care’ in health and social care settings; that is, delivery by the public sector. Thirteen databases were searched alongside specific journals related to social enterprise and non-profits. A total of 24,717 papers were retrieved, of which 771 were screened in full text. After excluding many of these papers, as they were not analysing social enterprise, nor assessing any type of health outcomes and/or not comparing an intervention with usual care, what was left was only 25 studies in total, of which 12 were quantitative, eight were qualitative and five mixed-method studies. Ten of the studies were conducted on social enterprise in the UK, four in Australia, four in Canada, three in the US, and one each in Bangladesh, China, Norway and Sweden. The interventions in question

were very heterogeneous, with a wide range of beneficiary populations, from physical activities for elderly people, to parenting programmes for at risk children. Table 1 includes details of the study design, data collection and data analysis used in the systematic review.

-----Insert Table 1 around here-----

Realist evaluation, meanwhile, focuses on how a specific intervention works, for whom, and in what circumstances (Pawson and Tilley, 1996), and is designed to identify the combination of generative mechanisms and contextual characteristics in achieving outcomes (Fletcher et al., 2016). Moore et al., (2015) consider that exploring the causal assumptions underpinning the intervention and understanding how interventions work in practice are important for building a robust evidence base that can inform policy and practice, and whether it can be replicated or reproduced in different contexts. Alongside the implementation and analysis of new interventions, realist principles are also used to inform pragmatic process evaluation of interventions already in routine practice (Evans et al., 2015), allowing exploration of whether or not interventions are effective, and if so, how effective they are (Craig et al., 2008).

The realist evaluation used in our research explored if and how Active Life contributes to health and social care when compared to a public sector organisation running a similar intervention in a similar setting elsewhere. The public sector organisation was selected as a comparable initiative due to its similarities to the social enterprise in terms of the services offered, the community-based approach and the geographical characteristics of the area in which it operates. An inductive approach to analysis based on stakeholder participation was undertaken (Brandon, 1998). A total of 68 in depth semi-structured interviews involving beneficiaries of both social enterprise and public sector organisations (we followed them with repeated interviews over time), managers of the social enterprise and the public organisation, health

professionals who deal with chronic conditions, social enterprise leaders, managers of the main partner organisations and grant-makers were conducted between January 2015 and August 2016. Table 2 includes details of the study design, data collection and data analysis used in the realist evaluation.

-----Insert Table 2 around here-----

Quasi-experimental investigation, meanwhile, aims to assess the effectiveness of specific health interventions (Craig et al., 2008) and determine what works in terms of measurable outcomes (Fletcher et al., 2016; Moore et al., 2015). With this methodology, the results could directly be attributed to the organisation without the barriers connected to the feasibility of random allocation of participants (Shadish et al., 2002). An assessment was undertaken of a quasi-experimental investigation (employing both retrospective and prospective data) to explore how Active Life could contribute to health and social care in comparison with no intervention (retrospective pilot study) and a similar intervention provided by a public sector organisation (prospective pilot study). The intervention group in both parts of the study was composed of beneficiaries of the social enterprise presenting with obesity. In the retrospective part of the study, the comparator group was comprised of beneficiaries who were not referred to any intervention but presented with a similar medical condition to the intervention group. Two groups adding up to a total of 60 beneficiaries (n=30 in the intervention group, and n=30 in the comparator group) were considered as a possibly adequate sample for a retrospective feasibility study. Patient records for beneficiaries of both the intervention and the control groups were identified and data extracted through the involvement of general practitioners (GPs). In the prospective part of the study, the comparator group comprised beneficiaries of a publicly run organisation and two groups of 10 beneficiaries each (n=10 in the intervention

group, and n=10 in the comparator group) were recruited. Data about well-being, mental health and physical health were collected with a questionnaire. Table 3 includes details of the methodologies used in the quasi experimental investigation.

-----Insert Table 3 around here-----

Ethical approval was requested and obtained from the NHS Highland R&D Committee for the realist evaluation and from the Caldicott Guardian Approval, NHS Highland, for the retrospective pilot study. In both research approaches, ethical approval was requested and obtained from the University ethics committee.

While a great deal of data was obtained to help extend our knowledge of the work of social enterprise as a health care provider, the challenges and limitations in adopting the three different methods were also recorded and analysed across three years of research. Three main limitations were identified, each of which is discussed in turn.

Challenges and limitations of using health-based approaches to evidence the contribution of social enterprises

The first challenge relates to issues concerning the appropriateness of comparator groups which could help to attribute the impact to the social enterprise. The second relates to the possibility of generalizing results to the heterogeneous world of social enterprise. The third concerns the challenges which were related to the specificity of the methodological approaches adopted.

Identifying a suitable comparator for evaluating social enterprise

Difficulties were faced in all three methods in identifying a suitable *comparator* for a social enterprise and consequently addressing the question of what would happen without the social

enterprise intervention (Maxwell, 2012; Shadish et al., 2002). In the systematic review, two different comparator groups were identified, one providing similar services to the social enterprise but with a different organisational structure (run by the local authorities) and one providing ‘usual care’ by the NHS. When social enterprise settles in a space which neither statutory agencies nor private companies have occupied previously, providing additional services, the beneficiaries who are referred to ‘usual care’ were considered as part of a comparator group. No analysis could be conducted to assess the differences made by social enterprises in comparison with other providers, such as other forms of third sector organisation, or for-profit companies.

The identification and selection of the comparator group were also challenging in the realist evaluation and quasi-experimental investigation, although two possible comparator groups were identified: one comprised patients with similar chronic conditions that were not referred to the social enterprise; the other involved sampling patients living in a similar area and served by a similar but publicly-run service. The lack of similar services provided by different organisations in the same area was the main challenge encountered in trying to identify and select the comparator group for both the realist evaluation and the quasi-experimental investigation.

Generalisability of results to the heterogeneous world of social enterprise

When we consider the generalisability of findings, we must also consider what we mean by ‘social enterprise’, given that the term is relatively new and covers a wide and complex variety of organisational forms (Peattie and Morley, 2008; Young et al., 2016). The relative novelty of the social enterprise term determined the low number of papers that actually evaluate organisations specifically labelled as such. For example, in the systematic review, a small

number of papers with a high level of heterogeneity, in terms of the nature of the social enterprises studied, and the outcomes they delivered, affected the generalisability of the results. Synthesising and making sense of such heterogeneity was not a straightforward task.

In the realist evaluation and quasi-experimental investigation, it was not possible to generalise results generated by a single case study to the heterogeneous world of social enterprise. Generalisability requires theory (in realist evaluation) or results (quasi-experimental investigation) to be tested in different contexts and with different comparator groups (such as for-profit providers, or other types of third sector organisations). In the quasi-experimental investigation, very little could be said about how the findings might be applied in new settings or among other populations. Thus, this approach could not assess the transferability of specific embedded and contextualised characteristics of social enterprises that could constitute an ‘added value’ (or not) to the status quo.

Limitations in design and data collection

In the realist evaluation, difficulties were experienced in understanding the extent to which the results could be attributed to the social enterprise. Although a comparator group was included, some other potential explanations for the efficacy of the programme were also identified. Participation in other programmes and interventions in the community could have impacted upon the development of specific mechanisms, affecting the attribution of the outcome patterns to the social enterprise initiative. That said, this is not altogether uncommon in evaluations of this type (Fletcher et al., 2016).

Our research also confirmed that conducting a realist evaluation is time- and labour-intensive. This was particularly true of the data collection phase, where the recruitment of an adequate number of beneficiaries over a short period of time was challenging and time consuming due

to the small number of beneficiaries involved in these initiatives, and the presence of complex medical issues which affected their participation. Moreover, it was difficult to include the views of the most marginalised beneficiaries, who were less likely to take part in our research. This could have affected some of results identified.

In the quasi-experimental investigation, difficulties in collecting complete datasets were faced in both parts of the feasibility study. In the prospective analysis, challenges in involving and retaining beneficiaries (not only in the research process, but also inside the intervention itself) and the interviewees' difficulties in recalling their medical results (such as blood sugar, blood pressure and weight) affected the completeness of the datasets. In the retrospective analysis, barriers to involving GPs in sharing data due to their lack of time and resources, alongside a lack of completeness of the records extracted, ensured that the process of undertaking statistical analysis and propensity score matching¹ was highly problematic. Even although GPs in the area had agreed to participate in these studies, practical difficulties in involving them in the collection of data were particularly challenging. A lack of resources and time were the main factors identified by GPs for not being able to commit to the study as hoped, potentially reducing the 'completeness' of the research. This meant, for example, that more than eight months passed between receiving ethical approval and the dataset being shared. In addition to that, different reasons were established for the absence of medical records at different stages of research. First, the database was created retrospectively, and data were originally collected for reasons not inherent to the study. Second, it is possible that some patients did not need to go back to their GP to have their physical health checked during the course of the research.

¹ Propensity score matching is a statistical matching technique that aims at matching sets of treated and untreated subjects who share a similar value of the propensity score. Once a matched sample has been formed the treatment effect can be estimated by directly comparing outcomes between treated and untreated subjects.

Finally, some of the clinical outcomes were not controlled within the time horizon of the research.

The path of causation between the intervention and the results was difficult to establish. In the prospective part of the study, the outcome changes within the social enterprise group showed positive results in relation to two outcomes. However, these changes were not statistically significant (or they were false-positive results) when we applied Bonferroni correction, which is a statistical adjustment when several dependent or independent statistical tests are being performed simultaneously on a single data set (Napierala, 2012). The findings did not consider the beneficiaries' adherence to the programme (few beneficiaries took part in the intervention in the long-term, quitting during the six months of the data collection for a wide range of different reasons), their participation in other interventions and/or the deterioration of the medical conditions, due to a variety of circumstances. Similar barriers were faced in the retrospective part of the study: it was not possible to explore the adherence of beneficiaries to the intervention, nor was it possible to explore the other activities they may have been undertaking that could have affected the study results. Moreover, the achievement of outcomes such as connectedness and social well-being were not easily detected through this approach.

What we have learned: considerations for the future evaluation of social enterprise

Reflecting upon the limitations of the three methods used in this study, three main lessons can be drawn. The first relates to choosing the most appropriate comparator, the second relates to the possibility of generalising results linked to the appropriateness of employing methods commonly used in public health into the context of social enterprise, while the third concerns the possibility of overcoming the limitations posed by the methodological approaches.

To choose the right comparator, the health literature discusses at least five different possibilities to consider (Setoguchi and Gerhard, 2013): alternative treatment; the standard (or ‘usual’) care; historical comparison; different data comparison; and no treatment. All these comparisons could be applied in research analysing the effectiveness of social enterprises. For example, in our research, alternative treatment and usual care were used as comparator groups across the three methods adopted. Selecting alternative treatment was useful in understanding the differences in mechanisms, contextual variables and outcome patterns between social enterprise and a similar service offered by a public provider and helped to build understanding of whether social enterprises are more, less or as effective than publicly run organisations providing the same services. We have also identified a second possible comparator group, composed of beneficiaries following usual care procedure. This comparator was selected to explore the outcomes of a specific social enterprise providing supplementary services in areas that are underserved by public or private providers. Historical comparisons and comparison groups from different data sources could also be used when exploring if a specific intervention provided by a social enterprise is effective.

Although all the possibilities discussed here could be considered as valid comparators, at the end of our research, we have understood that there is not only one right comparator for all the studies related to social enterprises. In order to choose the ‘right’ comparator, it is necessary to define clearly what it is meant by ‘usual care’ (in a specific setting), while recognising the diversity (for example, in terms of third sector provision) of the existing service providers and community initiatives. The comparator group selected should reflect meaningful choices in policy and in real world practice and should be chosen based on the consequent study question to be addressed. Context therefore is a key aspect to analyse to explore how to evaluate the intervention (Shoveller et al., 2016).

Difficulties concerning the generalisability of results were also faced in all three of the methodological approaches employed. In the systematic review, studies included were heterogeneous in terms of contexts, interventions and beneficiaries. Also, in the realist evaluation, it was not possible to generalise results, generated by a single case study, to the heterogeneous world of social enterprise. In the quasi-experimental investigation, very little could be said about how the findings might be applied in new settings or among other populations. We are acutely aware that local idiosyncrasies (Roy et al., 2017) and different historical, political, social, and cultural aspects can combine to influence the role that social enterprise plays within the provision of public services (Kerlin, 2013; Teasdale, 2012), which then must somehow be taken into consideration before any conclusive results can be claimed in relation to the general contribution of social enterprises to different health and social care settings.

Some scholars have discussed the importance of creating ‘one size fits all’ tools to compare the social value generated by very different social enterprises as a solution to the need to increase generalisability (Kroeger and Weber, 2014). However, contextual variables, as we have seen across the three methods, can greatly affect how the social enterprise ‘works’ and the achievement of outcomes. Thus, tools that do not consider the complexity of interventions and the contextual variations involved will be of limited use in adding to the evidence base and informing policymakers. Conducting the same process with the same tool for all organisations and measuring the same outcomes could actually also have the perverse result of producing negative effects on an organisation’s ability to address its social mission. It can affect the nature of the social enterprise and reduce the legitimacy of the organisation, particularly if it is done in connection with accessing public funds. The adoption by policymakers of only one common framework, based on specific common objectives, could create an instrument promoting only

specific areas of policy, or investing only in specific organisations, with a resultant reduction in the currently wide variety of social enterprises. It could also cause a growing tension in social enterprises between their social mission and financial sustainability and force the organisation to adopt more market-oriented behaviours or to adapt to the characteristics of the public sector. There are inherent risks both to the organisations themselves and to society as a whole in encouraging such behaviour (Eikenberry and Kluver, 2004; Zimmerman and Dart, 1998). To reduce this risk, the objectives, processes and the time horizon of the evaluation should be carefully considered in relation to the specific social enterprise or ‘intervention’ assessed. The choice of the design, in other words, should be in line with the rationale behind the evaluation to be conducted and the context in which social enterprises are operating.

The development of rigorous evidence-based studies which use integrative methodological approaches could instead help to resolve the challenges inherent to any single approach. Lessons on how social enterprises contribute to health and social care could be drawn from each of the approaches used in this research. Each of the approaches used provide some additional information that, when put together, increases the overall understanding of whether, how, and in what role social enterprise could add positive value to health and social care settings. Thanks to the systematic review, it was possible to understand that social enterprises may lead to improved health outcomes, particularly as regards well-being and mental health, when they are integrated into the existing service provision and relatively free to focus on social benefits rather than engage in intense competition for funding. Through the realist evaluation, it was possible to conclude that social enterprises have the ability to work flexibly and act as a ‘boundary spanner’ on behalf of beneficiaries. However, the advantages that social enterprises can accrue can also be undone without sufficient ongoing financial support and stability. The quasi-experimental investigation instead reinforced the importance of a physical activity

programme for people with complex health issues and the advantages that organising such activities via community-based organisations can provide, especially in rural settings.

The integration of different approaches could then reduce the limitations that each method might present if applied alone. For example, a realist evaluation could include a quantitative analysis such as a full quasi-experimental investigation, so as to inform the outcome patterns of both intervention and comparison groups, as suggested by Pawson (2013). Such work could be augmented by a systematic literature review to inform the theory behind how and in what circumstances social enterprises might provide positive results. A mixed methods study could fully explore if, how and for whom a specific social enterprise could achieve a positive impact in comparison with alternative providers or the usual care. Since realist evaluation has become recognised as a suitable method for evaluating complex health interventions, scholars have discussed the possibility of integrating this with different approaches which analyse differential outcomes and that are widely used in the evaluation of complex health interventions (Deaton and Cartwright, 2018). While some researchers imply that it is possible to develop ‘realist experimental investigation’ (Bonell et al., 2012), for example) others see this as an oxymoron because of irreconcilable ontological and epistemological positions which do not sufficiently allow for complexity (Hawkins, 2016; Marchal et al., 2013; Porter and O’Halloran, 2012). Our contribution shows that the ‘right’ methodological approach to choose depends again on the field of inquiry (context) being explored. Therefore, the methodological approaches involved in undertaking a realist evaluation should take the research questions into consideration, by determining what the best approaches are for unpacking the ‘black box’ of a programme. Pragmatically speaking, this could necessitate attempting to combining study design approaches that might be seen as having different epistemological and ontological positions

and develop new theoretical and methodological work (Shoveller et al.,2016). The context of the evaluation, in realist terms, then helps to reconcile the different philosophical positions.

Conclusion: lessons for social enterprise, and for social policy

Despite being framed as a potential means of delivering a wide range of public services, to date, comprehensive and rigorous processes of analysis of the contribution of community-based social enterprises to health and social care have been lacking. In the last couple of decades, investors, funders, policymakers and researchers have all tried to reach agreement on ‘a common one-size-fits-all-framework’ for evaluating such types of organisation (Dey and Gibbon, 2017). Different tools and approaches have been widely discussed to try to achieve a consensus on the ‘best way’ of measuring social enterprise value (Kah and Akenroye, 2020). However, the findings of this research raise doubts about the utility of adopting only one tool. Context in fact is a key aspect for understanding public health complex interventions and thus necessitates the employment of new theoretical and methodological approaches that are contextually dependant (Hervieux and Voltan, 2019; Shoveller et al., 2016). If the aim is to inform policymakers, then integrated approaches, such as those used in this research are likely more appropriate.

Reflecting upon our initial research question, as to whether methods from public health research could be used to assess social enterprise, we have come up with the following recommendations: First of all, in our research it has been argued that only through the involvement of different comparator groups, based on the research questions addressed, would it be possible to disentangle the embedded characteristics of organisations such as social enterprises. Selection of comparator groups (if a comparator group is important to address the research question) should be based on an in-depth analysis of the context in which the social

enterprise is embedded, and the comparator group chosen should be based upon the policy aims the evaluation wishes to address. Second, each of the methods adopted in this research is time-consuming and resource intensive and requires the researcher to possess advanced skills. Public officials should recognise the complexity and resource-intensive nature of such evaluation, and resource it accordingly, particularly to feature the resource implications of involvement of stakeholders and beneficiaries in the contexts and wider systems within which social enterprises operate, which may include specific funding for data collection. Finally, if the aim of policymakers is to understand the added value of social enterprise organisations, an integrative research approach combining different research methods and design should be implemented to improve generalisability. Ideally, future studies should use integrative and consistent methods and test results in different settings and context. This would allow, through combination of results, identification of the specific contributions of social enterprise and account for context more readily.

References

- Alcock, P., Millar, R., Hall, K., Lyon, F., Nicholls, A., Gabriel, M., 2012. Start-up and Growth: National Evaluation of the Social Enterprise Investment Fund (SEIF) (Report), Third Sector Research Centre. Third Sector Research Centre and Health Services Management Centre, Birmingham.
- Arvidson, M., Lyon, F., 2014. Social Impact Measurement and Non-profit Organisations: Compliance, Resistance, and Promotion. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations* 25, 869–886. <https://doi.org/10.1007/s11266-013-9373-6>
- Arvidson, M., Lyon, F., McKay, S., Moro, D., 2013. Valuing the Social? The Nature and Controversies of Measuring Social Return on Investment (SROI). *Voluntary Sector Review* 4, 3–18. <https://doi.org/10.1332/204080513X661554>
- Bagnoli, L., Megali, C., 2011. Measuring Performance in Social Enterprises. *Nonprofit and Voluntary Sector Quarterly* 40, 149–165.
- Bellucci, M., Nitti C., Franchi S., Testi E., Bagnoli, L., 2019. Accounting for social return on investment (SROI): The costs and benefits of family-centred care by the Ronald McDonald House Charities. *Social Enterprise Journal*, 15, 1, 46-75.
- Bonell, C., Fletcher, A., Morton, M., Lorenc, T., Moore, L., 2012. Realist Randomised Controlled Trials: A New Approach to Evaluating Complex Public Health Interventions. *Social Science & Medicine, Part Special Issue: Place, migration & health* 75, 2299–2306. <https://doi.org/10.1016/j.socscimed.2012.08.032>
- Borzaga, C., Defourny, J. (Eds.), 2001. *The Emergence of Social Enterprise*. Routledge, London.
- Bovaird, T., 2014. Efficiency in Third Sector Partnerships for Delivering Local Government Services: The role of economies of scale, scope and learning. *Public Management Review* 16, 1067–1090. <https://doi.org/10.1080/14719037.2014.930508>
- Buckingham, H., 2009. Competition and Contracts in the Voluntary Sector: Exploring the Implications for Homelessness Service Providers in Southampton. *Policy & Politics* 37, 235–254. <https://doi.org/10.1332/030557309X441045>
- Byford, S., Barrett, B., Dubourg, R., Francis, J., Sisk, J., 2010. The Role of Economic Evidence in Formulation of Public Policy and Practice, in: Shemilt, I., Mugford, M., Vale, L., Marsh, K., Donaldson, C. (Eds.), *Evidence-Based Decisions and Economics*. Wiley-Blackwell, Oxford, pp. 43–55. <https://doi.org/10.1002/9781444320398.ch4>
- Calò, F., Teasdale, S., Donaldson, C., Roy, M.J., Baglioni, S., 2018. Collaborator or Competitor: Assessing the Evidence Supporting the Role of Social Enterprise in Health and Social Care. *Public Management Review* 20, 1790–1814. <https://doi.org/10.1080/14719037.2017.1417467>

- Chalmers, I., Hedges, L.V., Cooper, H., 2002. A Brief History of Research Synthesis. *Evaluation & the Health Professions* 25, 12–37. <https://doi.org/10.1177/0163278702025001003>
- Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., Petticrew, M., 2008. Developing and Evaluating Complex Interventions: the New Medical Research Council Guidance. *BMJ* 337, a1655. <https://doi.org/10.1136/bmj.a1655>
- Deaton, A., Cartwright, N., 2018. Understanding and Misunderstanding Randomized Controlled Trials. *Social Science & Medicine* 210, 2–21. <https://doi.org/10.1016/j.socscimed.2017.12.005>
- Defourny, J., Nyssens, M., 2006. Defining Social Enterprise, in: Nyssens, M. (Ed.), *Social Enterprise: At the Crossroads of Market, Public Policies and Civil Society.*, Routledge Series in the Management of Voluntary and Non-Profit Organizations. Routledge, Abingdon, Oxon, pp. 3–26.
- Dey, C., Gibbon, J., 2017. Moving on from Scaling Up: Further Progress in Developing Social Impact Measurement in the Third Sector. *Social and Environmental Accountability Journal* 37, 66–72. <https://doi.org/10.1080/0969160X.2017.1285712>
- Donaldson, C., Mugford, M., Vale, L.D. (Eds.), 2002. *Evidence-Based Health Economics: From Effectiveness to Efficiency in Systematic Review.* BMJ Books, London.
- Eikenberry, A.M., Kluver, J.D., 2004. The Marketization of the Nonprofit Sector: Civil Society at Risk? *Public Administration Review* 64, 132–140. <https://doi.org/10.1111/j.1540-6210.2004.00355.x>
- Eisenhardt, K.M., Graebner, M.E., 2007. Theory Building From Cases: Opportunities And Challenges. *Academy of Management Journal* 50, 25–32. <https://doi.org/10.5465/AMJ.2007.24160888>
- Elmes, A.I., 2019. Health Impacts of a WISE: A Longitudinal Study. *Social Enterprise Journal* 15, 457–474. <https://doi.org/10.1108/SEJ-12-2018-0082>
- Epstein, D., Klerman, J.A., 2012. When is a program ready for rigorous impact evaluation? The role of a falsifiable logic model. *Eval Rev* 36, 375–401. <https://doi.org/10.1177/0193841X12474275>
- Evans, R., Scourfield, J., Murphy, S., 2015. Pragmatic, Formative Process Evaluations of Complex Interventions and Why We Need More of Them. *Journal of Epidemiology and Community Health* 925–926. <https://doi.org/10.1136/jech-2014-204806>
- Farmer, J., De Cotta, T., Kilpatrick, S., Barraket, J., Roy, M., Munoz, S.-A., 2019. How Work Integration Social Enterprises Help to Realize Capability: A Comparison of Three Australian Settings. *Journal of Social Entrepreneurship* 1–23. <https://doi.org/10.1080/19420676.2019.1671481>
- Farmer, J., Kamstra, P., Brennan-Horley, C., De Cotta, T., Roy, M., Barraket, J., Munoz, S.-A., Kilpatrick, S., 2020. Using Micro-Geography to Understand the Realisation of

- Wellbeing: A Qualitative GIS Study of Three Social Enterprises. *Health & Place*.
<https://doi.org/10.1016/j.healthplace.2020.102293>
- Farmer, J., Philip, L., King, G., Farrington, J., MacLeod, M., 2010. Territorial Tensions: Misaligned Management and Community Perspectives on Health Services for Older People in Remote Rural Areas. *Health & Place* 16, 275–283.
<https://doi.org/10.1016/j.healthplace.2009.10.010>
- Fletcher, A., Jamal, F., Moore, G., Evans, R.E., Murphy, S., Bonell, C., 2016. Realist Complex Intervention Science: Applying Realist Principles Across All Phases of the Medical Research Council Framework for Developing and Evaluating Complex Interventions. *Evaluation* 22, 286–303. <https://doi.org/10.1177/1356389016652743>
- Flockhart, A., 2005. Raising the Profile of Social Enterprises: The Use of Social Return on Investment (SROI) & Investment Ready Tools (IRT) to Bridge the Financial Credibility Gap. *Social Enterprise Journal* 1, 29–42.
- Hall, K., Alcock, P., Millar, R., 2012. Start Up and Sustainability: Marketisation and the Social Enterprise Investment Fund in England. *Journal of Social Policy* 41, 733–749.
<https://doi.org/10.1017/S0047279412000347>
- Hall, K., Arvidson, M., 2014. How Do We Know If Social Enterprise Works? Tools for Assessing Social Enterprise Performance, in: Denny, S., Seddon, F. (Eds.), *Social Enterprise: Accountability and Evaluation around the World*. Routledge, Abingdon, Oxon, pp. 141–159.
- Harlock, J., 2013. Impact Measurement Practice in the UK Third Sector: A Review of Emerging Evidence. Third Sector Research Centre Working Paper 106, Birmingham/Southampton.
- Hawkins, A.J., 2016. Realist evaluation and randomised controlled trials for testing program theory in complex social systems. *Evaluation* 22, 270–285.
<https://doi.org/10.1177/1356389016652744>
- Hervieux C., Voltan A. 2019. Toward a system approach to social impact assessment. *Social Enterprise Journal*, 15, 2, 264-285
- Kah S., Akenroye, T., 2020. Evaluation of social impact measurement tools and techniques: a systematic review of the literature. *Social Enterprise Journal*, Published online
- Kelly, D., Steiner, A., Mazzei, M., Baker, R., 2019. Filling a Void? The Role of Social Enterprise in Addressing Social Isolation and Loneliness in Rural Communities. *Journal of Rural Studies*. <https://doi.org/10.1016/j.jrurstud.2019.01.024>
- Kemm, J., 2006. The Limitations of ‘Evidence-Based’ Public Health. *Journal of Evaluation in Clinical Practice* 12, 319–324. <https://doi.org/10.1111/j.1365-2753.2006.00600.x>
- Kerlin, J.A., 2013. Defining Social Enterprise Across Different Contexts: A Conceptual Framework Based on Institutional Factors. *Nonprofit and Voluntary Sector Quarterly* 42, 84–108. <https://doi.org/10.1177/0899764011433040>

- Kerlin, J.A., 2006. Social Enterprise in the United States and Europe: Understanding and Learning from the Differences. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations* 17, 247–263.
- Kroeger, A., Weber, C., 2014. Developing a Conceptual Framework for Comparing Social Value Creation. *AMR* 39, 513–540. <https://doi.org/10.5465/amr.2012.0344>
- Leeuw, F.L., Vaessen, J., 2009. Impact Evaluations and Development: NONIE Guidance on Impact Evaluation. World Bank/Network of networks on impact evaluation, Washington, DC.
- Lyon, F., Gyateng, T., Pritchard, D., Vaze, P., Vickers, I., Webb, N., 2015. Opening access to administrative data for evaluating public services: The case of the Justice Data Lab. *Evaluation* 21, 232–247. <https://doi.org/10.1177/1356389015577507>
- Macaulay, B., Roy, M.J., Donaldson, C., Teasdale, S., Kay, A., 2017. Conceptualizing the Health and Well-Being Impacts of Social Enterprise: A UK-Based Study. *Health Promotion International*. <https://doi.org/10.1093/heapro/dax009>
- Marchal, B., Westhorp, G., Wong, G., Van Belle, S., Greenhalgh, T., Kegels, G., Pawson, R., 2013. Realist RCTs of Complex Interventions - an Oxymoron. *Soc Sci Med* 94, 124–128. <https://doi.org/10.1016/j.socscimed.2013.06.025>
- Maxwell, J.A., 2012. *A Realist Approach for Qualitative Research*. SAGE Publications, Thousand Oaks.
- Millar, R., Hall, K., 2013. Social Return on Investment (SROI) and Performance Measurement. *Public Management Review* 15, 923–941. <https://doi.org/10.1080/14719037.2012.698857>
- Milley, P., Szijarto, B., Svensson, K., Cousins, J.B., 2018. The evaluation of social innovation: A review and integration of the current empirical knowledge base. *Evaluation* 24, 237–258. <https://doi.org/10.1177/1356389018763242>
- Moore, G.F., Audrey, S., Barker, M., Bond, L., Bonell, C., Hardeman, W., Moore, L., O’Cathain, A., Tinati, T., Wight, D., Baird, J., 2015. Process Evaluation of Complex Interventions: Medical Research Council guidance. *British Medical Journal* 350, h1258. <https://doi.org/10.1136/bmj.h1258>
- Mulrow, C.D., 1994. Systematic Reviews: Rationale for Systematic Reviews. *BMJ* 309, 597–599. <https://doi.org/10.1136/bmj.309.6954.597>
- Napierala, M.A., 2012. What is the Bonferroni Calculation? *AAOS Now*.
- Pawson, R., 2013. *The Science of Evaluation: A Realist Manifesto*. SAGE, Thousand Oaks, CA.
- Peattie, K., Morley, A., 2008. Eight Paradoxes of the Social Enterprise Agenda. *Social Enterprise Journal* 4, 91–107.

- Porter, S., O'Halloran, P., 2012. The Use and Limitation of Realistic Evaluation as a Tool for Evidence-Based Practice: A Critical Realist Perspective. *Nursing Inquiry* 19, 18–28. <https://doi.org/10.1111/j.1440-1800.2011.00551.x>
- Roy, M.J., Donaldson, C., Baker, R., Kerr, S., 2014. The Potential of Social Enterprise to Enhance Health and Well-Being: A Model and Systematic Review. *Social Science & Medicine* 123, 182–193. <https://doi.org/10.1016/j.socscimed.2014.07.031>
- Roy, M.J., Lysaght, R., Krupa, T.M., 2017. Action on the Social Determinants of Health Through Social Enterprise. *Canadian Medical Association Journal (Journal de l'Association Medicale Canadienne)* 189, E440-441. <https://doi.org/10.1503/cmaj.160864>
- Setoguchi, S., Gerhard, T., 2013. Comparator Selection, in: *Developing a Protocol for Observational Comparative Effectiveness Research: A User's Guide*. Agency for Healthcare Research and Quality (US).
- Shadish, W.R., Cook, T.D., Campbell, D.T., 2002. *Experimental and Quasi-experimental Designs for Generalized Causal Inference*. Houghton Mifflin.
- Shemilt, I., Mugford, M., Vale, L., Marsh, K., Donaldson, C., 2010. *Evidence-Based Decisions and Economics: Health Care, Social Welfare, Education and Criminal Justice*. Wiley-Blackwell, Oxford.
- Smith, K., 2013. *Beyond Evidence Based Policy in Public Health: The Interplay of Ideas*, Palgrave Studies in Science, Knowledge and Policy. Palgrave Macmillan, Basingstoke, Hants.
- South, J., 2015. *A Guide to Community-Centred Approaches for Health and Wellbeing*. Public Health England, London.
- Suchowerska, R., Barraket, J., Qian, J., Mason, C., Farmer, J., Carey, G., Campbell, P., Joyce, A., 2019. An Organizational Approach to Understanding How Social Enterprises Address Health Inequities: A Scoping Review. *Journal of Social Entrepreneurship* 1–25. <https://doi.org/10.1080/19420676.2019.1640771>
- Teasdale, S., 2012. What's in a Name? Making Sense of Social Enterprise Discourses. *Public Policy and Administration* 27, 99–119. <https://doi.org/10.1177/0952076711401466>
- Vo, A.T., Christie, C.A., Rohanna, K., 2016. Understanding evaluation practice within the context of social investment. *Evaluation* 22, 470–488. <https://doi.org/10.1177/1356389016669447>
- Yin, R.K., 1994. *Case Study Research: Design and Methods*. SAGE Publications, Thousand Oaks.
- Young, D.R., Searing, E.A.M., Brewer, C.V., 2016. *The Social Enterprise Zoo: A Guide for Perplexed Scholars, Entrepreneurs, Philanthropists, Leaders, Investors, and Policymakers*. Edward Elgar Publishing, Cheltenham; Northampton.

Zimmerman, B., Dart, R., 1998. Charities Doing Commercial Ventures: Societal and Organizational Implications. Trillium Foundation; Canadian Policy Research Networks, Toronto; Ottawa.

Study Design	<p>The research question used for informing our study was: Do social enterprises provide improved outcomes in comparison with usual care in health and social care systems? Although not initially an intention of the study, when analysing the papers we began to notice differences between results which appeared to correlate to the contextual setting. Therefore, we subsequently differentiated results according to whether the social enterprise activity occurred in a competitive or collaborative healthcare system.</p>
Data Collection	<p>Inclusion</p> <p>To be included, studies had to meet the following criteria:</p> <ol style="list-style-type: none"> a) Primary research (any type of methods) evaluating and explaining the contribution in terms of efficacy and efficiency (i.e. quality of life, well-being and health) of social enterprise in comparison (either as a substitute or as a complement) to usual care providers in any publicly-funded health and social care; b) Only papers analysing social enterprises in publicly-funded health care systems or in sectors managed by the public part of non-publicly-funded health care systems were included. <p>The following databases were searched: Sociological Abstracts, Web of Knowledge, ASSIA, International Bibliography of social sciences, Cochrane, DARE, NHS EED and HTA, Campbell, PUBMED, Ethos, OpenGrey, PLANEX and CINHAL. In addition, specific journals relating to social enterprise and non-profits were screened. Of the 24,717 papers initially identified, 20% of the abstracts randomly selected after duplicates were removed were double screened to decrease possible biases.</p> <p>Quality Assessment</p> <p>The quality assessment of quantitative papers was undertaken using the EPHPP framework (Effective Public Health Practice Project). For qualitative papers, the framework by Mays and Pope (2000) was applied. For mixed methods papers the tool developed by Pluye et al. (2009) was used.</p>
Data Analysis	<p>Specific frameworks were used for extracting the data. The dimensions of these frameworks reflect the research methods used in the included studies, although the common categories used were study design, population, outcome and findings.</p> <p>Thematic analysis was brought to bear to identify common themes in the data (Barnett-Page and Thomas 2009).</p>

Table 1 – Systematic Review

Study Design	The research question used to inform our study was: How (mechanisms), and in what circumstances (contexts), social enterprises might produce positive health outcomes? The lack of previous research specifically framing social enterprise as a complex health intervention, and the significant lack of a realist evaluation in the field meant that an inductive approach for the realist evaluation was employed.
Data Collection	<p>A comprehensive approach to sampling based on stakeholder participation was undertaken.</p> <p>Different recruitment strategies were used in each case. Active Life beneficiaries were recruited directly by the researcher in situ. In contrast, beneficiaries in the control organisation were recruited through the fitness manager at Moving Well. After sharing their interest in being involved their details were passed on to the lead researcher who contacted them, explained the research and provided them with written materials to explain further, communicating their interest in being involved directly to the researcher without involving any members of Moving Well.</p> <p>A total of 68 in depth semi-structured interviews was undertaken. This involved beneficiaries (Active Life n = 22, Moving Well n = 25) of both organisations; managers of the social enterprise and the public organisation (Active Life n = 4, Moving Well n = 1), health professionals who deal with chronic conditions (n = 6); social enterprise leaders (n = 4); managers of the main partner organisations (n = 4) and grant-makers (n = 2). 65 interviews were recorded and transcribed ‘intelligent verbatim’, while the remaining three interviews were recorded via extensive field notes.</p> <p>The interview guidelines consisted up to 10 open-ended questions, depending on the groups of stakeholders involved.</p>
Data Analysis	The interviews were transcribed and, after ensuring that the transcripts were an accurate record of each interview, the data were imported into the computer-assisted qualitative data analysis software QSR NVivo to assist with two cycles of analysis.

Table 2 – Realist Evaluation

Study Design	The research question used to inform our study was: What is the contribution of a specific social enterprise to health and social care provision?
Data Collection	<p>Two phases of research were conducted:</p> <p>First Phase</p> <p>During the first phase an approach based on stakeholder participation was selected to gain insights into the research design. 24 interviews were conducted to explore research methods for future evaluation.</p> <p>Second Phase</p> <p>The feasibility of undertaking retrospective and prospective studies was assessed in the second phase of the research.</p> <p><i>Retrospective Study</i></p> <p>The intervention group was composed by beneficiaries of Active Life presenting with obesity (n=30 for retrospective study). In the retrospective pilot study the comparator group (n=30) comprised obese patients from the community (and surrounding areas) where Active Life is based, who followed traditional referral schemes (no physical activity intervention) but presented with a similar medical condition to the intervention group. Patient records for beneficiaries of both the intervention and the control groups were randomly identified by the local medical centre staff. Blood glucose, HbA1c, cholesterol (HDL and LDL), blood pressure, weight and BMI data were extracted (every six months)</p> <p><i>Prospective Study</i></p> <p>The intervention group was composed by beneficiaries of Active Life presenting with obesity (n=10). The comparator group (n=10) comprised beneficiaries of Moving Well. Beneficiaries in the intervention group were recruited directly by the researcher, while beneficiaries in the control group were recruited through the fitness manager at Moving Well. Mental health was assessed using the Warwick-Edinburgh Mental Well-being scale (WEMWBS) and Self-perceived health was assessed using EQ-5D. Data were collected at the baseline, after 3 months and after 6 months.</p>
Data Analysis	<p>First Phase</p> <p>In the first phase of research the interviews were transcribed by the author, and after ensuring that the transcripts were an accurate record of each interview, the data were then imported into the computer-assisted qualitative data analysis software QSR NVivo. The data collected were analysed in two cycles</p> <p>Second Phase</p> <p>In the retrospective study, an analysis of variation of data of Active Life beneficiaries over time was conducted. In the prospective analysis, variations of WEMWBS and EQ-5D of Active Life beneficiaries were explored, through a statistical analysis.</p>

Table 3 – Quasi-Experimental Investigation