



Process evaluations of mental health and psychosocial support interventions for populations affected by humanitarian crises

Alessandro Massazza^{a,*}, Carl R. May^{a,b}, Bayard Roberts^a, Wietse A. Tol^c, Sergiy Bogdanov^d, Abhijit Nadkarni^e, Daniela C. Fuhr^a

^a Department of Health Services Research and Policy, London School of Hygiene and Tropical Medicine, London, UK

^b NIHR North Thames Applied Research Collaborative, London, UK

^c Department of Public Health, University of Copenhagen, Copenhagen, Denmark

^d Centre for Mental Health and Psychosocial Support, National University of Kyiv-Mohyla Academy, Kyiv, Ukraine

^e Department of Population Health, London School of Hygiene and Tropical Medicine, London, UK

ARTICLE INFO

Keywords:

Process evaluation
Implementation science
Mental health and psychosocial support
Randomized controlled trial
Humanitarian crises

ABSTRACT

Background: Randomised controlled trials (RCTs) have been increasingly used to test the effectiveness of mental health and psychosocial support (MHPSS) interventions for populations affected by humanitarian crises. Process evaluations are often integrated within RCTs of psychological interventions to investigate the implementation of the intervention, the impact of context, and possible mechanisms of action. We aimed to explore limitations and strengths of how process evaluations are currently conceptualised and implemented within MHPSS RCTs specifically.

Methods: In April–June 2021 we conducted semi-structured interviews with 24 researchers involved in RCTs of MHPSS interventions in 23 different countries. Participants were selected based on systematic reviews of MHPSS interventions, funders' databases, and personal networks. Data were analysed using codebook thematic analysis.

Results: The conduct of process evaluations was characterized by high heterogeneity in perceived function, implementation outcomes assessed, and methods used. While process evaluations were overwhelmingly considered as an important component of an RCT, there were different opinions on their perceived quality. This could be explained by the varying prioritization of effectiveness data over implementation data, confusion around the nature of process evaluations, and challenges in the collection and analysis of process data in humanitarian settings. Various practical recommendations were made by participants to improve future process evaluations in relation to: (i) study design (e.g., embedding process evaluations in study protocol and overall study objectives); (ii) methods (e.g., use of mixed methods); and (iii) increased financial and human resources dedicated to process evaluations.

Conclusion: The current state of process evaluations in MHPSS RCTs is heterogeneous. The quality of process evaluations should be improved to strengthen implementation science of the growing number of evidence-informed MHPSS interventions.

1. Introduction

The randomised controlled clinical trial (RCT) is claimed to be the 'gold standard' for evaluating the effectiveness of health interventions (Mowat et al., 2018; Timmermans and Berg, 2010). An RCT is generally "a trial where subjects are randomly assigned to one of two groups: one (the experimental group) receiving the intervention that is being tested,

and the other (the comparison group or control) receiving an alternative (conventional) treatment" (Kendall, 2003, pg. 164). Critical research on RCTs has often focused on practices of design and evidence-making within trials and on the role of trials in forming the field of evidence-based medicine in high-income countries (Deaton and Cartwright, 2018; May et al., 2003; Rhodes and Lancaster, 2019; Will and Moreira, 2016). Common criticisms from the social sciences have

Abbreviations: RCT, Randomised controlled trial; MHPSS, Mental health and psychosocial support.

* Corresponding author. Department of Health Services Research and Policy, London School of Hygiene and Tropical Medicine, 15-17, Tavistock Place, London, WC1H 9SH, United Kingdom.

E-mail address: alessandro.massazza@lshtm.ac.uk (A. Massazza).

<https://doi.org/10.1016/j.socscimed.2022.114994>

Received 18 October 2021; Received in revised form 22 April 2022; Accepted 25 April 2022

Available online 29 April 2022

0277-9536/© 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

included epistemological challenges underpinning “evidence hierarchies”, limited relevance of average treatment effects for individual decision-making, the role of theory to inform the proper interpretation of RCTs, and issues concerning causal inference and external validity of RCT findings (Cohn et al., 2013; Horwitz and Singer, 2018). Taking a realist evaluation approach has been suggested as a possible solution to some of these limitations (Bonell et al., 2012). However, much less attention has been paid to the ways that RCT investigators understand and enact *within trials* the evaluation of technical and organisational processes through which trial outcomes are produced. Equally, we know little about the conduct of trials in the complex contexts of humanitarian settings, despite their unique characteristics.

1.1. Process evaluations in mental health and psychosocial support (MHPSS) interventions

Process evaluations are not new. Since the 1990s, there has been a growing movement to incorporate them within RCTs (Moore et al., 2015; Oakley et al., 2006), as both triallists and their funders have sought to understand how trial outcomes are reached, and as frameworks for understanding the processes through which trial interventions are realised and operationalised in practice have become available (Glasgow et al., 1999). For over 20 years, the Medical Research Council guidance on RCTs of ‘complex interventions’ has emphasised that process evaluations offer a way of understanding how dynamic interactions between different contextual, human, and socio-technical elements of trials lead to study outcomes, and materially shape the ‘evidence’ that such trials produce (Campbell et al., 2000; Moore et al., 2015; Skivington et al., 2021). These dynamics—and thus their evaluation—represent the everyday politics of knowledge production and evidence making (May and Ellis, 2001).

As these conceptual and methodological shifts have occurred, RCTs including process evaluations have tested the effectiveness of a diverse set of mental health and psychosocial support (MHPSS) interventions, including in humanitarian settings (Haroz et al., 2020; Purgato et al., 2018; Ryan et al., 2021). MHPSS has been defined as “any type of local or outside support that aims to protect or promote psychosocial well-being and/or prevent or treat mental disorders” (Inter-Agency Standing Committee, 2007; pg. 5) and includes a range of interventions implemented in settings affected by adversity such as armed conflict and disasters. Some recent RCTs of MHPSS interventions have included process evaluations (Greene et al., 2022; Harper Shehadeh et al., 2020; Uygun et al., 2020; van’t Hof et al., 2020). Such studies explore the intricate dynamics of interventions in contexts that are themselves very complex. However, little is known about the conduct of these trials in practice, or about the ways that they ultimately shape the implementation of complex interventions within trials and their dynamics (Panter-Brick et al., 2020). There is considerable debate amongst triallists and process evaluators about how to define and mobilise ideas about ‘context’ and whether to see it as a representation of socio-political, spatial, and sociotechnical circumstances, or as a representation of elements of a complex adaptive system that unfolds over time, and which then give rise to adaptive spaces and practices (May et al., 2016).

We decided to focus on RCTs of MHPSS interventions for various reasons. Firstly, humanitarian settings are characterised by multiple unique characteristics which have important implications for implementation research such as the complexity, turbulence, and socio-political instability inherent to these environments; lack of resources; and the multiple stressors that populations exposed to humanitarian crises experience (Mistry et al., 2021; Perera et al., 2020). Secondly, MHPSS interventions are increasingly becoming the norm in humanitarian response (Jones and Ventevogel, 2021), meaning that interrogating research surrounding their implementation constitutes an important priority for the humanitarian research community. Finally, we decided to focus on RCTs as they represent a common study design

used within the MHPSS field to test intervention effectiveness (Elrha, 2021).

1.2. The purpose of this paper

Because there is so little work on how process evaluations are formed and enacted in complex interventions in humanitarian settings, in this paper we begin to address this problem by focusing on trials conducted in humanitarian settings or among populations affected by humanitarian crises. We report on the ways that a group of investigators ($n = 24$, 13 female, 11 male) in such studies (a) account for the design and delivery of process evaluations within complex intervention RCTs of mental health and psychosocial support (MHPSS) interventions for populations affected by humanitarian crises, and (b) account for the ways that these process evaluations are shaped, not just by ideas about rigorous knowledge production and reliable evidence, but also by their socio-political and spatial contexts. Our study is necessarily descriptive. The assumption that the ensembles of beliefs, behaviours, and social practices with which we are concerned are best interrogated using theories—whether these are critical perspectives on knowledge production and practice, or practice theories intended to support evaluation and implementation—developed in well-resourced settings in high-income countries itself requires attention.

2. Methods

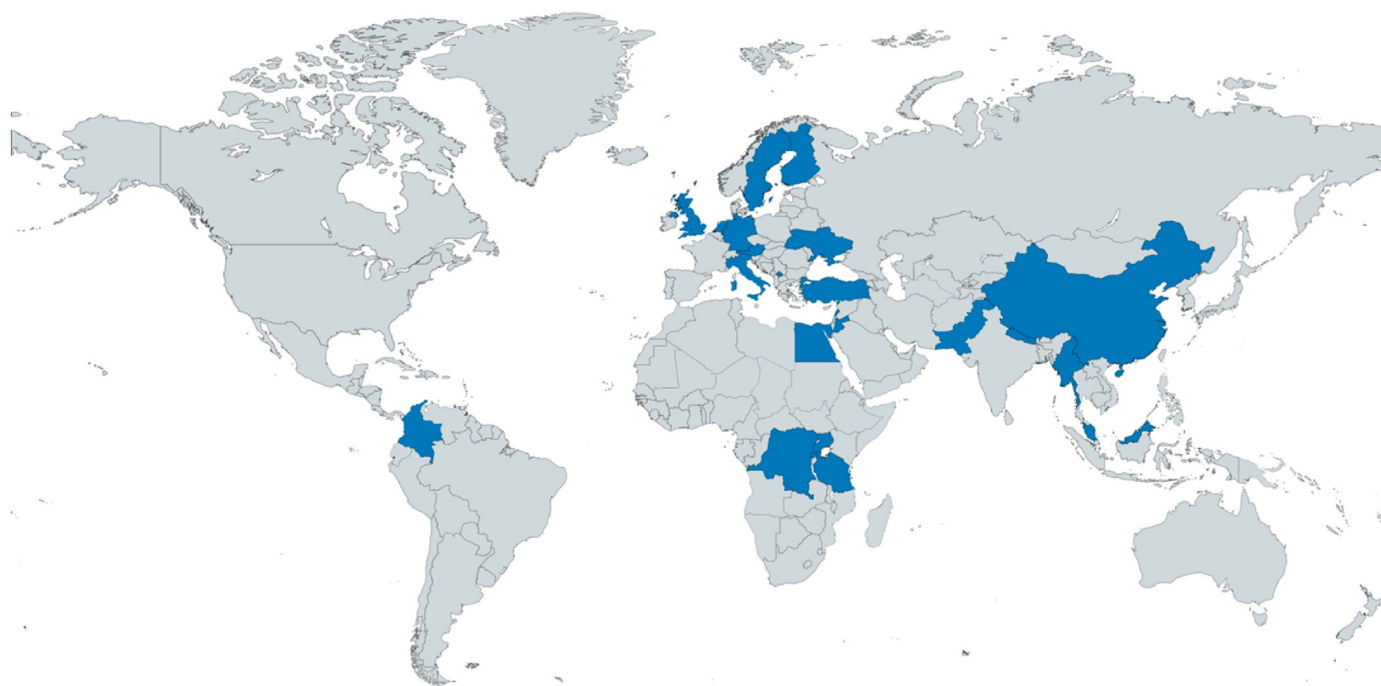
2.1. Participants and recruitment

Participants invited for interview were principal investigators and researchers who had been involved in the conduct of randomised controlled trials of MHPSS interventions among populations affected by humanitarian crises (e.g., populations currently residing in humanitarian settings or refugees displaced from humanitarian settings). We sought to interview the person who had been most involved in the process evaluation component of the RCT.

A purposive sampling strategy was used to identify participants. Participants were identified through existing systematic reviews of RCTs of MHPSS interventions (Purgato et al., 2018; Ryan et al., 2021), by assessing projects funded through the Elrha Research for Health in Humanitarian Crises (R2HC) programme (Tol et al., 2020), and through the personal networks of the authors who have had substantial involvement in the field. We deemed a participant eligible for participation if the RCT paper (or publications related to the RCT) described content falling under the MRC conceptualisation of a process evaluation (i.e., examining implementation outcomes such as fidelity and dose, examining mechanisms of impact, or examining the impact of context on implementation and outcomes) or if it was clearly stated that a process evaluation had been conducted as part of the RCT. This screening process was conducted by the first (AM) and last (DF) author. The email invitation clearly specified we were interested in the process evaluation component of their RCT. The geographical distribution of the RCTs covered by the participants in the current study is shown in Fig. 1.

2.2. Procedure

Semi-structured interviews were conducted in English with all the participants by the first author (AM). A topic guide was drafted by the authors and revised following the first two interviews, after which it remained unchanged (see Appendix 1). Prior to beginning the interview, a definition of process evaluation as per MRC guidance (Moore et al., 2015) was read out aloud to the participant to ensure participants had a shared understanding of the main construct under investigation. Participants were asked to focus specifically on the RCT for which they had been contacted but were advised they could draw on other relevant RCTs of MHPSS interventions they had been involved in. Interviews lasted approximately 1 h each and were conducted and audio recorded



Austria: 1; China: 1; Colombia: 1; Democratic Republic of the Congo: 1; Egypt: 1; Finland: 1; Germany: 2; Italy: 1; Jordan: 2; Kosovo: 1; Lebanon: 3; Malaysia: 1; Myanmar: 1; Netherlands: 1; Nepal: 2; Pakistan: 1; Rwanda: 1; Sweden: 1; Tanzania: 1; Turkey: 2; Uganda: 1; Ukraine: 1; United Kingdom: 1

Fig. 1. Geographical distribution of RCTs covered in the current study. Note: The higher number of RCTs than the number of participants is due to certain projects including multiple RCTs in different countries.

on Zoom (all interviews conducted with video except two, due to broadband issues) (*M*: 58 min, range: 30–70 min, *SD*: 9 min). Both principles of saturation (Hennink and Kaiser, 2022) and information power were used to determine our sample size with data collection stopping once no more participants fulfilling our inclusion criteria could be identified and once meaning and content saturation was reached as assessed during the interviews. Following the interview, participants were sent a brief demographic questionnaire.

Data collection took place from April–June 2021. The study was approved by the LSHTM Research Ethics Committee. Prior to taking part in the study, participants were provided with an information sheet over email and signed a written informed consent form.

2.3. Data analysis

All interviews were transcribed verbatim by the first author (AM). We used a codebook approach to thematic analysis as our method of choice for the analysis of the data (Braun and Clarke, 2021). During the process of transcription preliminary codes were noted down. Following a process of familiarisation with the entire dataset, a coding framework was developed. Another author (CM) read approximately 25% of all interviews and contributed to drafting the coding framework. The coding framework was then further discussed with the other authors and finalised (see Appendix 2). We used a hybrid approach towards the development of the coding framework encompassing a largely inductive framework with the inclusion of some methodological constructs (e.g., acceptability, appropriateness, costs, feasibility, fidelity, penetration, and sustainability) (Fereday and Muir-Cochrane, 2006; Proctor et al., 2011). All data was then coded in NVivo 12 by the first author (AM). As the coding took place, possible relationships between the codes were noted down using concept maps in NVivo 12.

3. Results

3.1. Characteristics of participants, interventions, and process evaluations

Participants had been involved in conducting interventions in 23 different countries (7 high-income countries and 16 low- and middle-income countries, see Fig. 1 for more precise breakdown). Multiple interventions were covered including Problem Management Plus, Common Elements Treatment Approach, Self-Help Plus, Step-by-Step, Community-Based Sociotherapy, Narrative Exposure Therapy, Interpersonal Therapy, Early Adolescent Skills for Emotions, and Integrative Adaptive Therapy, among others. See Table 1 for additional

Table 1
Demographic information on participants.

Construct	Percentage or mean
Gender	
Male	45% (<i>n</i> = 11)
Female	55% (<i>n</i> = 13)
Age	40.12 (<i>SD</i> = 9.05), range = 21–61
Country of residence (at time of trial)	
HIC	55% (<i>n</i> = 13)
LMIC	45% (<i>n</i> = 11)
Position within the RCT	
Principal investigator	38% (<i>n</i> = 9)
Project coordinator	17% (<i>n</i> = 4)
Co-investigator	13% (<i>n</i> = 3)
PhD student	13% (<i>n</i> = 3)
Postdoc/Research associate	8% (<i>n</i> = 2)
Research assistant	8% (<i>n</i> = 2)
Intervention expert/trainer	3% (<i>n</i> = 1)
Home discipline ^a	
Psychology or psychiatry	63% (<i>n</i> = 15)
Public health or epidemiology	33% (<i>n</i> = 8)
Social sciences (e.g., anthropology)	13% (<i>n</i> = 3)

^a Percentages for home discipline do not add up to 100% given some participants identifying with more than one discipline.

demographic information on the study participants. Additionally, descriptive information concerning the methods used and implementation outcomes assessed as part of the process evaluations discussed in this study are reported in [Appendix 3](#).

3.2. Perceived function and value of process evaluations as part of MHPSS RCTs

All participants had experience of process evaluations and most agreed that they were important components of RCTs. However, their perceived utility and purpose differed between participants. The function and value of process evaluations also changed depending on the stage of the project, with process evaluations conducted at the pilot RCT stage often focusing on acceptability and feasibility while process evaluations conducted at the definite RCT stage focusing more on potential for scale-up and fidelity.

For a substantial number of participants, a core function of process evaluations was that of providing insights into the implementation processes of an intervention, although they did so largely without recourse to ideas from implementation research itself. Process evaluation was described as a tool to understand the “experience on the ground”, what happens “in reality” or “behind the scenes” during an RCT, and the “nuts and bolts” of an intervention. Summative findings about effectiveness were regarded as insufficient, and so the purpose of process evaluations was seen as providing a foundation for the translation of experimental interventions into future practice by NGOs and other organisations. Even if an intervention was proven effective in an RCT, its future implementation beyond research settings was seen as largely dependent on factors such as feasibility and potential for scalability.

“As a researcher, you just want to know does it, what’s the effect size, but if you actually want to use those services, if you want to provide a certain treatment, you need to know what is required in order to do so, and you might have a treatment that has a great effect size, it’s really effective, but the implementation is so unsuitable for a particular context that you can’t use it”

[PI, Male, RCT conducted in LMIC]

Unveiling the implementation processes behind an intervention was also considered to be potentially helpful in going beyond the question of whether an intervention works or not, by focusing on the “how” and the “why” an intervention works or does not work. MHPSS interventions were described as “black boxes” or as objects that had to be “unpacked”. Participants often described that although interventions could be shown to be effective, it was not known “what makes an intervention work”.

“We need more of this [process evaluations], because we have such a wide range of programs and they all have different components and different active components and maybe not so active components right, [...] we’re like the, you know, throw the whole soup basket kind of thing and just figure out what sticks, and so, then we end up with this like mad litany of programs, and sometimes it seems like they’re effective and sometimes it seems like they’re not, but [...] there’s so much of that process stuff missing”

[Co-I, Female, RCT conducted in LMIC]

Participants’ accounts of process evaluation were formed around the importance of demonstrating that trials had been conducted in accordance with the design set out in their protocols, but process evaluations also provided a way of taking into consideration the impacts of context on intervention implementation. Process evaluation allowed participants to gain insight into the “parameters around the experiment [i.e., the RCT]” and “reality on the ground”. This was considered particularly important for various reasons. First, many RCTs were testing interventions that had not been developed locally but had been imported into humanitarian settings from high income countries. Process

evaluations therefore provided researchers with tools to explore assumptions around cross-cultural validity and consider the cultural appropriateness of mental health interventions.

“You can’t believe how many times I was proven wrong, I am from this population, and I had many assumptions that yes, this would work, or this would, and then I was shocked, no this is not what people want, they want different things”

[Project coordinator, Female, RCT conducted in LMIC]

Second, accounts of ‘context’ often referred to external disruptions that could have had significant impacts on the implementation of the intervention. These included a wide range of problems, ranging from the collapse of food distribution mechanisms in refugee camps, through natural disasters including flooding, to armed conflict that led to staff being evacuated from the region. The problem of context surrounding the intervention in humanitarian settings was overpowering, and two participants described feeling as though they had been “forced into” thinking about implementation research by working in such circumstances.

“It’s a mental health intervention, these are complex contexts, there’s so much going on, the idea that you can control for and kind of you know, try to write out a lot of that stuff to me is poppycock, it just it doesn’t make any sense”

[Co-I, Female, RCT conducted in LMIC]

“You know I think it’s [process evaluation] even so much more important in humanitarian settings than elsewhere, [...], they are kind of by definition chaotic settings right and so, [...] I think everything process related is just so much more critical to document because it is such a rapidly changing situation often so again just to kind of help make sense, of whatever you find so it’s not just such a black box in the middle ...”

[Co-I, Female, RCT conducted in LMIC]

The role of process evaluations in demonstrating fidelity to the protocol and accounting for significant external disruptions were particularly important when RCTs failed to show that interventions were effective, or when the RCT itself broke down once in process. One participant described how their process evaluation had transformed into a “post-mortem” assessment of the intervention once it became clear it would have not been possible to conduct the RCT. Another researcher described a situation whereby the same intervention had been implemented in two different settings with the intervention proving ‘effective’ in one but not the other, and that “*we were looking for everything, we were desperate to find what was making this difference*”. Process evaluations were therefore conceptualised as holding within themselves an explanatory potential that might be mobilised to understand failure.

“During some of our trials, we had COVID happening, and we had, [large scale national emergency] happening, [...], it [process evaluation] helps you understand your data better, particularly if you end up getting a negative trial, negative result that, you know, was there a big event that happened, was there a big conflict, [...] that might explain the results”

[Intervention expert/Trainer, Male]

Of course, in ‘successful’ studies, the process evaluation was described by some as a framework for reporting the implementation challenges encountered during the implementation of the intervention. Again, this was perceived as being particularly helpful in the context of volatile humanitarian settings, and for those who might translate the intervention into real world practice in the future.

“Particularly in the low- and middle-income country type work, humanitarian work, there are, we all do these trials, we all get burned, we all learn lessons the hard way, but then it doesn’t actually

get documented in an accessible way, so people learn those lessons, so that they don't repeat the same stupid mistakes"

[Co-PI, Male, RCT conducted in LMIC]

"I think the good thing about the process evaluation frameworks is they provide some structure to talking about challenges and failures and, maybe failures is a little bit strong but like there were some failures within this project for sure and so I think that [...] mainstreaming the integration of implementation outcomes could help normalize some of the disclosure of challenges or at least provide an avenue to do that"

[Research associate, Female, RCT conducted in LMIC]

3.3. Confusion, ambiguity, and lack of knowledge concerning process evaluation

Although many participants were committed to process evaluations, their ideas about what these consisted of were often uncertain. One participant described process evaluation as an "amorphous idea". Another described process evaluation as being about "all the other stuff" beyond the measurement of the primary outcome. There was also a degree of ambiguity concerning what specific implementation outcomes meant with participants at times mixing up different constructs one with another (e.g., fidelity as dose) or being confused about what counts as process evaluation and what does not.

I don't think it's [process evaluation] a fixed kind of entity with its own truth out there, it may be that there are libraries full out there with "The process evaluation" and there are specific courses and people get to know "The thing" [...] but that's not my, that's not how I've been working with process evaluation. For me, process evaluation is this kind of amorphous group of techniques, that you use to understand the processes and context where an intervention is being implemented, yeah everything else beyond just changes in symptoms"

[PI, Male, RCT conducted in HIC]

In a field heavily reliant on a primary research method—the RCT—what type of data counted as process evaluation was an issue. For some participants process evaluations were, by definition, only concerned with the collection of qualitative data. These participants often used the terms "process evaluation" and "qualitative evaluation" interchangeably, with the process evaluation being described by one participant as "the qualitative part" of the RCT. This conceptualisation of process evaluation as inherently qualitative had two consequences. First, it meant that for some participants quantitative process indicators were sometimes not considered part of a process evaluation. Quantitative indicators of fidelity or dose were described as part of standard monitoring in an RCT, "standard trial stuff", "record keeping", and not part of the "real" process evaluation, i.e., the qualitative component. One participant declined to participate in our study by responding to our invitation that *"I am not sure we have much to contribute to the discussion beyond the standard fidelity checks"*.

"Interviewer: Would you consider other sort of more quantitative outcomes, implementation outcomes such as fidelity, dose as part of the process evaluation?"

Participant: Not really, because, in my head, I have always seen those as more quantitative, as part of any trial, we always assess fidelity [...], this is standard trial stuff"

[Co-PI, Male, RCT conducted in LMIC]

One possible reason why process evaluation was perceived as being mainly qualitative might have to do with the nature of the questions that process evaluations were thought to be able to answer. Participants often described process evaluation as a tool to explore the *experience* and the *perception* of the intervention, to get feedback from stakeholders, to

investigate the feelings, the impressions, and the thoughts of participants in relation to the intervention, what participants liked and what they did not like, what worked and what did not work. One participant defined process evaluation as a tool to assess "the feel" of an intervention and to uncover "the human aspect of a trial". The unquantifiability of many of these questions meant that qualitative data collection was often perceived as the method of choice.

"I think, I think that a lot of times, some of the process evaluations I've seen like they are really just people walking in and asking like "How did this go for you?""

[Research associate, Female, RCT conducted in LMIC]

One further driver behind the confusion around process evaluation was the perceived lack of knowledge expressed by some participants concerning what a process evaluation is. A small number of participants struggled to answer the question of whether a process evaluation had been conducted as part of their RCT as they did not know what a process evaluation was. Even when participants did know the term process evaluation, some participants highlighted how they did not feel familiar with what the construct stood for or what a process evaluation entailed precisely.

"I don't know enough about process evaluations to say either way, I never studied process evaluation in school as a, as a thing out there, with specific methodology and specific guidelines, eh, I read a paper here and there"

[PI, Male, RCT conducted in LMIC]

Lack of knowledge does not, of course, equate to lack of interest. However, some participants implied that although process evaluation and implementation science perspectives were gaining momentum in the MHPSS field, they were really matters of fashion. One participant was explicit about this, describing them as "trendy". Additionally, a further justification for a lack of knowledge around these perspectives was the idea that current implementation science frameworks had not been developed with resource-constrained or humanitarian settings in mind and were therefore not appropriate for these contexts. Such considerations were never applied to the theory and method of the RCTs themselves.

"I've been in the past few years really like deep into the implementation science frameworks like, none of which really fit many of the contexts I work in, a lot of the other implementation science frameworks that I'm familiar with really developed primarily in high income settings [...] just really don't fit these contexts"

[Research associate, Female, RCT conducted in LMIC]

Beyond the methodological questions around process evaluation, the process evaluation was sometimes described as having been an "after-thought", as an area of the study that could have been improved, or as something that was done as a "box ticking exercise". As we have noted above, the relative importance of the process evaluation data compared with the effectiveness data was central to this. Several participants described how an epistemological hierarchy of importance underpinned the RCT, with the effectiveness findings taking precedence and being "prioritized" as "evidence", even when the process evaluation offered explanations for the outcome. For example, one participant described how, because of the word limit in a journal, they decided to remove the entire process evaluation data to maintain the reporting of the effectiveness data.

"The first thing that you want to answer is if it [the intervention] works or not. [...] The main big question was, all eyes were on the results of the effectiveness trial, the definitive trial that we had to do"

[PI, Male, RCT conducted in LMIC]

"It [process evaluation] is the first thing to go if a budget, if budgets are cut, I think it would be the first thing that would go, because the, the concern is, we need the scientific validity of a trial"

[Co-I, Female, RCT conducted in LMIC]

Process evaluation became important when effectiveness data lost its allure. This happened when RCTs failed to prove effectiveness, either because of methodological problems in the RCT design, or because of problems in their "context". This meant that the value of process evaluation outcomes tended to be conceptualised as being conditional on the state of the evidence around the effectiveness of the intervention: clinical effectiveness needed to be established before data was collected about the processes through which an intervention might be translated into everyday practice. Because of this, some participants described how they decided to re-focus their RCT on their process evaluation data only when they realised that the intervention or the trial was destined to fail.

"The primary outcome and secondary outcomes, it wasn't significant, of course, because it's just, you know, it's a small sample, so at the end we kind of like decided to give priority to the qualitative data, to the process evaluation"

[Research assistant, Female, RCT conducted in LMIC]

The difference between process and outcome results was thus related to wider concerns amongst participants about the relative prestige of the two research enterprises. The experience of participants was that "high impact" journals were mainly interested in publishing effectiveness results. As one participant described: *"you want to get out the high impact paper with the effect sizes into the top journal"*. Some participants believed that funders themselves would have also been more interested in the effectiveness results than in the process evaluation findings. Here, process evaluation always posed the risk of creating doubt about the RCT by undermining the pristine picture presented by reports of effectiveness findings—characterised by one participant as the "nice clean trial paper"—if it drew attention to the messiness of implementation in humanitarian settings. Participants suggested that the process evaluations received less attention as they often relied on qualitative data which, in turn, was perceived as being less valid and reliable than the quantitative outcome data that elided context-dependent factors.

"I thought we need to know about the efficacy, I thought, people will want to know whether it works, [...] I want to publish this in a decent journal so, if I, if I only have data on implementation what do I do with that? That might be helpful for NGOs, but I need, [...], the funder wants to see output from that so I need to have publications, and I thought it would be very difficult to publish anything on it without having some data to show that actually it does work"

[PI, Male, RCT conducted in LMIC]

The prioritization of the effectiveness data over the process data was heightened by the way that resources and effort were often directed towards strengthening the internal validity of the RCT. This was further reinforced by the complex and resource-constrained humanitarian settings in which many of these RCTs took place meaning that two participants described being able to do a process evaluation in the first place as being a "luxury".

"I am not going to think about a process evaluation on top of this you know [laughs], it's too much, it's not possible, it was one after one, it's like earthquake, then two months after insecurity, then in the middle of these floods, and then we need to recruit someone and then this and that, and then we need to look, to search I mean for more money because we are running out of money and we have been so delayed and we need to raise funds and so on, it's a non-ending story"

[Project coordinator, Female, RCT conducted in LMIC]

Ensuring the proper conduct of an RCT was often described as a

hectic endeavour that ate up all energies and resources of the research and implementation teams leaving little space and time for thinking about the process evaluation. One participant described that they had been so busy thinking about properly setting up the RCT that they had forgotten as a team to include the process evaluation component in the grant write-up.

"Because doing a full RCT in a conflict setting itself is a big credit actually, [...] so there is so much on your head, so the process evaluation might not be the first thing on your mind, maybe it might be a second last thing at that moment in time"

[PI, Male, RCT conducted in LMIC]

Resources for managing and conducting the RCT were frequently described by participants as over-stretched and without sufficient time to properly manage the collection, analysis, and write-up of process evaluation data.

"The research program manager, who is responsible for collecting these data and doing these data was so overburden that I don't think, I'm not sure if they actually sat through the whole session [to assess fidelity], I think they would do what they could, show up when they could, try to fill this out and then go do the hundreds of other things they had to coordinate and do."

[Research associate, Female, RCT conducted in LMIC]

Taken together, all the above-mentioned factors contributed to heterogeneity in the quality of process evaluations in MHPSS RCTs, as perceived by study participants. These links are summarised visually below in Fig. 2. As a result, multiple practical recommendations were made by participants to improve the quality of process evaluations in MHPSS RCTs in the future. The full list of recommendations is shown below in Table 2. Generally, recommendations touched on (i) study design (e.g., embedding process evaluations in study protocol and overall study objectives); (ii) methods (e.g., use of mixed methods); and (iii) increased financial and human resources dedicated to process evaluations.

4. Discussion

Process evaluations are being increasingly integrated within RCTs of interventions for populations affected by humanitarian crises. As humanitarian settings are characterised by unique implementation challenges due to their unstable and volatile nature, investigating how process evaluations are conducted within such settings represents a valuable contribution to the implementation research literature. Our paper represents the first exploration of the ways that investigators account for their conceptualisation and enacting of procedures, methods, and implementation constructs to do this work in these settings. Although this is work that is done in highly unstable socio-political and diverse geographical settings, it involves importing into them methods and models developed in very different contexts. One implication of our work is that although these are crucial to understanding the ways in which such RCTs "succeed" or "fail", process evaluations are rarely accorded prestige or priority. Indeed, they often possess a secondary or insurance value as ways of accounting for circumstances in which the trialists themselves are unable to deliver the RCT according to its protocol and theory of change, or circumstances in which some external source of disruption makes the RCT impossible. While participants often focused attention on the normative expectation for publications in prestigious medical journals that focused on results, there was little evidence that they took a different view in most cases. Indeed, their accounts of process evaluations were often characterised by epistemological and methodological uncertainty.

This meant that in everyday practice process evaluations were often focused on the internal elements of the RCT, concerning how the intervention is implemented or received. Less attention was given to

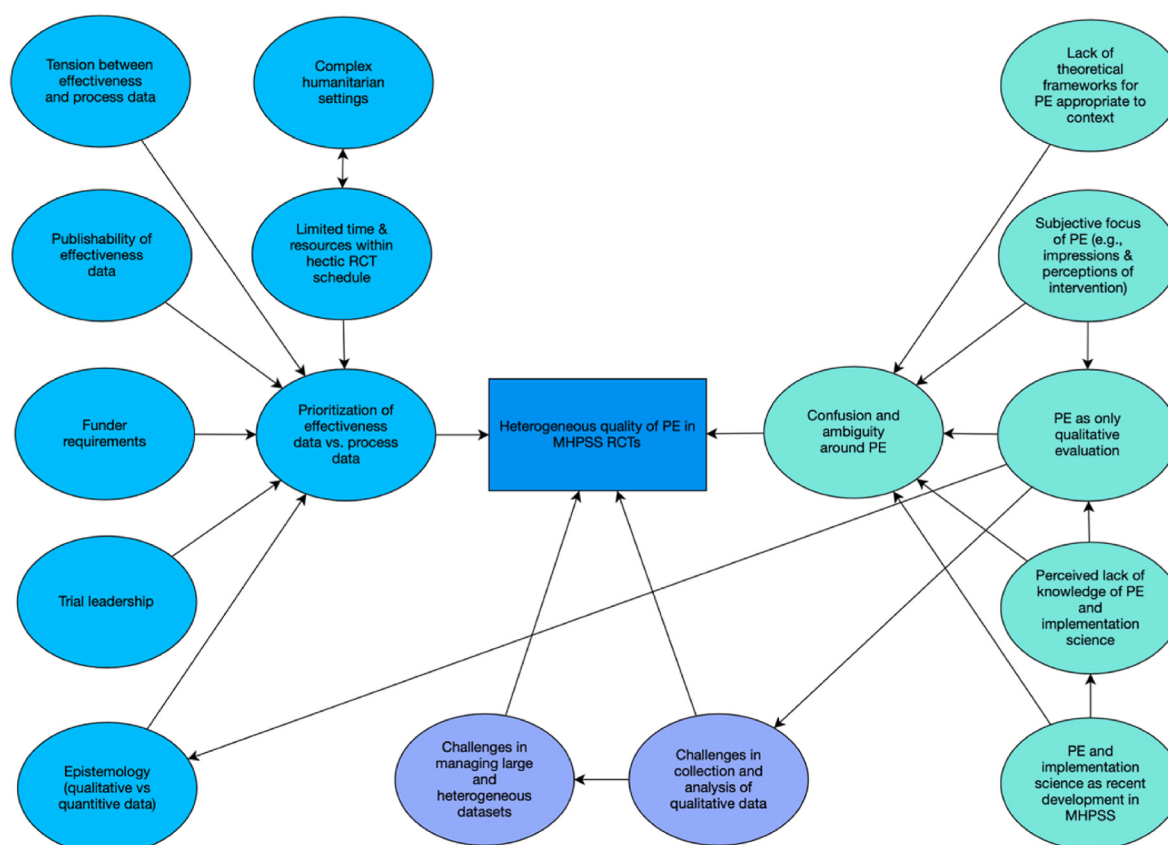


Fig. 2. Graphic summary of factors shaping the perceived quality of process evaluations of randomised controlled trials of mental health and psychosocial support interventions.

Table 2

Recommendations by participants to improve the conduct of process evaluations as part of MHPSS RCTs.

Study design

- Integrating process evaluation components throughout the project cycle (i.e., not only once the intervention has been concluded)
- Specifying objectives and anticipated outcomes of the process evaluation in advance, and supporting this by including process evaluations in RCT protocols
- Using frameworks or theories drawn from the implementation science literature to support the objectives and outcomes (bearing in mind need to adapt framework and/or theory to local context)
- Having a more iterative and flexible approach to the process evaluation (e.g., having previous rounds of process evaluation inform subsequent rounds of data collection)

- Ensuring some degree of independence between the trial team and the process evaluation team

- Thinking about process evaluation from the very beginning of the project

Methods

- Ensuring a mixed-methods approach is taken during the process evaluation
- Ensuring data is collected from a heterogeneous sample (both in terms of different types of stakeholders and different individuals within the same stakeholder group)
- Using frameworks and theories that help integrate heterogeneous types of data (e.g., large amounts of notes, logbooks, qualitative interviews, checklists etc.) and complex mixed-methods datasets

Increased financial and human resources

- Having members of the RCT team with expertise in process evaluation, implementation science, and social sciences
- Setting aside the proper amount of resources and time for the process evaluation (e.g., qualitative research is provided with the correct amount of resources to ensure high-quality data collection, analysis, and write-up)

Note. Recommendations were included in this box when they were provided by several participants. Quotes in support of the recommendations are provided in Appendix 4.

understanding the interaction between the intervention and the local context of implementation. This is not unique to process evaluations of interventions in humanitarian settings. In most research domains, process evaluations of complex intervention RCTs have largely focused on the “closed world of the intervention” (Morgan-Trimmer, 2015, pg. 298). The clinical literature in this field has thus remained *accontextual* (Wells et al., 2012). While there is a substantial empirical and theoretical literature from the social sciences on the construct of context as a multi-level, social, and dynamic factor (Morgan-Trimmer, 2015), this is not epistemologically consistent with the underpinnings of the RCT, even though complex systems perspectives may be particularly useful in the nonlinear, unpredictable, and constantly evolving nature of humanitarian settings (McGill et al., 2020).

Even so, participants considered process evaluations to have potential and sometimes real value. Participants described how the importance of process evaluations was amplified in humanitarian settings, due to the likely impact of the complex context on the intervention and RCT implementation. These discussions were however underpinned by implicit epistemological hierarchies concerning what counts as “best evidence” for clinical practice (May, 2006), with quantitative effect sizes from RCTs being actively placed at the pinnacle of this hierarchy. While this approach is largely reflexive of core narratives within the evidence-based medicine movement, it has also been subject to criticism with alternative frameworks such as “practice-based evidence” being proposed, including within MHPSS (Kienzler, 2019).

A certain degree of confusion around what constitutes a process evaluation also contributed to the perceived heterogeneous quality of process evaluations, with a small minority of participants struggling to answer whether a process evaluation had been conducted as part of their trial despite their answer clearly indicating that they did. A lack of clarity around what a process evaluation is and what its aims are is not

unique to the current sample, with other studies describing similar findings in other fields (Scott et al., 2019). Possible reasons for this are the current lack of clear operational guidance on how to conduct a process evaluation (Grant et al., 2013) as well as the complex nature of the implementation science field (e.g., more than 100 different published frameworks) (Best et al., 2021).

The lack of published high-quality process evaluations conducted in humanitarian settings is also a likely driver behind the confusion and lack of clarity. In the MHPSS field, process evaluations are often briefly reported in a sub-section of the main RCT paper and, occasionally, they are simply not published. While some high-quality stand-alone process evaluations conducted in LMICs have been published within the mental health field (Munodawafa et al., 2018), there is a dearth of examples in humanitarian settings specifically. This is true for implementation science research in general, with little work conducted in LMICs in general (Alonge et al., 2019) or humanitarian settings (Shahabuddin et al., 2020). Future studies should aim at always publishing their process evaluation findings in appropriate outlets and allow enough space for in-depth description of the data, if necessary, through a separate publication from the effectiveness findings.

4.1. Strengths and limitations

The current study has various limitations. A first limitation is the retrospective nature of the study. Various participants had worked on their RCT several years ago meaning that recall might have been an issue. An additional limitation was our sole focus on process evaluations in the context of RCTs, despite other study designs being available for the study of MHPSS interventions (Kohrt et al., 2015). However, we are not aware of process evaluations having been used with other study designs testing the effectiveness of MHPSS interventions. Finally, the scope of the study was limited by focusing only on process evaluations conducted as part of RCTs of MHPSS interventions. There are likely important insights that can be drawn from studies that included implementation components without specifically calling them a “process evaluation” and by process evaluations of interventions for other health conditions beyond mental health and conducted in other low-resource contexts.

Our paper also highlights several recommendations made by participants to ensure high-quality process evaluations in MHPSS RCTs in relation to (i) study design (e.g., embedding process evaluations in study protocol and overall study objectives); (ii) methods (e.g., use of mixed methods); and (iii) increased financial and human resources dedicated to process evaluations.

5. Conclusion

Many of the challenges facing global mental health today are concerned with implementation-type questions such as how to scale-up effective interventions to populations in need (Fuhr et al., 2020). Exploration of the development of MHPSS interventions in the last decade has highlighted the need for research to move beyond RCTs, and to begin to facilitate understanding of “how an intervention actually works in complex humanitarian settings with the myriad of challenges that arise and cannot always be predicted” (Elrha, 2021, pg. 44). Although some examples of such work exist (Cohen and Yaeger, 2021; Dickson and Bangpan, 2018), there is still space for improvement when it comes to integrating insights from implementation science into the MHPSS field. Our paper contributes to addressing this gap by providing an in-depth overview of conduct, barriers, and possible recommendations for the future of process evaluations in MHPSS RCTs.

Credit author statement

Alessandro Massazza: Conceptualisation, Formal analysis, Methodology, Writing; **Carl R. May:** Conceptualisation, Formal analysis, Methodology, Writing, Supervision; **Bayard Roberts:** Conceptualisation, Writing, Supervision; **Wietse A. Tol:** Conceptualisation, Writing; **Sergiy Bogdanov:** Conceptualisation, Writing; **Abhijit Nadkarni:** Conceptualisation, Writing; **Daniela C. Fuhr:** Conceptualisation, Formal analysis, Methodology, Writing, Supervision.

Acknowledgments

We would like to thank the participants for offering their time and their insights into this topic. This work was supported by the National Institute for Health Research (NIHR) (using the UK's Official Development Assistance (ODA) Funding) and Wellcome (grant reference number 219468/Z/19/Z) under the NIHR-Wellcome Partnership for Global Health Research. CM's contribution was also supported by NIHR Applied Research Collaborative, North Thames. The views expressed are those of the authors and not necessarily those of Wellcome, the NIHR, or the Department of Health and Social Care.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.socscimed.2022.114994>.

References

- Alonge, O., Rodriguez, D.C., Brandes, N., Geng, E., Reveiz, L., Peters, D.H., 2019. How is implementation research applied to advance health in low-income and middle-income countries? *BMJ Glob. Health* 4 (2), e001257. <https://doi.org/10.1136/bmjgh-2018-001257>.
- Best, S., Hutchinson, K., Gul, H., Ryder, T., Rapport, F., Hibbert, P., 2021. Does Implementation Science Have an Implementation Problem? https://www.mq.edu.au/_data/assets/pdf_file/0011/1178435/EIS-Report-Final-20210719-website.pdf.
- Bonell, C., Fletcher, A., Morton, M., Lorenc, T., Moore, L., 2012. Realist randomised controlled trials: a new approach to evaluating complex public health interventions. *Soc. Sci. Med.* 75 (12), 2299–2306. <https://doi.org/10.1016/j.socscimed.2012.08.032>, 1982.
- Braun, V., Clarke, V., 2021. Conceptual and Design Thinking for Thematic Analysis. *Qualitative Psychology*. [doi:10.1037/qp0000196](https://doi.org/10.1037/qp0000196).
- Campbell, M., Fitzpatrick, R., Haines, A., Kinmonth, A.L., Sandercock, P., Spiegelhalter, D., Tyrer, P., 2000. Framework for design and evaluation of complex interventions to improve health. *Br. Med. J.* 321 (7262), 694–696. <https://doi.org/10.1136/bmj.321.7262.694>.
- Cohen, F., Yaeger, L., 2021. Task-shifting for refugee mental health and psychosocial support: a scoping review of services in humanitarian settings through the lens of RE-AIM. *Implement. Res. Pract.* 2 <https://doi.org/10.1177/2633489521998790>, 2633489521998790.
- Cohn, S., Clinch, M., Bunn, C., Stronge, P., 2013. Entangled complexity: why complex interventions are just not complicated enough. *J. Health Serv. Res. Pol.* 18 (1), 40–43. <https://doi.org/10.1258/jhsrp.2012.012036>.
- Deaton, A., Cartwright, N., 2018. Understanding and misunderstanding randomized controlled trials. *Soc. Sci. Med.* 210, 2–21. <https://doi.org/10.1016/j.socscimed.2017.12.005>.
- Dickson, K., Bangpan, M., 2018. What are the barriers to, and facilitators of, implementing and receiving MHPSS programmes delivered to populations affected by humanitarian emergencies? A qualitative evidence synthesis. *Global Mental Health* 5. <https://doi.org/10.1017/gmh.2018.12>.
- Elrha, 2021. Review and Assessment of Mental Health and Psychosocial Support Intervention Research in Humanitarian Settings. <https://www.elrha.org/wp-content/uploads/2021/02/MHPSS-research-review-2021.pdf>.
- Fereday, J., Muir-Cochrane, E., 2006. Demonstrating rigor using thematic analysis: a hybrid approach of inductive and deductive coding and theme development. *Int. J. Qual. Methods* 5 (1), 80–92. <https://doi.org/10.1177/160940690600500107>.
- Fuhr, D.C., Acarturk, C., Sijbrandij, M., Brown, F.L., Jordans, M.J.D., Woodward, A., McGrath, M., Sondorp, E., Ventevogel, P., Ikkursun, Z., El Chammay, R., Cuijpers, P., Roberts, B., 2020. Planning the scale up of brief psychological interventions using theory of change. *BMC Health Serv. Res.* 20 (1), 801. <https://doi.org/10.1186/s12913-020-05677-6>.

- Glasgow, R.E., Vogt, T.M., Boles, S.M., 1999. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. *Am. J. Publ. Health* 89 (9), 1322–1327. <https://doi.org/10.2105/ajph.89.9.1322>.
- Grant, A., Treweek, S., Dreischulte, T., Foy, R., Guthrie, B., 2013. Process evaluations for cluster-randomised trials of complex interventions: a proposed framework for design and reporting. *Trials* 14 (1), 15. <https://doi.org/10.1186/1745-6215-14-15>.
- Greene, M.C., Scognamiglio, T., Likindikoki, S.L., Misinzo, L., Njau, T., Bonz, A., Ventevogel, P., Mbwambo, J.K.K., Tol, W.A., 2022. Examining implementation of an intervention to reduce psychological distress and intimate partner violence in a refugee camp setting. *Global Publ. Health* 1–15. <https://doi.org/10.1080/17441692.2022.2029926>.
- Haroz, E.E., Nguyen, A.J., Lee, C.I., Tol, W.A., Fine, S.L., Bolton, P., 2020. What works in psychosocial programming in humanitarian contexts in low- and middle-income countries: a systematic review of the evidence. *Intervention* 18 (1), 3. https://doi.org/10.4103/INTV.INTV_6_19.
- Harper Shehadeh, M.J., Abi Ramia, J., Cuijpers, P., El Chammay, R., Heim, E., Kheir, W., Saeed, K., van Ommeren, M., van't Hof, E., Watts, S., Wenger, A., Zoghbi, E., Carswell, K., 2020. Step-by-Step, an e-mental health intervention for depression: a mixed methods pilot study from Lebanon. *Front. Psychiatr.* 10, 986. <https://doi.org/10.3389/fpsy.2019.00986>.
- Hennink, M., Kaiser, B.N., 2022. Sample sizes for saturation in qualitative research: a systematic review of empirical tests. *Soc. Sci. Med.* 292, 114523. <https://doi.org/10.1016/j.socscimed.2021.114523>.
- Horwitz, R.I., Singer, B., 2018. Introduction. What works? And for whom? *Soc. Sci. Med.* 210, 22–25. <https://doi.org/10.1016/j.socscimed.2018.05.013>.
- Inter-Agency Standing Committee, 2007. IASC Guidelines on Mental Health and Psychosocial Support in Emergency Settings. Inter-Agency Standing Committee. <https://doi.org/10.1037/e518422011-002>.
- Jones, L., Ventevogel, P., 2021. From exception to the norm: how mental health interventions have become part and parcel of the humanitarian response. *World Psychiatr.* 20 (1), 2–3. <https://doi.org/10.1002/wps.20808>.
- Kendall, J.M., 2003. Designing a research project: randomised controlled trials and their principles. *Emerg. Med. J.* 20 (2), 164–168. <https://doi.org/10.1136/emj.20.2.164>.
- Kienzler, H., 2019. Mental health system reform in contexts of humanitarian emergencies: toward a theory of 'practice-based evidence. *Cult. Med. Psychiatr.* 43 (4), 636–662. <https://doi.org/10.1007/s11013-019-09641-w>.
- Kohrt, B.A., Burkey, M., Stuart, E.A., Koirala, S., 2015. Alternative approaches for studying humanitarian interventions: propensity score methods to evaluate reintegration packages impact on depression, PTSD, and function impairment among child soldiers in Nepal. *Global Mental Health* 2, e16. <https://doi.org/10.1017/gmh.2015.13>.
- May, C., 2006. Mobilising modern facts: health technology assessment and the politics of evidence. *Sociol. Health Illness* 28 (5), 513–532. <https://doi.org/10.1111/j.1467-9566.2006.00505.x>.
- May, C., Ellis, N.T., 2001. When protocols fail: technical evaluation, biomedical knowledge, and the social production of 'facts' about a telemedicine clinic. *Soc. Sci. Med.* 53 (8), 989–1002. [https://doi.org/10.1016/s0277-9536\(00\)00394-4](https://doi.org/10.1016/s0277-9536(00)00394-4).
- May, C., Johnson, M., Finch, T., 2016. Implementation, context and complexity. *Implement. Sci.* 11 (1), 141. <https://doi.org/10.1186/s13012-016-0506-3>.
- May, C., Mort, M., Williams, T., Mair, F., Gask, L., 2003. Health technology assessment in its local contexts: studies of telehealthcare. *Soc. Sci. Med.* 57 (4), 697–710. [https://doi.org/10.1016/s0277-9536\(02\)00419-7](https://doi.org/10.1016/s0277-9536(02)00419-7).
- McGill, E., Marks, D., Er, V., Penney, T., Petticrew, M., Egan, M., 2020. Qualitative process evaluation from a complex systems perspective: a systematic review and framework for public health evaluators. *PLoS Med.* 17 (11), e1003368. <https://doi.org/10.1371/journal.pmed.1003368>.
- Mistry, A.S., Kohrt, B.A., Beecroft, B., Anand, N., Nuwayhid, I., 2021. Introduction to collection: confronting the challenges of health research in humanitarian crises. *Conflict Health* 15 (1), 38. <https://doi.org/10.1186/s13031-021-00371-8>.
- 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. *BMJ* 350, h1258. <https://doi.org/10.1136/bmj.h1258>.
- Morgan-Trimmer, S., 2015. Improving process evaluations of health behavior interventions: learning from the social sciences. *Eval. Health Prof.* 38 (3), 295–314. <https://doi.org/10.1177/0163278713497363>.
- Mowat, R., Subramanian, S.V., Kawachi, I., 2018. Randomized controlled trials and evidence-based policy: a multidisciplinary dialogue. *Soc. Sci. Med.* 210, 1. <https://doi.org/10.1016/j.socscimed.2018.05.003>.
- Munodawafa, M., Mall, S., Lund, C., Schneider, M., 2018. Process evaluations of task sharing interventions for perinatal depression in low and middle income countries (LMIC): a systematic review and qualitative meta-synthesis. *BMC Health Serv. Res.* 18 (1), 205. <https://doi.org/10.1186/s12913-018-3030-0>.
- Oakley, A., Strange, V., Bonell, C., Allen, E., Stephenson, J., RPPLE Study Team, 2006. Process evaluation in randomised controlled trials of complex interventions. *Br. Med. J.* 332 (7538), 413–416. <https://doi.org/10.1136/bmj.332.7538.413>.
- Panther-Brick, C., Eggerman, M., Ager, A., Hadfield, K., Dajani, R., 2020. Measuring the psychosocial, biological, and cognitive signatures of profound stress in humanitarian settings: impacts, challenges, and strategies in the field. *Conflict Health* 14 (1), 40. <https://doi.org/10.1186/s13031-020-00286-w>.
- Perera, C., Salamanca-Sanabria, A., Caballero-Bernal, J., Feldman, L., Hansen, M., Bird, M., Hansen, P., Dinesen, C., Wiedemann, N., Vallières, F., 2020. No implementation without cultural adaptation: a process for culturally adapting low-intensity psychological interventions in humanitarian settings. *Conflict Health* 14 (1), 46. <https://doi.org/10.1186/s13031-020-00290-0>.
- Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., Griffey, R., Hensley, M., 2011. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Admin. Pol. Mental Health* 38 (2), 65–76. <https://doi.org/10.1007/s10488-010-0319-7>.
- Purgato, M., Gastaldon, C., Papola, D., van Ommeren, M., Barbui, C., Tol, W.A., 2018. Psychological therapies for the treatment of mental disorders in low- and middle-income countries affected by humanitarian crises. *Cochrane Database Syst. Rev.* 7, CD011849. <https://doi.org/10.1002/14651858.CD011849.pub2>.
- Rhodes, T., Lancaster, K., 2019. Evidence-making interventions in health: a conceptual framing. *Soc. Sci. Med.* 238, 112488. <https://doi.org/10.1016/j.socscimed.2019.112488>.
- Ryan, G.K., Bauer, A., Endale, T., Qureshi, O., Doukani, A., Cerga-Pashoja, A., Brar, S.K., Eaton, J., Bass, J.K., 2021. Lay-delivered talk therapies for adults affected by humanitarian crises in low- and middle-income countries. *Conflict Health* 15 (1), 30. <https://doi.org/10.1186/s13031-021-00363-8>.
- Scott, S.D., Rotter, T., Flynn, R., Brooks, H.M., Plesuk, T., Bannar-Martin, K.H., Chambers, T., Hartling, L., 2019. Systematic review of the use of process evaluations in knowledge translation research. *Syst. Rev.* 8 (1), 266. <https://doi.org/10.1186/s13643-019-1161-y>.
- Shahabuddin, A.S.M., B Sharkey, A., Jackson, D., Rutter, P., Hasman, A., Sarker, M., 2020. Carrying out embedded implementation research in humanitarian settings: a qualitative study in Cox's Bazar, Bangladesh. *PLoS Med.* 17 (7), e1003148. <https://doi.org/10.1371/journal.pmed.1003148>.
- Skivington, K., Matthews, L., Simpson, S.A., Craig, P., Baird, J., Blazeby, J.M., Boyd, K.A., Craig, N., French, D.P., McIntosh, E., Petticrew, M., Rycroft-Malone, J., White, M., Moore, L., 2021. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. *BMJ* 374, n2061. <https://doi.org/10.1136/bmj.n2061>.
- Timmermans, S., Berg, M., 2010. *The Gold Standard: the Challenge of Evidence-Based Medicine*. Temple University Press.
- Tol, W.A., Ager, A., Bizouerne, C., Bryant, R., El Chammay, R., Colebunders, R., García-Moreno, C., Hamdani, S.U., James, L.E., Jansen, S.C.J., Leku, M.R., Likindikoki, S., Panther-Brick, C., Pluess, M., Robinson, C., Rutenberg, L., Savage, K., Welton-Mitchell, C., Hall, B.J., van Ommeren, M., 2020. Improving mental health and psychosocial wellbeing in humanitarian settings: reflections on research funded through R2HC. *Conflict Health* 14 (1), 71. <https://doi.org/10.1186/s13031-020-00317-6>.
- Uygun, E., Ilkkursun, Z., Sijbrandij, M., Aker, A.T., Bryant, R., Cuijpers, P., Fuhr, D.C., de Graaff, A.M., de Jong, J., McDaid, D., Morina, N., Park, A.-L., Roberts, B., Ventevogel, P., Yurtbakan, T., Acarturk, C., STRENGTHS consortium., 2020. Protocol for a randomized controlled trial: peer-to-peer group problem management Plus (PM+) for adult Syrian refugees in Turkey. *Trials* 21 (1), 283. <https://doi.org/10.1186/s13063-020-4166-x>.
- van't Hof, E., Sangraula, M., Luitel, N.P., Turner, E.L., Marahatta, K., van Ommeren, M., Shrestha, P., Bryant, R., Kohrt, B.A., Jordans, M.J.D., 2020. Effectiveness of Group Problem Management Plus (Group-PM+) for adults affected by humanitarian crises in Nepal: study protocol for a cluster randomized controlled trial. *Trials* 21 (1), 343. <https://doi.org/10.1186/s13063-020-04263-9>.
- Wells, M., Williams, B., Treweek, S., Coyle, J., Taylor, J., 2012. Intervention description is not enough: evidence from an in-depth multiple case study on the untold role and impact of context in randomised controlled trials of seven complex interventions. *Trials* 13, 95. <https://doi.org/10.1186/1745-6215-13-95>.
- Will, C., Moreira, T., 2016. *Medical Proofs, Social Experiments: Clinical Trials in Shifting Contexts*. Routledge.