RANDOMIZED CONTROLLED TRIALS IN WOHP INTERVENTIONS: A REVIEW AND GUIDELINES FOR USE

Deirdre O’Shea*, Brenda H. O’Connellb & Stephen Gallagherb

*Kemmy Business School, University of Limerick, Ireland

bCentre for Social Issues Research, Department of Psychology, University of Limerick, Ireland

in press

Applied Psychology: An International Review

*Address for correspondence: Dr. Deirdre O’Shea, Department of Personnel & Employment Relations, Kemmy Business School, University of Limerick, Limerick, Ireland. Email: deirdre.oshea@ul.ie

Acknowledgements:

We would like to acknowledge funding support from the Irish Research Council.

Please cite this article as:

Randomized controlled trials (RCTs) have long been considered a gold standard for intervention design and the most rigorous method for understanding causal mechanisms. However, their implementation in work and organizational health psychology (WOHP) can be challenging. We review the use of RCTs in WOHP interventions and demonstrate that their adoption has been relatively small in comparison to areas such as health psychology and medical sciences. For WOHP researchers to be able to compare the effectiveness of their work-specific health interventions to other interventions in health and medicine, it is important that the design methodology, rigor and reporting are comparable. Thus, there is a need for a clearer road map and guidance for WOHP researchers to encourage higher use of RCTs in WOHP intervention research. In the paper, we provide an overview of RCTs, and review past research that has utilized an RCT design when evaluating WOHP interventions. We develop an adapted RCT checklist for use in WOHP settings, which takes specific organizational issues into account. Thus, our paper provides a clearer road map for the design and reporting of WOHP RCT studies for future researchers.

**Keywords:** Interventions, occupational health psychology, randomized controlled trials, RCTs, research method, work and organizational health psychology.
The field of work and organizational health psychology (WOHP) is an area of both academic and practical value, shedding light on how to improve the experience of work for all who are employed. Recently, the bridge between research and practice has come into focus, leading to an increased consideration of WOHP interventions. Psychological interventions focusing on worker well-being can be defined as planned, behavioral, theory-based actions that aim to improve employee health and well-being (Nielsen & Abildgaard, 2013; Nielsen, Randall, Holten, & Gonzalez, 2010a), and are considered a type of quasi-experiment (Grant & Wall, 2009). We define WOHP interventions as a specific type of psychological intervention designed to “focus on creating healthy workplaces by enhancing employee’s well-being (e.g. high employee engagement) and reducing their health-related work problems (e.g. high sickness absence)” (Yang, Chang, & Lim, 2014; p. 564). Key questions regarding WOHP interventions pertain to their design, mechanisms and evaluation (Michel, O' Shea, & Hoppe, 2015). Early research on WOHP interventions were heavily weighted towards considering whether a particular intervention had an effect, but more recently, attention has shifted to understanding the mechanisms or processes through which such interventions work (Michel et al., 2015). This focus on mechanisms has two main considerations. The first pertains to a theoretical understanding of the psychological processes that are impacted through engaging in an intervention, which explain how the desired change (e.g. in well-being) occurs (Lippke & Ziegelmann, 2008; Michie, Johnston, Francis, Hardeman, & Eccles, 2008). The second pertains to the design of the research study itself, and how this can have a significant impact on the outcomes of interest. Less WOHP literature has focused on this latter issue of research design and method, and given its importance, we attend to it in this paper.
Issues of research design can prompt psychological processes such as increased self-awareness, reflection and expectation that may influence the outcome, and without awareness, such effects can confound the results (Boot, Simons, Stothart, & Stutts, 2013; Sitzmann & Wang, 2015). There has been general criticism of the WOHP field for a lack of evaluation research and strong research designs (Brough & O'Driscoll, 2010; Kompier, Cooper, & Geurts, 2000), while also acknowledging the difficulties of adopting true experimental designs with random allocation of subjects to treatment or control groups (Cox, Karanika, Griffiths, & Houdmont, 2007; Grant & Wall, 2009; Kompier et al., 2000). Nielsen and colleagues have called for two different types of evaluations for WOHP interventions, distinguishing outcome evaluations from process evaluations (Nielsen, Fredslund, Christensen, & Albertsen, 2006; Nielsen & Randall, 2012b; Nielsen, Randall, & Albertsen, 2007; Nielsen, Taris, & Cox, 2010b). Focusing on process evaluations has demonstrated that intervention effectiveness (or indeed lack of intervention effectiveness) may at times be due to programme implementation rather than issues with the theoretical rationale for the intervention (Nielsen et al., 2006). Randomized controlled trials (RCTs) are acknowledged as an important design component in the evaluation and understanding of causal processes that underlie theory-based interventions, thus contributing to evidence-based interventions (Michie & Abraham, 2004; Michie et al., 2005; Michie et al., 2008). Although RCTs do not guarantee a successful or meaningful outcome (Nielsen et al., 2006), they do contribute to the validity and rigor of intervention design, accounting for and reporting potential confounding effects of programme implementation.

In fields such as medical sciences, the use of RCTs to select individuals into different experimental conditions is well established, and considered a gold standard (Kaptchuk, 2001; Macdonald, Veen, & Tones, 1996; Richardson & Rothstein, 2008). However, to date, there has been little attention paid to how RCT methods can strengthen research designs focusing
on WOHP interventions, nor to practical issues regarding their implementation, which are significant to say the least (Sauter et al., 2002). We argue that the WOHP presents unique contextual and procedural issues that need to be incorporated into the RCT design to allow for wider implementation of RCT designs in organizational contexts. In this paper, we review the extent to which RCTs have been utilized in WOHP intervention research to date, and examine some of the challenges of implementing RCT in organizational contexts. We present an adapted RCT model that more strongly accounts for context and process issues relevant to WOHP, and discuss ways in which this helps WOHP researchers to overcome barriers to the implementation of RCTs. Our aim is that by providing clear and practical guidelines for RCT implementation specific to organizational contexts, researchers will be encouraged to adopt such methodological designs to a greater extent in the future, answering past calls for such approaches (Kelly et al., 2008; Kossek, Hammer, Kelly, & Moen, 2014). Secondly, by incorporating these guidelines as reporting standards, it will allow for more direct comparison across WOHP interventions, as well as with psychological health interventions in other disciplines such as medicine and public health.

**Randomized controlled trials: Overview**

The RCT is considered one of the most effective and powerful tools available to researchers. RCTs are used to evaluate the efficacy of an intervention/treatment/programme on any number of relevant outcomes (Stolberg, Norman, & Trop, 2004) and have become a criterion against which the quality of interventions are measured (Shadish, Cook, & Campbell, 2002). The core qualities of the RCT design lie in the process of allocating the participants to treatment/intervention and control groups, in such a way that every participant has an equal probability of assignment to any of the groups, which can be performed through several techniques (see Altman & Bland, 1999; Bowling & Ebrahim, 2005; Kang, Ragan, & Park, 2008). This gives the RCT the empirical rigor it is hailed for, in that it adds validity to
the statistical test, minimizes confounding, and reduces a number of biases due to the nature of the randomization process (see Kao, Tyson, Blakely, & Lally, 2008 for synopsis). Further, other aspects of the RCT design, namely masking/blinding, safeguards against ascertainment bias, while allocation concealment prevents selection bias (Altman & Schulz, 2001; Viera & Bangdiwala, 2007).

Within the RCT protocol there are a number of specific design options available to the researcher that determine how the participants are exposed to the intervention, for example parallel group, factorial, cross-over, or cluster design (see Table 1 for an explanation of each option). The scientific literature is replete with promoters of the RCT design as the most effective method for testing intervention effectiveness (Moher, Schulz, & Altman, 2001; Reynolds & Trinder, 2008), be they medical trials, educational programs, or health interventions in an organizational setting. Given the ability of RCTs to aid efforts in confirming the value of interventions, proponents of evidence-based practice identify RCTs as one of the best forms of scientific evidence in the causal inference framework, surpassing all other designs (Kitson, Harvey, & McCormack, 1998). Further, in what is known as the hierarchy of research designs, the findings obtained from RCTs are identified as evidence of the highest grade and value, by organizations such as the Medical Research Council and the U.S. Preventive Services Task Force (Campbell et al., 2000; Harris et al., 2001). This has encouraged the extension and promotion of RCT designs in the evaluation of interventions in the public health domain (Ham, Hunter, & Robinson, 1995).

To deal with potential sources of bias and to improve transparency and reporting of RCTs, the CONSORT statement was first developed in 1996 by an international group as
worldwide-accepted guidelines for reporting of RCTs. The CONSORT statement comprises a minimum set of standards, that has been updated twice in 2001 and 2010 (Moher et al., 2001; Schulz, Altman, & Moher, 2010), and is an evidence-based checklist of 25-items (see Table 1 or see Text S1 on http://www.consort-statement.org/) that are recommended for researchers conducting RCTs so that it aids completion of transparent reporting, critical appraisal, and interpretation of the published RCT. It is worth noting that the CONSORT statement is only concerned with the reporting of what was done and what was found, it does not include recommendations for designing, conducting and analyzing trials (Schulz et al., 2010).

**Prevalence of the CONSORT Statement guidelines**

Since the introduction of the CONSORT statement there has been widespread adoption and support of the guidelines among the scientific community. Indeed, there are now over 400 academic journals globally that explicitly support the CONSORT statement. This is not confined to medical journals, it is also supported by psychology journals, and with a revised CONSORT for psychology and sociology currently being developed (CONSORT-SPI; Grant et al., 2013; Montgomery et al., 2013) these reporting standards will be increasingly common in psychology. Moreover, other disciplines, including occupational therapy, and planning and development are actively pursuing similar guidelines (Dahl Rasmussen, Malchow-Møller, & Barnebeck Andersen, 2011), implying that the CONSORT is seen as a way of moving the science of interventions forward. The implications of this are obvious; if we as a discipline want to adhere to best practice then we need to apply comparable methods as those in other sciences, and be consistent with other fields of psychology. However, we recognize that some adaptations are necessary and we address these below.

Further, given the support of journals for adopting the statement it is quite common now to see the Journal’s Guide for Authors section include a reference to the application of the
RCTs in WOHP interventions

The CONSORT statement for reporting interventions. In addition, bodies such as the World Health Organization are major supporters of the CONSORT statement; thus, there is both academic and practical value in their incorporation. As the practice becomes more widespread among the academic and scientific community, it is imperative for WOHP as a discipline that we are operating in line with these recommendations.

Additionally, a basic principal of research methods training requires sufficient detail to allow for replication. The CONSORT statement is no different in this regard, and some argue that when studies do not report sufficient methodological detail the link between research and practice is weakened which is a waste of resources (Glasziou et al., 2010). In fact, Montgomery, Peters, and Little (2003) suggest that if research is to have its intended impact then the processes for reporting intervention trials need to be as robust and rigorously applied as the methods for actually conducting them.

**RCT in WOHP intervention research**

Although significant strides have been made in recent years to enhance the quality of intervention research in work and organizational psychology (Cox et al., 2007; Cox, Taris, & Nielsen, 2010; Nielsen, 2013; Nielsen et al., 2010a; Nielsen et al., 2010b), there are still indications that improvements are needed. For example, in a recent systematic review of resilience interventions in the workplace, Robertson, Cooper, Sarkar, and Curran (2015) found that eight out of fourteen studies used RCTs.

In order to examine the extent and manner with which RCTs have been utilized in work and organizational psychology to date, we consulted fifteen journals from the top twenty journals in the “Psychology Applied” category of the ISI Web of Science journal rankings (we excluded 5 of these 20 as they did not focus on work and organizational psychology). In each of these journals, we searched individual articles since 1996 (the year the CONSORT
RCTs IN WOHP INTERVENTIONS

Guidelines were first introduced) using the individual search terms: ‘intervention’, ‘random*’, ‘randomized controlled trial’, “RCT” and “CONSORT”. We also searched for the term ‘controlled trials’ in order to include papers that had partially adhered to an RCT design. We included articles that: (i) included an intervention, (ii) where either the outcome or the intervention had a WOHP focus, (iii) had implemented an RCT design, or approximated an RCT design, e.g. a controlled trial (CT) without randomization (we excluded papers that did not include any control group). We excluded papers that (i) focused on a physical intervention (i.e. did not have a psychological focus) or an occupational therapy intervention, (ii) research that utilized a non-worker sample (e.g. student samples and clinical samples were excluded), and (iii) pure laboratory experiments. Although we acknowledge many of these approaches have merit, they did not fit the purpose of our review, which was to examine RCT designs in work and organizational health psychology intervention research. In total, we found 33 papers meeting our criteria. We excluded 11 papers which utilized a RCT or CT design in a work context, but which did not have a WOHP focus. Table 2 provides a summary of the papers by journal.

-----------------------------

INSERT TABLE 2 HERE

-----------------------------

To review the extent to which these papers adhered to standard recommendations for RCT designs, their methodological quality were evaluated using the Cochrane risk of bias assessment tool (see Figure 1). This assessment is grounded in the information reported, rather than the quality of the actual study conducted. This tool includes screening for potential risks of bias in selection, performance, detection, attrition, and response in RCTs. Specifically individual papers were classified as having a high, low, or unclear risk of bias for (a) use of random sequence generation, (b) allocation concealment, (c) blinding of
researchers and participants, (d) blinding of outcome assessment, (e) incomplete outcome data, (f) selective reporting and (g) other bias (i.e. baseline bias).

Of the thirty-three studies, only nine were considered to be at low risk concerning the random sequence generation, while twelve were considered to be at an unclear risk, due to inadequate reporting of the randomization process. A further twelve were considered to be at high risk, as they did not use random assignment; although many studies justified the reasons for this, which were primarily pragmatic. We consider this to be good practice where randomization is not practical. Allocation to intervention versus control groups was poorly reported across the studies, with seventeen rated as having an unclear risk and a further thirteen were rated as having a high risk of bias as allocation was not well concealed. Only three studies had sufficient information to be considered at a low risk of bias. This points to an area where WOHP intervention researchers can provide clearer information when reporting their studies in future. For example, Kröger et al. (2014) implemented a matched randomization procedure, where eligible participants were matched on both gender and age (± 2 years), and then each pair was randomly allocated to one of the two treatment conditions (cognitive behavioral therapy as usual; CBT-AU, or work-related cognitive-behavioral therapy; W-CBT). In this example, random sequencing generation was deemed low risk as they employed a matched or paired randomization procedure, but they did not explicitly say how they (or whom) randomized the pairs so random allocation sequence risk of bias was unclear.

In contrast, Salmela-Aro, Mutanen, and Vuori (2012) and Vuori, Toppinen-Tanner, and Mutanen (2012) used sealed envelopes to conceal group allocation, and in this situation the random allocation sequence risk was considered low.

Blinding can be conducted in a number of ways: the participants can be blind to the condition they are in, the researchers can similarly be blind to the condition they are working with, or both. Due to the nature of these interventions it is often very difficult, or impossible
to blind participants as to what condition they are in. Consequently, thirty-two of the studies were considered to be at high risk of performance bias and detection bias, and only one was considered at low risk through employment of a blinding method. However, as noted previously by Zwarenstein et al. (2008), studies should still make efforts to blind assessors or obtain objective means for evaluating study outcomes, as it is the most robust and therefore desirable approach. In the Kröger et al. (2014) study, the researchers employed trained psychotherapists to conduct the eligibility screening interviews, who were blinded to the allocation of the treatment groups. However, given the two different treatment types, there was a requirement for the psychotherapists offering the CBT to know which therapy to provide. It was unclear as to whether attempts were made to blind participants to the intervention condition, but they would have to be aware of the content of the intervention, so it would have been very difficult to achieve participant blinding in this study.

For incomplete data, twenty-four studies were considered to be at a low risk of bias. Six were classified as having an unclear risk of bias, due to insufficient detail, or no attrition analysis being conducted, while three studies were considered to be at a high risk of bias due to results of attrition analysis. For example, Kröger et al. (2014; p, 3) included a CONSORT flowchart that clearly indicates the reasons for exclusion and the number of people excluded for each reason, as well as the attrition at each phase (allocation, follow-up and analysis). Also, a full information maximum likelihood estimation procedure was employed to handle the missing data, which when incorrectly dealt with or ignored, reduces statistical power and introduces bias in the results (Enders & Bandalos, 2001).

For selective outcome reporting, all studies were considered to be at an unclear risk of bias, as a judgment of this type of bias can only be made if a protocol is available, to compare the pre-specified outcomes with the published paper. This highlights the potential for future WOHP intervention researchers to register proposed trials, to aid in the transparency and
accuracy of the final published study. Finally, three studies were considered to be at a high risk of baseline bias as heterogeneity between groups at baseline was not statistically controlled for in subsequent analyses. Twenty-one of the thirty-three studies reported either no differences at baseline, or statistically controlled for these differences in the analysis and so were at a low risk of bias, while nine studies had not reported sufficient data and so were at an unclear risk of bias.

This risk assessment provides some guidance on areas where future WOHP intervention research can provide further clarity in reporting. Biron, Gatrell, and Cooper (2010) argue that risk assessments must be combined with appropriate implementation plans in order to produce any positive results, recommending the need to evaluate process and contextual issues in occupational health interventions. We discuss this further below.

It must be acknowledged that previous WOHP researchers expressed some doubts regarding the adequacy of the natural science paradigm on its own in providing an effective framework for evaluating interventions in applied domains (Cox et al., 2007). RCTs are associated with the traditional scientific approach to research, and thus, it may be that WOHP researchers have considered them impractical to implement (Grant & Wall, 2009). Despite such reservations, our review revealed that although the application of RCT is small in work and organizational psychology, it is possible. Indeed, when one considers implementing alternatives to the typical parallel group design (see Table 1), in particular cluster designs, then it is possible to offset some of the earlier criticisms. We argue that increased implementation of RCTs, in combination with standardized reporting guidelines, will
RCTs IN WOHP INTERVENTIONS

enhance the evidence-base and improve the validity and comparability of WOHP interventions, which have been criticized in the past (Cox et al., 2007).

Nielsen and Randall (2012b) discuss the need to consider the circumstances under which an intervention will work as an important component of process evaluation of interventions. In a similar fashion, it is important to consider the circumstances under which it is both necessary and feasible to implement RCTs when conducting WOHP interventions. Thus, we analyzed the WOHP intervention research identified from our review that had used an RCT. Although a small number of papers focused on a sample that was pre-assessed as experiencing significant psychological distress, burnout, depression or similar (De Vente, Kamphuis, Emmelkamp, & Blonk, 2008; Kröger et al., 2014; McGonagle, Beatty, & Joffe, 2014; Salmela-Aro et al., 2012; Salmela-aro, Näätänen, & Nurmi, 2004), others were based on particular areas of job design that were lacking for employees, such as psychosocial work factors (Hasson et al., 2014; Martin, Reece, Lauder, & McClelland, 2011) or on unemployed individuals in a job search programme (Vuori & Vinokur, 2005). The rest of the papers in our review did not select a distressed sample, but many focused on the reduction of ‘ill-being’ such as stress, depressive symptoms and distress (Flaxman & Bond, 2010; Füllemann, Jenny, Brauchli, & Bauer, 2015; Gardner, Rose, Mason, Tyler, & Cushway, 2005; Kröger et al., 2014; Le Blanc, Hox, Schaufeli, Taris, & Peeters, 2007; Lloyd, Bond, & Flaxman, 2013; Querstret, Cropley, Kruger, & Heron, 2015, in press; Schaer, Bodenmann, & Klink, 2008; van Dierendonck, Schaufeli, & Buunk, 1998; Vuori et al., 2012; Wolever et al., 2012).

Some recent papers examined how to improve psychological well-being for workers by enhancing their psychological and work-related resources. For example, a number of papers focused on interventions to improve the work context or perceptions of work conditions (Leiter, Day, Oore, & Spence Laschinger, 2012; Leiter, Laschinger, Day, & Oore, 2011; Logan & Ganster, 2005; Mikkelsen, Saksvik, & Landsbergis, 2000; Moen, Kelly, & Lam,
RCTs IN WOHP INTERVENTIONS

2013), on recovery from work and work-life balance (Hahn, Binnewies, Sonnentag, & Mojza, 2011; Hammer, Kossek, Anger, Bodner, & Zimmerman, 2011), job crafting (van den Heuvel, Demerouti, & Peeters, 2015) and career management (Vuori et al., 2012). Other papers focused on enhancing psychological resources such as mindfulness (Hülsheger, Alberts, Feinholdt, & Lang, 2013; Michel, Bosch, & Rexroth, 2014) and emotion regulation (Hülsheger, Lang, Schewe, & Zijlstra, 2015). This focus on prevention is an important step forward, which will strengthen WOHP interventions cross-comparison with research evaluating similar interventions in fields such as health psychology and other applied fields. Current recommendations suggest that primary interventions, where there is a preventive or proactive goal are more effective than secondary ameliorative interventions (developing skills and resources to cope with stressors), and in turn are more effective than tertiary interventions, which are more reactive in nature and aim to treat workers already experiencing psychological health issues (LaMontagne, Keegel, Louie, Ostry, & Landsbergis, 2007). For example, the positive psychology movement is a case in point. While little RCT research on positive psychology interventions is evident in WOHP, there is evidence for this in other applied domains of psychology (for recent meta-analyses see Boiler et al., 2013; Sin & Lyubomirsky, 2009). Although the majority of reviewed papers focused on the reduction of distress and negative well-being, which is the typical approach adopted in prevention interventions (Israel, Baker, Goldenhar, & Heaney, 1996), a number did draw on positive interventions such as resource-building (Vuori et al., 2012), mindfulness (Hülsheger et al., 2013; Wolever et al., 2012) and coaching (McGonagle et al., 2014). Mindfulness interventions, in particular, have been examined by multiple researchers using RCTs in workplace settings (Flook, Goldberg, Pinger, Bonus, & Davidson, 2013; Michel et al., 2014; Pidgeon, Ford, & Klaassen, 2013; Pipe et al., 2009). Taken together, it appears that the utilization of RCTs in WOHP intervention research is both possible and practical.
The WOHP Intervention Checklist

It is clear from our review that there are challenges to the implementation of traditional RCT designs in WOHP intervention research. Previous research has identified the process of intervention implementation as being a key issue in the success of interventions (Biron et al., 2010). Cox et al. (2007) define process as “the flow of activities; essentially who did what, when, why and to what effect” (p. 353). This is very similar to Johns (2001, 2006) call for a stronger focus on methodological context in organizational behavior research more generally. To ensure the feasibility of RCT designs in WOHP intervention research, we present an adapted the CONSORT checklist in the WOHP intervention checklist (see Table 3). For ease of comparison, we first present the traditional CONSORT checklist on the left-hand side of this table, then our adaptations on the right hand side (changes in italics). Many of the CONSORT items pertain to good reporting and thus require little adaptation. For example, in terms of improving reports in the title and abstract (Items 1a and 1b), two easy improvements are to identify in the title that the paper focuses on a randomized (and/or controlled) intervention, and to summarize the intervention design, methods and results in the abstract. Most of the papers in our review did include a summary of the intervention in the abstract, but only 5 identified this in the title (De Vente et al., 2008; Martin et al., 2011; Salmela-Aro et al., 2012; Vuori et al., 2012; Wolever et al., 2012).

INSERT TABLE 3 HERE

Moving to the introduction, we recommend that researchers clearly outline the scientific background, explanation and rationale for the design of the intervention, which should include the theorized mechanisms and outcomes of the intervention (Item 2a), in line with similar past recommendations (Kossek et al., 2014). This may seem like an obvious
statement. However, it is surprising how often this is omitted in WOHP intervention research. For example, in their systematic review of resilience interventions in work contexts, Robertson et al. (2015) demonstrate that although all fourteen papers in the review labeled their interventions as resilience interventions, only three of the fourteen studies assessed changes in resilience. Michie et al. (2008) discuss a range of techniques for moving from theory to intervention, specifically pertaining to changing behaviors, but which have general applicability in clarifying the design and expected outcomes of intervention research. Clearly outlining the hypothesized intervention mechanisms will also contribute to the outlining of clearer objectives that the intervention was designed to answer (Item 2b). For example, Michel et al. (2014) provide a strong rationale for the cognitive-emotional segmentation mechanism inherent in practicing mindfulness and why this should promote work-life balance, drawing on mindfulness as a self-regulation of attention mechanism (Bishop et al., 2004) as well as boundary theory (Ashforth, Kreiner, & Fugate, 2000; Nippert-Eng, 1996).

We propose a number of additional considerations in the methods and results sections relevant for the reporting, and evaluation of WOHP interventions. Two of the core adaptations we suggest pertain to the consideration of contextual and process issues. First, in considering the methods, we suggest that researchers should incorporate a section outlining the context within which the intervention is being delivered (Item 3), which is important in terms of understanding the processes by which the intervention and outcomes are brought about (Nielsen et al., 2010b). For example, organizational contextual influences such as ongoing issues (e.g. job insecurity, high turnover), organizational culture, and level of managerial support for the intervention are important to include (Kelly et al., 2008). In addition, broader environmental influences (e.g. economic recession versus boom) and psychosocial-environmental influences (e.g. major life events, daily hassles, chronic strains,
ambient environment; Israel et al., 1996) may significantly impact the responsiveness of participants to the intervention, and thus should be mentioned (if relevant). Johns (2006) describes an array of ways in which context can manifest including: the salience of situational features, the situational strength, a cross-level effect, a configuration of stimuli, an event, and a shaper of meaning. Each of these aspects should be considered by researchers in describing potential contextual influences to include in this section. Johns (2006) suggests that amongst others, two ways to contextualize research are to study processes and to collect qualitative data. The study of process is inherent in identifying the mechanisms by which an intervention is expected to have its effects. We cover qualitative contributions in process evaluation (Item 21) below. To provide an example of the detailed reporting of context, we refer readers to the paper by Vuori et al. (2012), which includes a text box outlining the organizational context issues pertinent to the WOHP intervention examined. We advise the inclusion of similar discussions of contextual influences in WOHP intervention going forward.

Moving to reporting intervention designs (Item 4a), we recognize that it is often challenging to find suitable control groups in field research (Nielsen et al., 2010b). To address this, we suggest that WOHP researchers should consider alternatives to the individual parallel group randomization, which is often difficult to achieve in organizational research. One could use cluster designs for franchise operations or multinational companies where one location receives the intervention while the other receives the control arm. For example, Le Blanc et al. (2007) used a cluster RCT design, where they randomized at the ward level in their examination of a team-based burnout intervention programme for oncology ward staff.

A further alternative is to consider assessing the equivalence, non-inferiority and/or superiority of a new intervention. The majority of WOHP interventions focus on comparisons with a “do nothing” control group. This tells us little about an interventions’ comparability with already existing approaches. Utilizing more ‘active’ control groups or
comparisons to existing interventions with similar objectives to the research intervention of interest also has practical advantages when working with organizations and in considering organizational leaders’ willingness to facilitate rigorous research designs, incorporating control group comparisons. For example, Kröger et al. (2014) compared a work related cognitive behavioral therapy (W-CBT) with a cognitive behavioral treatment as usual (CBT-AU) in order to compare the effectiveness of a new version of CBT (W-CBT) with more traditional approaches (CBT-AU). Prevalent concerns in organizational contexts are contamination effects from the experimental to control groups if the intervention is conducted in the same organization. Although there is direct comparability if a control group is included from the same organization as the intervention group, contamination may be an issue if, for example, participants in both groups work together or are in frequent communication (Nielsen et al., 2010b). However, the ideal approach when testing for equivalency is likely to be an experimental condition compared with both an active control group, and a ‘do nothing’ control group (Temple & Ellenberg, 2000), particularly when conducting an intervention in one organizational context. This can help to rule out context effects.¹

In terms of reporting the intervention design, we encourage researchers to consider reporting activity features such as those outlined by Lyubomirsky and Layous (2013). They suggest that the features of dosage (i.e. frequency and timing), variety (i.e. the same intervention activity and a variety of activities), trigger (e.g. what psychological processes the intervention may prompt), and level of social support are relevant across all interventions. Thus, we suggest these should be reported in intervention studies going forward. Activity features that may distinguish one intervention from another include: temporal focus (past, present, future), self vs. other oriented, and whether the intervention is social vs. reflective. For example, mindfulness interventions (Hülsheger et al., 2013; Michel et al., 2014) tend to

¹ Our thanks to Ekaterina Pogrebtsova for bringing this point to our attention.
focus on the present, while savoring interventions may focus on the past, present or future (Quoidbach, Dunn, Petrides, & Mikolajczak, 2010). Similarly, identifying whether an intervention was focused on the self or other can have important differences. O’Connell, O’Shea, and Gallagher (2015, in press) demonstrated that engaging in either a gratitude or kindness activity that was other-focused, rather than self-focused, resulted in greater increases in relationship satisfaction and perceptions of friendship improvement. Füllemann et al. (2015) found that shared participation (i.e. doing an activity socially rather than reflecting in private) played a role in the effectiveness of an occupational self-efficacy intervention, which they explained using a combination of social identity and shared mental models. Thus, the activity features can have a clear impact on the effectiveness of the intervention, and should be reported.

The remainder of the methods checklist items require only minor edits in order to implement in the context of WOHP interventions. We have added an emphasis on mechanisms in the outcomes (Item 7a). We have retained the recommendations pertaining the randomization (Item 8) and blinding (Item 11), although we recognize that where a WOHP intervention is being implemented to address organizational and/or employee issues it may not be feasible to implement allocation concealment or blinding of all involved. However, whether these occurred or not, the details should be reported.

In terms of reporting the results, it was quite surprising that only two of the papers in our review (De Vente et al., 2008; McGonagle et al., 2014) included a participant flow diagram (Item 14a). This is an easy improvement to make in WOHP intervention research going forward, and we encourage researchers to do so. In keeping with this, recruitment processes (Item 14), baseline data (Item 15) for participants in each group and numbers of participants in the analyses (Item 16) and effects sizes (Item 17) should be clearly reported. A further addition we recommend pertains to the outcomes and estimations (Item 18a), where
we suggest that to assess the specified mechanisms of an intervention, it may be appropriate to use mediation analysis to assess the indirect effect of the intervention on the outcomes via the hypothesized mechanisms.

The major addition we recommend in terms of reporting the results is to incorporate a process evaluation (Item 21). Nielsen and colleagues (Nielsen & Abildgaard, 2013; Nielsen et al., 2006; Nielsen & Randall, 2012b; Nielsen et al., 2010b) have outlined the merits of conducting both process and outcome evaluations of organizational level interventions, and the value of process evaluation is also recognized as important in the evaluation of RCTs (Oakley, Strange, Bonell, Allen, & Stephenson, 2006). Neilson and colleagues (Nielsen & Randall, 2012a; Nielsen et al., 2007) have shown that issues such as employee appraisals, participation, and perceptions of changes are important influencing factors (separate to the intervention design and implementation) than can influence the intervention effectiveness and are a useful component to include. Nielsen et al. (2006) demonstrate that qualitative reports can be very beneficial in understanding the effectiveness or otherwise of interventions, which can contribute to deeper understanding of the activities of the intervention programme, the project organization and involvement of employees, identification of ownership of the intervention, and the influence of other organizational changes and processes on the intervention. Indeed, incorporating qualitative components to intervention studies may be critical for understanding the usability and feasibility, and aid in the implementation of interventions that are both evidence-based and user informed in nature.

We acknowledge that it may not always be feasible to include all the aforementioned information in the main article, due to word restrictions in journals; in these instances we encourage authors to include this information as supplementary material when possible.

**Utilizing RCT: some considerations**
In this next section, we move to consider practical considerations and recommendations which we hope will encourage WOHP researchers to adopt RCT designs more frequently in future research. Despite RCTs being seen as the gold standard for conducting health and behavioral interventions, caution is warranted, as not all RCTs are equitable. In fact, when they lack methodological rigor the results may be biased (Jüni, Altman, & Egger, 2001; Schulz, Chalmers, Hayes, & Altman, 1995). These potential biases concern issues surrounding feasibility, participant selection and attrition, and measurement and performance/treatment variation Oakley et al. (2003), which we review below.

Feasibility. One of the more pervasive critiques in the literature is how feasible it is to implement and replicate an RCT, including questions pertaining to external validity due to the emphasis placed on securing internal validity (Glasgow et al., 2006; Rothwell, 2005). This issue of applicability and transferability of results generated from rigid RCT designs is an important one, resulting in many calls and recommendations for a validated framework of systematically addressing issues of external validity in all health publications (for review see Burchett, Umoquit, & Dobrow, 2011). Therefore, researchers employing an RCT design are advised not to neglect external validity, particularly in the reporting and interpretation of findings (Rothwell, 2005). In fact, a good theory-based design will allow researchers to understand the context in which the intervention has or has not worked, and where and how it can be implemented in the future. The field of WOHP is favorably placed in that it has the ability to overcome this imbalance, given its focus on field research (Edmondson & McManus, 2007; Grant & Wall, 2009). Thus incorporating RCTs into current WOHP will facilitate the achievement of high standards of both internal validity and external validity, producing research that is translational in nature with equivalent scientific rigor, practicality, and applied value.
RCTs can also be viewed as unfeasible due to the complex nature of applied field research in WOHP. However, it is also argued that, in fact, it is this complexity that necessitates the use of RCTs (Bonell, Hargreaves, Strange, Pronyk, & Porter, 2006; Sheldon & Oakley, 2002). The very nature of the design, in its random allocation process, ensures the formation of homogeneous treatment assignment, and unbiased comparison groups, which optimizes the likelihood that differences observed between the experimental and control intervention groups can be attributed to the intervention under investigation, rather than other confounding factors not explicitly accounted for (Kunz, Vist, & Oxman, 2007; Sheldon & Oakley, 2002). Although confounding is also possible in randomized experiments (see Greenland & Morgenstern, 2001) this minimization of bias grants the researcher a sound and reliable basis for making casual inferences regarding the effects of the intervention when conducted appropriately.

**Participant Selection and Attrition.** A key issue, which may impact the ability of an intervention to enhance well-being in the workplace, is the choice of sample. Well-being interventions in other domains of psychology (e.g. clinical, health) often select their sample from those with an established need (e.g. a previous diagnosis), resulting in a focus more frequently on the reduction of distress and so-called ‘ill-being’; the workplace presents a different context. While interventions may be focused on reducing environmental stressors and resulting strain, or on employees with established needs regarding the reduction of burnout, stress, amongst others, there is also a significant body of research, which now focuses on enhancing ‘well-being’ and the positive aspects of occupational health. These include engagement, vitality, and psychological resources, amongst others. Although this is beneficial in terms of preventative approaches, there is the potential that employees with an interest in such areas self-select into WOHP intervention studies. However, if an RCT design is used potential self-selection biases can be minimized, although it cannot completely
remove the effect of initial self-selection to participate in the first place. Utilizing RCT designs minimizes any pre-existing differences in participants across conditions, thereby facilitating the examination of other design effects such as dosage and variety (Lyubomirsky & Layous, 2013).

A further bias pertains to attrition rates. If one treatment condition had a higher rate of attrition than another this may influence the results as it is hard to know whether they did better or worse than the remaining cohort; similarly, if there is slight variation for selection procedures for each group or lack of clarity of how the randomization procedure worked this may also cause bias. Qualitative studies included as part of the intervention may provide more information on this particular issue.

**Measurement and performance/treatment variation.** If people assessing the outcomes know the group allocation, this level of knowledge may influence their assessment and analytic decision-making. One way to avoid this potential bias is to have trainers who are unaware of the research goals. As an example, Kröger et al. (2014) had external psychotherapists delivering the training, who were unaware of the research hypotheses. While this may not always be possible, it is a strength when it can be achieved. Further, regarding performance, all groups should be treated equally and expect to do as well as the other, as outcome expectancy has been found to influence trial outcomes (Boot et al., 2013). In the case of non-active control groups, it could be that any differences are a result of doing something rather than nothing (Sitzmann & Wang, 2015); thus one cannot really infer causation to treatment based on this study design. For this reason, we provided alternative designs, including active control groups, and comparison with ‘treatment as normal’ interventions.
A prominent objection to RCTs focuses on the allocation of a control group as inherently unethical and exploitative in nature. This ethical position is particularly valid in medical research that adopts placebo controls that are deemed invasive in nature and pose risk of serious harm to participants, which are unnecessary for scientific validity or worthwhile contributions (Cyna, Costi, & Middleton, 2011). Of particular significance to the use of control groups in WOHP research pertains to situations when the treatment or intervention is deemed to be of some value (due to strong theoretical evidence, pilot testing and/or expert opinion, for example) and is deliberately withheld from participants. For example, an intervention designed to moderate work-related stress and reduce employee burnout is given to half of the employees in an organization whilst the other half receive either an active control treatment, placebo control treatment, or nothing at all. Although allocation concealment and random allocation deals with the issue of selection bias (Viera & Bangdiwala, 2007), the appropriateness of a control group for testing applied health intervention is still subject to debate. This is despite its necessity in ensuring a rigorous evaluation design through producing socially equivalent experimental and control groups, if done successfully (Oakley, 1998). As stated by Solomon and colleagues (2008), research and organizational practice ethics must take precedence when planning, designing, and implementing psychosocial RCTs. There are many options available to WOHP researchers in carefully choosing a scientifically, ethically, and practically suitable control group against which to compare the intervention under investigation (De Vente et al., 2008; McGonagle et al., 2014). For example, using a waiting list control group appears as a common approach to dealing with this issue (Flaxman & Bond, 2010; Martin et al., 2011; McGonagle et al., 2014; Michel et al., 2014). As previously discussed, a further approach to dealing with such issues may be to adopt an ‘active-control’ approach where the control group themselves receive an
intervention, but one which only contains the ‘non-active’ or neutral components of the experimental manipulation.

Conclusion

In conclusion, when optimally conducted, RCTs can provide robust evidence to inform evidenced-based practice, with the goal of benefiting health at an individual and organizational level. Although it is acknowledged that it is challenging to employ an RCT design in an applied setting, WOHP researchers should strive to trial theory-driven interventions with methodological rigor, whilst maintaining their applied value (e.g. see Zapf, Dormann, & Frese, 1996). In this paper we have provided an overview of RCTs, discussed WOHP exemplars, and provided guidelines for broader implementation. Finally, we recommend incorporating RCTs into WOHP interventions where possible going forward; this methodological advancement coupled with the discipline’s existing applied utility, will aid in achieving increased legitimacy in the wider research community and inform future evidence-based practice.

REFERENCES

RCTs IN WOHP INTERVENTIONS


<table>
<thead>
<tr>
<th>Description of Design</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parallel Group Design</td>
<td>This is a basic, two-armed trial (two conditions) in which every participant is randomized to one of two groups. This is usually the intervention/experimental group or the control group, be it a placebo, active, waitlist, or no-treatment control group (see Hulley, Cummings, Browner, Grady, &amp; Newman, 2013).</td>
</tr>
<tr>
<td>Factorial Design</td>
<td>This design involves randomizing each participant to more than one intervention or condition in a single experiment. In its simplest form, a factorial design evaluating two interventions involves randomizing each participant to either receive both interventions, one of the two interventions, or neither of the interventions. This allows the researcher to evaluate the efficacy of more than one intervention, either independently, or as complementary interventions (see Montgomery, Astin, &amp; Peters, 2011; Montgomery et al., 2003).</td>
</tr>
<tr>
<td>Clustered Design</td>
<td>This involves randomizing at a group-level, for example, by community, class, school, or organization, rather than by each individual participant. These groups or “clusters” of people may be allocated to an intervention or control group (see Bowling &amp; Ebrahim, 2005; Hahn, Puffer, Torgerson, &amp; Watson, 2005).</td>
</tr>
<tr>
<td>Cross-over Design</td>
<td>In this design each participant is randomized to receive the control and intervention in random order. In this way half of the participants firstly receive the intervention followed later by the control, whilst the other half of the participants receives the control followed later by the intervention. Sufficient time is needed before participants are moved to the alternative condition to reduce carryover effects (see Hulley et al., 2013).</td>
</tr>
</tbody>
</table>
Table 2. Summary of past research on WOHP CT or RCT interventions in top work and organizational psychology journals.

<table>
<thead>
<tr>
<th>Journal</th>
<th>Total Number WOHP Intervention articles</th>
<th>CT WOHP intervention</th>
<th>RCT WOHP intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Journal of Occupational Health Psychology</td>
<td>10</td>
<td>Hahn et al. (2011); Leiter et al. (2011); Moen et al. (2013); Rydstedt, Johansson, and Evans (1998)</td>
<td>Bond and Bunce (2001); De Vente et al. (2008); Flaxman and Bond (2010); Kröger et al. (2014); McGonagle et al. (2014); Wolfever et al. (2012)</td>
</tr>
<tr>
<td>2. Journal of Applied Psychology</td>
<td>7</td>
<td>Leiter et al. (2011); van Dierendonck et al. (1998)</td>
<td>Hammer et al. (2011); Hülshgeger et al. (2013); Hülshgeger, Feinholdt, and Nübold (2015); Le Blanc et al. (2007); Vuori et al. (2012)</td>
</tr>
<tr>
<td>3. Work &amp; Stress</td>
<td>6</td>
<td>Hasson et al. (2014)</td>
<td>Cheng, Kogan, and Chio (2012); Gardner et al. (2005); Lloyd et al. (2013); Mikkelsen et al. (2000); Salmela-aro et al. (2004)</td>
</tr>
<tr>
<td>7. Applied Psychology – Health &amp; Well-being2</td>
<td>1</td>
<td>None</td>
<td>Martin et al. (2011)</td>
</tr>
<tr>
<td>10. Organizational Behavior and Human Decision Processes</td>
<td>0</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>11. Journal of Behavioral Decision Making</td>
<td>0</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>12. Journal of</td>
<td>0</td>
<td>None</td>
<td>None</td>
</tr>
</tbody>
</table>

2 Applied Psychology – Health & Well-being: from 2009 onwards
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Count</td>
<td>0</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Total</td>
<td>33</td>
<td>10</td>
<td>23</td>
<td></td>
</tr>
</tbody>
</table>

*CT = Controlled trial; an intervention which included an experimental and control group but where allocation into each group was not randomized.

*RCT = Randomized controlled trial; an intervention in which allocation into experimental group(s) and control group(s) was randomized.
### Table 3. Comparison between the CONSORT checklist and the adapted WOHP Intervention Checklist.

<table>
<thead>
<tr>
<th>Section/Topic</th>
<th>Item No</th>
<th>Consort Checklist</th>
<th>WOHP RCT Intervention Checklist</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Title and abstract</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1a</td>
<td>Identification as a randomized trial in the title</td>
<td>Identification as a randomized intervention in the title</td>
</tr>
<tr>
<td></td>
<td>1b</td>
<td>Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)</td>
<td>Structured summary of intervention design, methods, results and conclusions</td>
</tr>
<tr>
<td><strong>Introduction</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Background and objectives</td>
<td>2a</td>
<td>Scientific background and explanation of rationale</td>
<td>Scientific background and explanation of rationale, including theorized mechanisms and outcomes</td>
</tr>
<tr>
<td></td>
<td>2b</td>
<td>Specific objectives or hypotheses</td>
<td>Objectives are the questions that the intervention was designed to answer</td>
</tr>
<tr>
<td><strong>Methods</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Context</td>
<td>3</td>
<td>Discussion of the organizational/environmental context in which the intervention was implemented, including any pre-existing information which may have informed the choice and design of the intervention</td>
<td></td>
</tr>
<tr>
<td>Trial design</td>
<td>3a</td>
<td>Description of trial design (such as parallel, factorial) including allocation ratio</td>
<td>Description of intervention design (such as parallel, factorial) including allocation ratio</td>
</tr>
<tr>
<td></td>
<td>3b</td>
<td>Important changes to methods after trial commencement (such as eligibility criteria), with reasons</td>
<td>Important changes to methods after trial commencement (such as eligibility criteria), with reasons</td>
</tr>
<tr>
<td>Participants</td>
<td>4a</td>
<td>Eligibility criteria for participants</td>
<td>Eligibility criteria for participants</td>
</tr>
<tr>
<td></td>
<td>4b</td>
<td>Settings and locations where the data were collected</td>
<td>Settings and locations where the data were collected</td>
</tr>
<tr>
<td>Interventions</td>
<td>5</td>
<td>The interventions for each group with sufficient details to allow replication, including how and when they were actually administered</td>
<td>The interventions for each group with sufficient details to allow replication, including how and when they were actually administered</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outcomes</td>
<td>6a</td>
<td>Completely defined pre-specified primary and secondary outcome measures, including how and when</td>
<td>Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed. These should</td>
</tr>
<tr>
<td>Component</td>
<td>A</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>---</td>
<td>-------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td><strong>RCTs IN WOHP INTERVENTIONS</strong></td>
<td>35</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Sample size</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7a How sample size was determined</td>
<td>7a</td>
<td>How sample size was determined</td>
<td></td>
</tr>
<tr>
<td>6b Any changes to trial outcomes after the trial commenced, with reasons</td>
<td>8b</td>
<td>Any changes to outcome after the intervention commenced, with reasons</td>
<td></td>
</tr>
<tr>
<td><strong>Randomization</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8a Method used to generate the random allocation sequence</td>
<td>9a</td>
<td>Method used to generate the random allocation sequence</td>
<td></td>
</tr>
<tr>
<td>8b Type of randomization; details of any restriction (such as blocking and block size)</td>
<td>9b</td>
<td>Type of randomization; details of any restriction (such as blocking and block size)</td>
<td></td>
</tr>
<tr>
<td><strong>Allocation concealment mechanism</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned</td>
<td>10</td>
<td>Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned</td>
<td></td>
</tr>
<tr>
<td><strong>Implementation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions</td>
<td>11</td>
<td>Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions</td>
<td></td>
</tr>
<tr>
<td><strong>Blinding</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11a If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how</td>
<td>12a</td>
<td>If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how</td>
<td></td>
</tr>
<tr>
<td>11b If relevant, description of the similarity of interventions</td>
<td>12b</td>
<td>If relevant, description of the similarity of interventions</td>
<td></td>
</tr>
<tr>
<td><strong>Statistical methods</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12a Statistical methods used to compare groups for primary and secondary outcomes</td>
<td>13a</td>
<td>Statistical methods used to compare groups for primary and secondary outcomes, and tests of mechanisms of the intervention</td>
<td></td>
</tr>
<tr>
<td>12b Methods for additional analyses, such as subgroup analyses and adjusted analyses</td>
<td>13b</td>
<td>Methods for additional analyses, such as subgroup analyses and adjusted analyses</td>
<td></td>
</tr>
<tr>
<td><strong>Results</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participant flow (a diagram is strongly recommended)</td>
<td>13a</td>
<td>For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analyzed for the primary outcome</td>
<td></td>
</tr>
<tr>
<td>13b For each group, losses and exclusions after randomization, together with reasons</td>
<td>14b</td>
<td>For each group, losses and exclusions after randomization, together with reasons</td>
<td></td>
</tr>
<tr>
<td>Recruitment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14a Dates defining the periods of recruitment and follow-up</td>
<td>15a</td>
<td>Dates defining the periods of recruitment and follow-up</td>
<td></td>
</tr>
<tr>
<td>14b Why the trial ended or was stopped</td>
<td>15b</td>
<td>If relevant, why the trial ended or was stopped</td>
<td></td>
</tr>
<tr>
<td>Baseline data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15 A table showing baseline demographic and clinical characteristics for each group</td>
<td>16</td>
<td>A table showing baseline demographic and relevant characteristics for each group</td>
<td></td>
</tr>
</tbody>
</table>
### RCTs IN WOHP INTERVENTIONS

<table>
<thead>
<tr>
<th><strong>Numbers analyzed</strong></th>
<th>16</th>
<th>For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups</th>
<th>17</th>
<th>For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Outcomes and estimation</strong></td>
<td>17a</td>
<td>For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)</td>
<td>18a</td>
<td>For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval). If relevant, mediation analysis to assess the indirect effect of the intervention on the outcomes via hypothesized mechanisms.</td>
</tr>
<tr>
<td>Ancillary analyses</td>
<td>17b</td>
<td>For binary outcomes, presentation of both absolute and relative effect sizes is recommended</td>
<td>18b</td>
<td>For binary outcomes, presentation of both absolute and relative effect sizes is recommended</td>
</tr>
<tr>
<td><strong>Ancillary analyses</strong></td>
<td>18</td>
<td>Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory</td>
<td>19</td>
<td>Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory</td>
</tr>
<tr>
<td><strong>Harms</strong></td>
<td>19</td>
<td>All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)</td>
<td>20</td>
<td>All important harms or unintended effects in each group</td>
</tr>
<tr>
<td><strong>Process Evaluation</strong></td>
<td>21</td>
<td>Reporting and evaluating: (1) the activities of the intervention programme, (2) the project organization and the involvement of participants, (3) identification of ownership of the intervention project and activities, and (4) the influence of other organizational or contextual changes and processes on the intervention project (Nielsen et al., 2006)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Discussion</strong></td>
<td>20</td>
<td>Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses</td>
<td>22</td>
<td>Intervention limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses, and non-empirical limitations that may have arisen from qualitative process evaluation. Generalizability (external validity, applicability) of the intervention findings</td>
</tr>
<tr>
<td>Limitations</td>
<td>Generalizability</td>
<td>Generalizability (external validity, applicability) of the trial findings</td>
<td>23</td>
<td>Generalizability (external validity, applicability) of the intervention findings</td>
</tr>
<tr>
<td>Interpretation</td>
<td>22</td>
<td>Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence</td>
<td>24</td>
<td>Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence</td>
</tr>
<tr>
<td><strong>Other information</strong></td>
<td>23</td>
<td>Registration number and name of trial registry</td>
<td>25</td>
<td>Where the full intervention protocol can be accessed, if available</td>
</tr>
<tr>
<td>Registration Protocol</td>
<td>24</td>
<td>Where the full trial protocol can be accessed, if available</td>
<td>26</td>
<td>Sources of funding and other support, role of funders</td>
</tr>
<tr>
<td>Funding</td>
<td>25</td>
<td>Sources of funding and other support (such as supply of drugs), role of funders</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 1. Pooled risk of bias results using the Cochrane Risk of Bias Assessment tool.

Numbers reflect the number of studies out of thirty-three.