Co-designed or evidence-based? Developing digital self-management interventions for long-term conditions. McCallum, Campbell, Hackett, Vines

ABSTRACT

Self-management interventions for long-term conditions may improve quality of life. Psychosocial techniques can be delivered via digital technologies such as smartphone applications. Interventions should be useful and meaningful to users, but also theory- and evidence-based. This chapter explores some of the processes and challenges involved in developing self-management interventions using co-design, theory and evidence. Drawing on app development research focused on supporting those with Sjögren's syndrome (an autoimmune and rheumatic disease with complex symptom patterns), the chapter discusses the practical difficulties we had in generating and respecting individuals' ideas, preferences, requirements and creativity, while acknowledging the need for an intervention to include established cognitive components, behaviour change techniques, and "active ingredients". The chapter is organised around three tensions from our project where our ambition to co-design the intervention with participants was challenged and contradicted. In such situations, we often found ourselves having to judge the relevancy or compatibility of participants' viewpoints and lived experiences against the evidence base. We close with critical reflections on these tensions, discussing the potential limits of co-design and user-led approaches to creating health and care interventions in a disciplinary culture of evidence-based practice. We offer a series of implications for conducting design research in the area of self-managing long-term conditions, identifying ways for the competing paradigms of co-design and evidence-based design to be bridged.

Introduction

Intervention developers are encouraged to use evidence to decide and justify their inclusion of specific therapeutic content for health interventions. The Medical Research Council (MRC) guidance states that "Best practice is to develop interventions systematically, using the best available evidence and appropriate theory" (p.2) [Craig et al., 2008]. What constitutes evidence is not explicitly stated, but generally refers to previous studies evaluating the effectiveness of an intervention in changing behaviours or health outcomes. The guidance to use the 'best available' evidence also highlights its hierarchical nature, where high quality evidence from many randomised control trials (RCTs) should be prioritised over non-randomised trials or expert opinion when making decisions [Evans 2003]. Interventions are more likely to be effective if their therapeutic content is theory-based [Craig et al 2008]. Behaviour change techniques (BCTs) are an example of theory-based intervention approaches. These techniques or "active ingredients" of interventions have been theorised to influence key drivers of behaviour change [Michie et al 2011].

As well as being evidence- and theory-based, interventions should be developed with potential future users. MRC guidelines suggest that following initial intervention development using evidence and theory, the intervention should be assessed for acceptability and feasibility [Craig et al 2008], to ensure it is likely to be used. Yet, there is growing recognition that health interventions should involve potential users in intervention decision-making from the outset [Yardley et al 2015][Michie et al 2017]. A 'user-centred' intervention will typically involve researchers speaking to participants to understand their specific needs, building a product designed to suit those needs, and perhaps going back to them to check the acceptability or make minor changes (i.e. user-centred design). A step beyond this is co-design, where participants are involved throughout the process, and crucially,

are given genuine decision-making powers (Sanders & Stappers, 2008). In the context of developing health interventions, this might mean codesigners contribute to decisions regarding therapeutic content (e.g. which behaviour change techniques to include) and/or its delivery (e.g. digitally via a smartphone, choosing various displays on-screen, and/or by a human healthcare professional).

The use of co-design approaches to develop interventions is not new, yet how to combine these with evidence-based and theory-based approaches is not clear. Some recent papers on digital intervention development do describe step-by-step methodologies that integrate these approaches, which appear planned and well-knitted together [Curtis et al 2015, Korpershoek et al 2020, O'Brien et al 2016]. However what is missing is an account of the contradictions and difficulties experienced in bringing together user-led and theory- and evidence-based approaches in practice. In this paper, we endeavour to be highly reflexive: when creating a self-management intervention for a complex long-term condition, we found it extremely challenging to base intervention development decisions on evidence, theory, and co-design findings. We draw upon a project undertaken between 2019-2021 funded by Versus Arthritis (22026). This aimed towards developing and evaluating a smartphone app that supports those living with Sjogren's syndrome (SS), a long-term rheumatic autoimmune condition, to self-manage their key symptoms (dryness, sleep disturbance, pain, and fatigue [Mariette & Criswell 2018, Lewis et al 2019]).

Our app development methodology involved two parallel workstreams. We used co-design methods to understand the experiences and perspectives of those with SS, including how they experience and self-manage this complex condition, any associated challenges they face, and their needs. This was to ensure any app developed would be useful, targeted these needs, and supported existing self-management activities. This workstream involved facilitating a series of co-design workshops with those living with SS, over several months, comprising various design activities: developing magic machines [Andersen & Wakkary 2019]), sketching metaphors of symptoms and self-management experiences, prototyping an application interface, and focussed group discussions.

In parallel with the co-design workshops, the research team conducted 'desk research', which had a different objective: to identify evidence on the effectiveness of existing relevant interventions, and theories of behaviour change, to inform the intervention therapeutic content. In practice this involved literature searches of trials evaluating intervention effectiveness, and reviewing manuals of interventions found to be effective. As SS is an understudied condition, experimental evidence was sparse and often unavailable. On occasion, qualitative studies (e.g. of how people with SS experience their symptoms or perceive therapeutic approaches) were considered to be the 'best available' evidence.

The research team felt that a co-design approach, alongside pulling together existing research, was essential: not only because published evidence was sparse, but because individuals with SS have expertise that we did not in self-managing this complex condition, and a wealth of knowledge of 'what works' for them (and importantly 'when', given the condition's long-term nature). Yet, our research team encountered several instances where bringing co-design principles together with the use of evidence and theory appeared to be paradoxical.

This chapter is organised around three tensions from our project where our ambition to co-design the intervention with participants was challenged and contradicted. In such situations, we found ourselves having to judge the relevancy or compatibility of participants' viewpoints and lived experiences against the evidence base. We will report tensions surrounding: : ii) the input of co-designers and existing behavioural evidence and theory, ii) how co-designers' views contradicted

prior qualitative research on user needs and preferences iii) how the research team had to carefully reduce the risk of "overfitting" the design to co-designers' experiences and preferences, and judge the relevancy of these for a typical (and fictional) user with Sjögren's syndrome. In several cases, co-designers ideas or experiences were discounted.

This research was funded by Versus Arthritis [grant number 22026].

Tensions we experienced

Tension 1. Co-designers' self-management practices and desired therapeutic content was not supported by theory or evidence

Our experience

A key incentive for this work was the lack of psychosocial and behavioural interventions for people with SS [Hackett 2015]. An earlier phase of development had established that due to the complexity of the condition and scattered service provision, people were having to act as their own case managers. Participants of a qualitative study had identified which intervention targets and outcomes were important to them [Hackett at al 2018]. Our approach to being theory- and evidence-based was to bring together and digitalise the content of multiple existing interventions targeting these symptoms, which clinical trials had evidenced to be effective (such as cognitive behavioural therapy for insomnia to aid sleep disruptions, and energy management techniques such as pacing to reduce rheumatic fatigue and pain).

Our objective was not to generate new ideas for therapeutic content and active ingredients "from scratch" with participants (an approach used elsewhere [Curtis et al 2015]), but to deliver existing and evidence-based intervention content in a way which supported self-management and fitted into participants' lives. Nevertheless, we did seek to understand the self-management practices and techniques that co-designers used. This was so the app could support rather than disregard these, and also incorporate them as tips; we believed other people with SS could learn from their expertise in managing their own condition. However, given we required all content to be 'evidence-based', and we planned to prioritise and digitalise evidence-based behavioural techniques, we soon realised that we were constrained in the therapeutic content we could include. In hindsight, codesigners were always restricted to making decisions on how content should be delivered.

Unsurprisingly, this tension was evident during interactions with participants. Despite wanting to incorporate into the app a few self-management techniques that codesigners' used, some told us about techniques and views that were directly counter-indicative to evidence. For example, some napped during the day to manage fatigue (which can make sleep disruptions worse), and others listened to music to try and sleep (which constitutes sleep effort and should not normally be engaged in). Many spoke about living in a state of opposition to their illness, and described "fighting it" and pushing through it, rather than using recommended pacing techniques to manage and pace themselves through the day.

We encountered a similar challenge regarding the therapeutic content of the app when, unsolicited by the research team, participants explicitly mentioned therapeutic techniques and features they

wanted in the app. For instance, a participant suggested that she would like to see a video demonstrating how to massage her eyes to stimulate tears:

> *I've never been shown how to massage my eyes and what I should be doing. So, on the app it would be lovely to have like a little video to show how to do it*

This request had the potential to be theory-based: it mapped onto an existing BCT: "6.1 Demonstration of the behaviour" (BCTT). However, such ideas had not been explored for their effectiveness in any previous trials for Sjogren's syndrome or similar conditions, and were not included in any of the interventions we had decided to digitalise. Furthermore, it was a costly feature, which required additional resources for recording and hosting. Ultimately, the mere potential of untrialled ideas being effective was just too small of a promise when it was resource intensive to implement; not least test.

Our approach

In both scenarios and throughout the process of development we found ourselves needing to think about therapeutic content as something separate from its delivery method (the smartphone app). Although we were interested in the self-management experiences and techniques codesigners' used, and even in codesigners' unsolicited ideas for therapeutic content: when evidence did not support them, we to an extent discarded them, or at least considered them low priority. However, we did attempt to sensitise our fixed plans to these views held, behaviours engaged in, and content requested by codesigners. We saw these as useful insights into what future users might think and do when arriving at our app (or continue doing despite our app suggesting otherwise). Our solution was sometimes to call for user discretion, where for example the app would read: "if you need to nap, keep it short". We operated from a place that was not eager to take away agency. We also did keep a record of feasible ideas and suggestions for future app development and testing. In hindsight, and to facilitate more shared decision making, we did need a mechanism to discuss and inform codesigners' about which of their ideas did not make the cut.

Were we really conducting co-design? Perhaps not so much at this point, but discarding participant suggestions was not something we did lightly or carelessly. It is fair to point out that if we consider the intervention process from its very inception, the co-design balance looks a lot more attainable. In the first phases of development participants had a say over the very fundamentals of the intervention; i.e. the intervention targets. The researchers stepped up their input in the second phase on techniques to improve these symptom targets. Yet, it is not at all unfair to say that to a certain extent we, the researchers, saw the task at hand to be improving the acceptability of how intervention techniques are delivered. Was this a case of passive paternalism? (i.e. "We experts know what's good for you – we just need to find a way to make you accept it"). Perhaps. But there is only finite room for compromise when we talk about the evidence-based therapeutic techniques or theoretical active ingredients of any given treatment.

Tension 2: Co-designers' perspectives on therapeutic content and delivery differed to those previously published

Our experience

In addition to co-designers requesting and suggesting therapeutic content we did not ultimately include, they also challenged us on the therapeutic content we did plan to include and our intentions for its delivery. As researchers, we did approach the co-design workshops with our own views, expectations and agendas. These agendas were mostly informed by research previously gathered to motivate the study. However, as well as arriving with an awareness of the evidence-base, we had funding to complete a project that had to go in a particular direction.

Planned therapeutic content for the intervention was based on systematic reviews and RCTs of intervention effectiveness where possible, but we were also focus groups on the acceptability of these, conducted by a member of the research team (Hackett et al., 2017; Hackett et al., 2018; Lewis et al., 2019). However at times co-designers' input contradicted or differed to previously published qualitative studies surrounding the experiences, perspectives and preferences in those with SS. For instance few participants were not fully accepting of the potential inclusion of CBT in the app:

> *I don't personally think it is for me … I can't see how that would stop me being woken up by the things that wake me up, that's the issue because that is the underlying thing. McKenna said the same in his book. I didn't find that it worked for me.*

Being regarded as acceptable by a previous group with SS did not automatically mean CBT-I was palatable to our codesigners. This same issue arose in the core delivery method of a smartphone app: previous qualitative gathered by a member of the research team suggested people with SS wanted a menu driven application and an accessible digital tool (Hackett dissertation). These prior participants had felt this could be something potentially accessible to many, and especially useful for those unable to access psychosocial and educational support. These discussions formed part of the rationale for our funding application to develop a digital self-management tool, where we then worked out the app development costs and timescales. This meant that if differing ideas came out of the workshops, we had limited ourselves in fully taking these on board. Therefore, when one or two participants in this current study seemed to suggest they didn't want a smartphone, a tension arose. One female workshop participant said:

> *My daughter persuaded me to accept her partner's old phone which I've got here and it's got apps on it and I'm thinking, what are you trying to do to me? … The other one I had was really cheap. ... All it does is send texts and makes calls. That's all. I didn't want any, all of this argy bargy.*

The intervention delivery method of a digital technology, which was suggested by previous participants with SS, was not agreed on by all codesigners in the present project. Despite being invited to take part in workshops to co-create a smartphone app for people with SS, there were one or two participants who either didn't currently own a smartphone, or despite having recently acquired one, the thought of using apps on one seemed to be too new and overwhelming for them to consider.

Our approach

In both of the above situations where codesigners' views and experiences contrasted the evidence used to motivate the study, we chose to continue on the same path we had set foot on. This may seem like we discounted participants' concerns without thought. On the contrary, we had several research team discussions which helped to consider our actions. For example, in the case of including CBT-I: (i) there were other co-designers who were keen for the inclusion of CBT-I (ii) this component formed only one part of the app (on a menu of options as previously suggested by people with SS), and therefore not everyone using it needed to engage with this optional section and (iii) barriers to implementing the CBT-I techniques (such as getting up in the night if unable to sleep) are to be expected, whether the intervention takes place online or in a face-to-face setting with a clinician (Yeung et al 2015).

Similar to how we dealt with Tension 1, we did not fully discard co-designers' input: we used it to sensitise our work. We realised the importance of providing a rationale for CBT-I within the app, as a therapist would when delivering similar components in a face-to-face setting. And although some members of the co-development group did not have nor intend to use a smartphone in future, we chose to continue working with them to understand their contexts and experiences (e.g. scenarios where their SS impacts their lives and in discussing metaphors to describe their experiences). Their seemingly incompatible views did not preclude them from informing the development of the app in any way; we believe they enriched it by providing more diverse experiences and perspectives.

Overall, we learned not to expect that our codesigners' views and experiences would align well with those of participants with SS in previous qualitative studies, but experienced challenges in deciding whose views to prioritise - especially given that it was the previous work that informed the funding proposal.

Tension 3: The research team judged the relevance and applicability of co-designers' experiences for future users

Our experience

We aimed to design the app for as many people with Sjogren's syndrome as possible, and ensure the content was relevant to and likely to be effective for all of them. Attempting to do this in a way that was 'evidence-based' led us to being reductionist. Despite co-design principles of inclusivity, our approach ended up being exclusive, homogenous and potentially less useful and relevant to the people it was designed with and for.

To help promote a sense of relatedness, validation and belonging for our users [as informed by Self-Determination Theory; Deci & Ryan 2008], the app included quotes and metaphors from co-designers of what it felt like to experience and self-manage Sjogren's syndrome and its often invisible symptoms such as fatigue, pain, and dryness. A key principle in co-design is to be inclusive: understand the diversity in people's needs and experiences and ensure the ultimate design accounts for and supports these [Clarkson et al., 2013]. To do this, we recruited for diversity within the target population of people with SS: specifically in age, gender, time since diagnosis, severity and across countries. To increase the applicability of our app to a range of people with Sjogren's syndrome, we recruited both those who had only Sjogren's syndrome (Primary SS) and those who had additional

comorbid conditions (secondary SS). Codesigners together had a range of comorbidities: rheumatoid arthritis, fibromyalgia, neuropathy, vasculitis, bowel problems, kidney disease, asthma, blood disorders, and cancer.

Despite gathering diverse experiences and perspectives from codesigners, our research team quickly discarded some of these in both the qualitative analysis and in the design of the app. Specifically; to try and ensure the content was relevant to everyone with Sjogren's syndrome, we felt the need to strip out details which we judged to be too specific to individual users. This was particularly evident in relation to experiences of comorbidities that impacted how codesigners self-managed their condition, such as one female with kidney disease:

> *I have no sleep because of my kidneys… every two hours I have to wake up to sort my kidneys out...I'm on a totally different planet because of the sleeplessness*

These experiences were key in this codesigners' daily self-management experiences, yet kidney disease is seen only in 4-5% of those with Sjogren's syndrome [Aiyegbusi et al 2021)]. Some codesigners reported symptoms that they were unsure whether they were connected to SS, such as sciatic nerve issues, intestinal issues and hypertension. Several were also aware of individual differences in how Sjogren's syndrome presents:

> *Everyone has different symptoms… everybody seems to have different little quirks*

> *We've got the same Sjögren's, haven't we, but it affects us in different ways… which complicates things.*

Recognising this, some codesigners showed reluctance to make design decisions on behalf of others, and felt strongly there were differences in the responsibilities between themselves as codesigners and us researchers:

> *Well that's a toughie, that's your job. It's something that I would mull over rather than immediately come up with something. This is your work and you've done lots of research... So, we'll not dwell on that because I think that is too much of a me thing.*

> *I understand how difficult it must be for yourselves because it is one thing having us bantering and talking, oh well this helps me, but as a clinical professional you've got to put it across evidence based and you cannot recommend something to somebody that [is] not factual or something that could damage someone.*

Therefore, although codesigners were motivated to get involved in the study to advocate for the needs and perspectives of people living with Sjögren's syndrome; in recognition that their needs were often very different some would hesitate when it came to shaping design decisions, instead relying on the direction of the researchers.

Our approach

Despite being important to codesigners, we struggled to decide whether to include experiences and perspectives in the app for symptoms that seemed less related to Sjogren's syndrome. During qualitative analysis and decision-making for the app, experiences we considered to be "too much another condition and not enough Sjogren's syndrome" were typically set aside. In other cases we found ourselves attempting to use disease prevalence rates to justify our decisions: as fibromyalgia is often comorbid with and presents similarly to SS [Kang & Lin, 2010], we included these

overlapping symptom experiences over comorbidities that epidemiological research suggested were rare - and thus unlikely to be relevant to several SS users.

Researchers and designers naturally must reduce scope and narrow down their target users for several reasons, such as to keep to time and funding constraints, and to avoid a product containing an overwhelming amount of content that risks being irrelevant to some people. However, attempting to homogenize our users' experiences to focus on what we as researchers believed were more core to SS was especially problematic when codesigining a chronic condition self-management app, for two reasons.

First, most people with SS have comorbidities of some form: very few people have "only Sjogren's syndrome" [Kang & Lin, 2010]. This meant we were essentially designing for a fictional user and an app unlikely to be useful for the majority of people with SS. Similar logic is found in highly controlled clinical trials which exclude participants with comorbidities. In these reductionist trials, homogeneity (i.e. highly similar participants) is prioritized to give an intervention the best chance of being evidenced as effective (as variation could make any effects of the app harder to detect). However, this ends up reducing the usefulness and applicability of the trial findings to most people outside the trial, who do have comorbidities, and for whom it remains unknown if the intervention would be effective. In the same way, when we reduce co-designers and future users to a homogenous group, the final app loses some of its usefulness to real-world users in managing the realities of their condition, beyond those considered by researchers. Overall, some of the key benefits of co-design were perhaps lost.

The second reason that homogenizing codesigners' experiences was problematic was that it was us as researchers, not them, who judged the relevancy of their experiences (i.e. whether it would be relatable for other people with Sjogren's syndrome). This issue of whether those with the condition should speculate and design for others has been briefly discussed [Yardley et al 2015], and suggested not to be the task of participants. However, this then leaves it to researchers, who are potentially even less qualified. Overall, we did not fully represent or weave in struggles surrounding co-morbidities within the narrative of the app. In future, we might discuss such scope decisions with co-designers. Ultimately we attempted to generalize and make the app more widely applicable by stripping out contextual details too specific to users, or reducing their weight when making decisions.

Discussion and conclusions

In this chapter we discussed our attempts to bring together the competing paradigms of co-design together with evidence- and theory-based intervention development. We outlined three key tensions we experienced and how we sought to respect codesigners input in the design of the app, through: calling for user discretion around techniques unsupported by evidence; sensitising and rationalising the use of core (and less acceptable) techniques; and reducing the weight of highly contextual details or unique characteristics when making design decisions. Far from proposing other intervention developers do the same, we document what happened and provide a reflexive account of moments during the project that questioned whether our approach involved co-design, despite our best intentions.

What counts as evidence and what should we prioritise when co-designing?

We often found ourselves checking codesigners' ideas, experiences and perspectives against 'evidence' if it was available, and prioritising the use of evidence to make final decisions by default. Codesigners' own self-management practices, techniques and suggestions were checked for their

compatibility with behaviour change theory and the content of existing interventions that trials suggested were effective (Tension 1). Codesigners' perspectives on therapeutic techniques were compared to what others in their population had said they wanted in previous studies (Tension 2). And finally, their experiences of symptoms and comorbid conditions were even compared to epidemiological studies (Tension 3). Crucially, this 'checking' was not systematic; it was a rough and iterative way (i.e. a quick check of the literature) to try and balance co-design with theory and evidence. Neither approach was fully committed to.

We now reflect on what we counted as "evidence". Initially we considered this to be trial data on effectiveness when available [Craig et al., 2008][Evans 2003], but in hindsight, it seems we considered evidence to simply mean research that was published or generated before the study. This was especially highlighted by the times we prioritised insight from published qualitative studies (which is low on the evidence hierarchy [Evans 2003]) over new qualitative data acquired from codesigners (as in Tension 2). In short, we clung to the concept of an 'evidence base' even when it was potentially inappropriate or produced an app potentially less relevant to our co-designers.

At what point does new data gathered with codesigners' become "evidence"? Should we automatically consider it as evidence, given that it was reported by an expert by experience? If so, then what do we prioritise when making design decisions? Guidance for comparing qualitative research in evidence-based practice [Daley et al 2007] suggests we should prioritise whichever qualitative data has been more thoroughly analysed. Yet comprehensive analysis methods may be challenging to achieve in ways that keep up with the rapid pace of app development, meaning codesign findings may never be prioritised to inform design decisions, even over qualitative studies.

An alternative avenue to prioritising analytical quality [Daly et al., 2007] or study design [Evans 2003] would be to orient design decisions around data "relevance"' (i.e. whether from prior studies or current codesigners). The concept of relevance aligns with proposals that evidence from N-of-1 trials should be prioritised over typical RCTs when making treatment decisions for individuals [Vohra et al., 2015]. Yet in our study, as we saw in tension 3, we tried to design for any smartphone user with SS (through insights and input from a small pool of codesigners). This is when using evidence from large-scale studies to hold more weight in app development decisions felt, and still feels, more appropriate than one or two individuals' experiences. The problem with relevance is that it requires knowing the characteristics of the final target users: when creating something only for codesigners and people exactly like them, we might prioritise data relevance over high quality evidence. But when generating an app for many people to use (such as in a large-scale clinical trial of people with SS), we might be more wary of "overfitting" the app design to only codesigners' needs and experiences. Instead it may be appropriate to prioritise high quality evidence from large-scale studies, as we did.

There are also ethical considerations regarding whether to prioritise high quality evidence or codesigners' input. During workshops, codesigners offered their time, personal experiences, creativity, and ideas to help to make decisions. Although such input was rarely fully discarded (and often used to at least sensitise or align therapeutic content), the amount of energy required to generate these especially for those with fatigue may not have been commensurate with what actually made it to the app, for either codesigners or other people with SS to use. From another perspective, it can of course be more ethical to lean on the evidence-base. Learning from prior studies and using evidence and theory to develop interventions can give them the best chance of being effective [Craig et al., 2008] and as such, more worthwhile for patients to engage in, rather than unnecessarily burdensome. We also saw how (in tension 3) some codesigners might be

reluctant to engage in codesign activities, and perhaps should not be expected to consider what other people 'like them' might find beneficial [Yardley et al 2013].

In hindsight, and with a more honest picture of what our own objectives were surrounding the use of codesign methods (i.e. to inform the delivery of therapeutic content, rather than the content itself), we could have restricted codesign workshop activities accordingly, or made plans for balancing codesigners' input (even if unsolicited) against the evidence base and communicating these. Having learned from our experiences, we might overall set clear expectations for codesigners and potentially refrain from considering work to be 'codesign', unless there is substantial ability for codesigners' to help decide therapeutic content (e.g. techniques): not just how it will be delivered.

Who knows best? Power and paternalism

We found ourselves making all the ultimate decisions despite using co-design methods: judging the relevancy of co-designers experiences to those of our future users and deciding 'what goes in' to the app or 'makes the cut', accordingly. It is important to note where codesign deviates from qualitative research in its deliverables and ultimate goals. Crudely, research is to advance knowledge, and design is to develop something applied for a particular purpose [Gaver, 2012]. In designing and developing an intervention we ultimately need to design a 'thing' and therefore make design decisions; not endlessly ponder or conclude about our confidence in advancing knowledge and leave it at that. This decision making opens up the process of design to power dynamics. As we saw from tension 3, codesigners were very aware of this.

In most projects, researchers (answerable to funders, the academic community and public tax-payers) are likely to hold the final say over what makes it in the app and what is discarded. Perhaps what is needed is methodological innovation where co-designers can go beyond providing disparate insights of their experiences to generating consensus without researcher input. Alternatively it could be the order in which we do things. Rather than co-designing "up to a point" after which the research is sealed off and turned into published evidence, and the final product is developed; codesigners could have the final decision of what makes it into an app after seeing both evidence and experiences, and approve it before and after its evaluation. Currently, the ultimate way co-designers and target users have their final say is by discarding the developed intervention, or parts of it deemed irrelevant. Although potentially a waste of resources, co-designers who cannot ever hope to force researchers to include therapeutic content they do not want to, still hold the ultimate power through their ability to either shut down the success of an app, or more optimistically, take what they need from it.

Conclusions: how do we move forward?

Our aims to combine co-design and theory- and evidence-based approaches meant we were tackling several known differences in the paradigms underpinning theory versus contextual work and generalisability in design [Höök & Löwgren, 2012][Stolterman 2008]. Applying theory and evidence is inherently a 'top down' deductive approach to analysis as opposed to inductive methods used in co-design. Theory is abstract and reductionist by its very nature, often aiming to strip out contextual details, to ensure it is as widely applicable as possible. This contrasts with more inclusive design approaches and those that detail context to provide others with an understanding of how well the findings could transfer to other contexts and people.

One approach we suggested to bridge co-design and evidence- and theory-based approaches in intervention development is to restrict co-design activities to intervention delivery, not therapeutic

content. Here it would be important to establish these constraints early and set expectations, so as not to waste the energy of co-designers (especially those experiencing fatigue). Yet, this essentially drifts from co-design to become user-centred design, where we seek users' opinions on established components rather than give them decision-making powers. With this perspective, the aim would be to guard yet sensitise our therapeutic content (e.g. as we did; by providing rationales for incorporating particular theory- and evidence-based techniques (tension 2), or asking users to ensure naps are short if they must take them at all (tension 1)). After all, co-design is unlikely to ever be a 50-50 transaction between researchers and potential users. Co-design could be seen as researchers being compassionate and listening to what experts by experience have to bring to the table, but without placing undue burden on them and without ignoring our own training.

Alternatively, we consider co-designers' ideas for therapeutic content (solicited or otherwise) and explore the extent to which they are supported by the current evidence base. While there is perhaps a stronger case for discarding ideas for BCTs or self-management practices which evidence suggests could be harmful or counter-productive (such as napping): in cases where co-designers and researchers identify a gap in evidence, the aim would be to actively pause, prioritise and pivot to genuinely explore these ideas, rather than continue to force existing evidence onto co-designers. Only then can we truly co-design therapeutic content: when we have the resources to take forward ideas such as BCTs, or potential correlations suggested by co-designers, and test them. Designs such as N-of-1s [Vohra et al 2015] can do this in a way that does not require large-scale RCTs, but rather, facilitate quick and lean experimentation tailored to individuals (e.g. with comorbidities), and then resume the development process. These might be able to bypass clashes in paradigms between the need for large numbers of participants to establish some form of universal ground truth - to truly respecting that a self-management practice may work for an individual given their characteristics (and potentially others who share them).

It has been recognised that true co-design is often difficult to achieve and often falls short [Vines et al 2013; Green et al 2020]. This chapter provides insight into the detailed practical decisions required or types of input co-designers can realistically give - even in a domain of self-management that lends itself to users as experts in managing their long term condition.

References

Andersen, K., & Wakkary, R. (2019). The magic machine workshops: making personal design knowledge. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1-13).

Aiyegbusi, O., McGregor, L., McGeoch, L. *et al.* Renal Disease in Primary Sjögren's Syndrome. *Rheumatol Ther* 8, 63–80 (2021)

Clarkson, P. J., Coleman, R., Keates, S., & Lebbon, C. (2003). Inclusive design: Design for the whole population. Springer, London.

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*.

Curtis, K. E., Lahiri, S., & Brown, K. E. (2015). Targeting parents for childhood weight management: development of a theory-driven and user-centered healthy eating app. *JMIR mHealth and uHealth*, *3*(2), e3857.

Daly et al. (2007). A hierarchy of evidence for assessing qualitative health research, Journal of Clinical Epidemiology, 60 (1). 43-49.

Deci, E. L., & Ryan, R. M. (2008). Self-determination theory: A macrotheory of human motivation, development, and health. *Canadian psychology/Psychologie canadienne*, *49*(3), 182.

Evans, D. (2003). Hierarchy of evidence: a framework for ranking evidence evaluating healthcare interventions. *Journal of clinical nursing*, *12*(1), 77-84

Gaver, W. (2012). What should we expect from research through design?. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 937-946).

Green T, Bonner A, Teleni L, et al (2020). Use and reporting of experience-based codesign studies in the healthcare setting: a systematic reviewBMJ Quality & Safety;29:64-76

Hackett, K. L., Deane, K. H., Newton, J. L., Deary, V., Bowman, S. J., Rapley, T., ... & MacKay, K. (2018). Mixed-methods study identifying key intervention targets to improve participation in daily living activities in primary Sjögren's syndrome patients. *Arthritis care & research*, *70*(7), 1064-1073.

Hackett KL, Deary V, Deane KH, et al. (2018) Experience of sleep disruption in primary Sjogren's syndrome: A focus group study. *Br J Occup Ther* 81: 218-226.

Hackett KL, Gotts ZM, Ellis J, et al. (2017) An investigation into the prevalence of sleep disturbances in primary Sjogren's syndrome: a systematic review of the literature. *Rheumatology (Oxford)* 56: 570-580.

Hackett, K. L., Deane, K. H., Strassheim, V., Deary, V., Rapley, T., Newton, J. L., & Ng, W. F. (2015). A systematic review of non-pharmacological interventions for primary Sjögren's syndrome. *Rheumatology*, *54*(11), 2025-2032

Höök, K., & Löwgren, J. (2012). Strong concepts: Intermediate-level knowledge in interaction design research. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *19*(3), 1-18.

Kang JH, Lin HC. Comorbidities in patients with primary Sjogren's syndrome: a registry-based case-control study. J Rheumatol. 2010 Jun;37(6):1188-94.

Korpershoek, Y. J., Hermsen, S., Schoonhoven, L., Schuurmans, M. J., & Trappenburg, J. C. (2020). User-centered design of a mobile health intervention to enhance exacerbation-related self-management in patients with chronic obstructive pulmonary disease (copilot): mixed methods study. *Journal of medical Internet research*, *22*(6), e15449.

Lewis I, Hackett KL, Ng WF, et al. (2019) A two-phase cohort study of the sleep phenotype within primary Sjogren's syndrome and its clinical correlates. *Clin Exp Rheumatol* 37 Suppl 118: 78-82.

Mariette X and Criswell LA. (2018) Primary Sjogren's Syndrome. N Engl J Med 379: 97.

Michie, S., Van Stralen, M. M., & West, R. (2011). The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implementation science*, *6*(1), 1-12.

Michie, S., Yardley, L., West, R., Patrick, K., & Greaves, F. (2017). Developing and evaluating digital interventions to promote behavior change in health and health care: recommendations resulting from an international workshop. *Journal of medical Internet research*, *19*(6), e7126.

Morrison, L. G. (2015). Theory-based strategies for enhancing the impact and usage of digital health behaviour change interventions: A review. *Digital Health*, *1*, 2055207615595335.

O'Brien, N., Heaven, B., Teal, G., Evans, E. H., Cleland, C., Moffatt, S., ... & Moynihan, P. (2016). Integrating evidence from systematic reviews, qualitative research, and expert knowledge using co-design techniques to develop a web-based intervention for people in the retirement transition. *Journal of medical Internet research*, *18*(8), e5790.

Sanders, E. B. N., & Stappers, P. J. (2008). Co-creation and the new landscapes of design. *Co-design*, *4*(1), 5-18.

Stolterman, E. (2008). The nature of design practice and implications for interaction design research. *International Journal of Design*, *2*(1).

Vines, J., Clarke,R, Wright, P, McCarthy, J and Olivier, P. 2013. Configuring participation: on how we involve people in design. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13).

Vohra, S., Shamseer, L., Sampson, M., Bukutu, C., Schmid, C. H., Tate, R., ... & Moher, D. (2015). CONSORT extension for reporting N-of-1 trials (CENT) 2015 Statement. *BMJ*, *350*.

Yardley, L., Morrison, L., Bradbury, K., & Muller, I. (2015). The person-based approach to intervention development: application to digital health-related behavior change interventions. *Journal of Medical Internet research*, *17*(1), e4055.