

Meilleures pratiques en matière de recherche

L'importance de procéder à une étude pilote pour les essais cliniques aléatoires en matière d'intervention

**Nancy Feeley, Sylvie Cossette, José Côté, Marjolaine Héon,
Robyn Stremler, Geraldine Martorella et Margaret Purden**

L'étude pilote fournit une occasion unique de déterminer les difficultés que pose l'évaluation d'une intervention et de s'y préparer. Au bout du compte, cette façon de faire permettra d'améliorer la rigueur et la valeur scientifique de l'étude à grande échelle. Bien que certaines revues publient les comptes rendus d'études pilotes, on accorde très peu d'attention à ces types de questions et de thèmes de recherche qui sont propres à ces mêmes études. Le présent article porte principalement sur l'utilité de procéder à un essai clinique pilote sur un échantillon aléatoire comme première étape avant d'effectuer un essai clinique aléatoire. On y décrit les principaux objectifs d'un essai clinique pilote sur un échantillon aléatoire : à savoir, évaluer la faisabilité et l'acceptabilité de l'intervention, du plan de recherche et des méthodes ainsi que déterminer plus facilement les valeurs des effets à des fins de calculs pour la taille de l'échantillon.

Mots clés : études pilotes, essais cliniques aléatoires, méthodes, faisabilité, acceptabilité

Best Practices for Research

The Importance of Piloting an RCT Intervention

**Nancy Feeley, Sylvie Cossette, José Côté, Marjolaine Héon,
Robyn Stremler, Geraldine Martorella, and Margaret Purden**

The pilot study provides a unique opportunity to identify and prepare for the challenges of evaluating an intervention. Ultimately, it will enhance the scientific rigour and value of the full-scale study. Although some journals publish reports of pilot studies, little attention has been given to the types of research questions and issues specific to these studies. This article focuses on the utility of a pilot randomized controlled trial (RCT) as a first step towards conducting an RCT. Three major objectives of a pilot RCT are discussed: assessing the feasibility and acceptability of the intervention, assessing the feasibility and acceptability of the design and procedures, and facilitating the determination of effect sizes for use in sample-size calculations.

Keywords: pilot studies, randomized controlled trials, methods, feasibility, acceptability

Introduction

Evidence-based practice is currently a goal of the nursing profession and is thought to achieve optimal outcomes for patients (Melnyk & Fineout-Overholt, 2005). Meta-analyses of randomized controlled trials (RCTs) are considered the strongest source of evidence on which to base practice (Melnyk & Fineout-Overholt, 2005). Thus the RCT is an important evidence-building tool for nursing practice and is increasingly being utilized in nursing research. As noted by Sidani and Braden (1998), the evaluation of nursing interventions is challenging because numerous factors can hinder one's ability to implement the experimental design and deliver the intervention as planned. The pilot RCT provides the researcher with an invaluable opportunity to identify these challenges before conducting a full-scale RCT. Moreover, the pilot study can be an important first step in securing funding for a full-scale RCT. Since RCTs are expensive, evidence that a procedure is feasible will be invaluable (Gardner, Gardner, MacLellan, & Osborne, 2003).

Although nurse researchers usually learn how to conduct experimental studies during their training (Bennett, 2005), they do not learn about pilot studies, and the majority of research textbooks fail to address this topic. Moreover, although a small number of nursing journals publish reports of pilot studies, the types of research questions that can be addressed in a pilot study have received little attention (Gardner et al., 2003; Jairath, Hogerney, & Parsons, 2000). In this article we outline the potential objectives of a pilot study and strategies to address these in order to assist investigators with the planning and conduct of an RCT.

Objectives of Pilot Studies

A pilot study precedes and is closely related to a full-scale study (Hinds & Gattuso, 1991; Perry, 2001). It is used to assess the design, methodology, and feasibility of the larger study (Gardner et al., 2003). Pilot studies should have well-defined objectives, or questions, to ensure methodological rigour and scientific validity (Lancaster, Dodd, & Williamson, 2004). Although the pilot study is often designed in much the same way as the subsequent full-scale study (Jairath et al., 2000), its research questions are different, with refinement of the processes and methods being the central focus (Campbell et al., 2007; Oakley et al., 2006).

When developing a pilot study as a preliminary step in an RCT, the researcher has several design options from which to choose. Myriad factors may influence this decision. However, we advocate a pilot RCT as the best option because only this design can fully test the feasibility of the main features of an RCT, such as random assignment of participants to a group. Hence in this article we focus on the pilot RCT and describe three major objectives: (a) assessing the feasibility and acceptability of the intervention, (b) assessing the feasibility and acceptability of the design and procedures, and (c) facilitating the determination of effect sizes to use in sample-size calculations for a full-scale trial. We conclude with a discussion of the issues surrounding the publication of pilot RCTs.

Defining Feasibility and Acceptability

While many published pilot studies assess feasibility and acceptability, these terms have not been explicitly defined, and many authors do not differentiate between the two. This may be partly because the terms are closely intertwined. For example, if an intervention is not acceptable to participants, then it is unlikely that it will be feasible, and vice versa. Nonetheless, it is helpful to distinguish between the two terms in the interests of conceptual clarity. Feasibility is defined as the ease or convenience of execution (Soanes & Stevenson, 2005), while acceptability is defined as the suitability or favourability of reception (“Acceptableness,”

Assessment of feasibility	Determines whether the intervention, study design, and procedures can be successfully executed by the researcher and delivered to the participants as planned.
Assessment of acceptability	Determines the suitability of the intervention and the study procedures from the perspective of the clinical population of interest, the intervention providers, or the health professionals who provide care to the population of interest.
Assessment of intervention fidelity	Determines the extent to which the intervention can be provided as intended. These data can be used as indicators of feasibility in a pilot RCT.

2008). Therefore, we propose that feasibility is primarily concerned with the researcher's ability to execute the plan — that is, to provide the intervention and complete the study procedures — whereas acceptability is concerned with the suitability of the intervention or the research design from the perspective of the recipients, the intervention providers, or health-care professionals (Table 1).

Assessing the Feasibility and Acceptability of the Intervention

Before a pilot RCT can be conducted, the experimental intervention must be developed to the point where pre-testing is possible. The process of intervention development has been described by others (Conn, Rantz, Wipke-Tevis, & Maas, 2001; van Meijel, Gamel, van Swieten-Duijffjes, & Grypdonck, 2004) and is beyond the scope of this article. Feasibility and acceptability issues need to be considered early in the process of intervention development and should be systematically assessed in the pilot RCT, because non-significant results in the full-scale RCT might be a result of problems with feasibility (e.g., ineffective delivery) or acceptability (e.g., ineffective uptake) rather than ineffectiveness of the intervention itself (Santacroce, Maccarelli, & Grey, 2004).

Feasibility

Many different aspects of intervention feasibility can be examined, including the dose (i.e., number, frequency, and timing), content, and methods of delivery. One indicator of intervention feasibility is the percentage of sessions delivered to participants. In a pilot study assessing an intervention with mothers of very low birth-weight (VLBW) infants,

one objective was to determine the feasibility of delivering the intervention in six sessions while the mothers were in the neonatal intensive care unit (NICU) (Feeley et al., 2008). The intervention began shortly after the mothers had given birth, a time often marked by high levels of maternal psychological distress and precarious infant health. The number of sessions received by each participant was documented, and findings revealed that the intervention was feasible, since 83% of mothers received all six sessions. Although the percentage of sessions delivered is an indicator of feasibility, it also reflects (to some extent) the acceptability of the intervention for the study population. In cases where the researcher has reason to believe that one element of the intervention might be of particular importance for intervention effectiveness, the percentage of sessions provided may not be the best indicator; it may be more appropriate to determine whether the critical elements of the intervention have been provided.

Intervention fidelity is defined as the extent to which the intervention can be delivered as intended (Bruckenthal & Broderick, 2007). Most researchers will be familiar with the notion of monitoring intervention fidelity in a full-scale RCT. However, Bruckenthal and Broderick (2007) propose that intervention fidelity be assessed in a pilot RCT. This can reveal problems related to implementation that can be remedied and thus lead to improved delivery in the full-scale RCT. For example, in a pilot study for a coping skills intervention, Bruckenthal and Broderick assessed whether nurse providers implemented the intervention content according to the manual. Audiotapes of intervention sessions were analyzed using a checklist that included the essential content of each session. Although the findings indicated that most of the content (i.e., 86%) was delivered, one important element was omitted: the providers often failed to ask whether participants practised their new skills. To address this problem the researchers extended provider training and placed greater emphasis on how to follow up on participants' use of the coping skills they had been taught. It is hoped that enhanced training will translate into improved implementation of the intervention in the full-scale RCT.

Acceptability

There are many stakeholders in an RCT, and therefore a pilot RCT should examine the acceptability of an intervention from the perspective of several key groups, including patients, intervention providers, and health professionals. The nursing literature includes many examples of pilot RCTs that examine intervention acceptability. When researchers measure acceptability, they typically assess satisfaction, or perceptions of the helpfulness, credibility, comprehensibility, and user-friendliness of the intervention (Vandelanotte & De Bourdeaudhuij, 2003). For example,

Dennis (2003) administered a reliable and valid questionnaire to assess mothers' perceptions of a peer-support program. The measure examined the quality of each mother's relationship with the peer intervener and her satisfaction with the support. It also examined the peer interveners' perceptions of their experience in providing the intervention. In a pilot RCT for an intervention to help mothers manage their preterm infants' health problems, researchers assessed the usefulness of the intervention from the perspective of the participating mothers, the infants' primary health-care provider, and the nurses who delivered the intervention (Pridham et al., 2006). These data pointed to the need for modifications to the intervention. For instance, written guidelines on how to manage infant health problems were seldom used, and mothers suggested how these could be modified to enhance their utility.

It is evident from these examples that feasibility and acceptability assessment may reveal the need for some adjustment to an intervention. If careful attention has been paid to these issues in the early stages of intervention development, the RCT should require only fine-tuning, such as minor changes to the content and methods. As van Meijel and colleagues (2004) suggest, if the pilot study leads to a major modification of the intervention, the researchers should continue with the pilot procedure, until the intervention is optimized, before proceeding to the full-scale RCT.

Assessing the Feasibility and Acceptability of the Study Design and Procedures

Another major objective of a pilot RCT is to determine the feasibility and acceptability of the design and procedures (van Teijlingen & Hundley, 2002). This assessment determines if and how the design should be modified for the full-scale RCT (Hinds & Gattuso, 1991). It can also reveal threats to validity, such as contamination (Becker, Roberts, & Voelmeck, 2003). Although many design features can be examined, in this article we will focus on those most pertinent to RCTs.

Feasibility

Questions concerning the recruitment of participants need to be considered in the pilot RCT (Friedman, Furberg, & DeMets, 1998). An obvious indicator of recruitment feasibility is the percentage of eligible persons agreeing to participate. Of course this may also reflect the acceptability of the intervention or study procedures, as a person's willingness to take part will be influenced by his or her perception of these factors. To assess the extent to which the acceptability of the intervention and procedures may affect recruitment, a questionnaire can be adminis-

tered at the time of recruitment to identify the specific reasons for refusal (i.e., study procedures too time-consuming or intervention not appropriate).

Data concerning the rate of recruitment can be useful in determining the overall timeline for the full-scale RCT as well as in drawing up a budget (Lancaster et al., 2004). They may also be used to confirm the adequacy of proposed recruitment sites and the need for additional sites for the full-scale study. As well, they can be useful for assessing the effectiveness of recruitment methods, revealing obstacles to recruitment, and revising methods accordingly.

Inclusion and exclusion criteria can be scrutinized (Hinds & Gattuso, 1991) to determine whether the study population, as defined by the inclusion and exclusion criteria, is sufficiently large. If the pilot study reveals otherwise, revision of the criteria may be advisable, to capture a narrower or broader range of participants. Nevertheless the researcher will have to ensure that the intervention is appropriate for the revised sample. In pilot testing an intervention with caregivers of Parkinson's patients, Habermann and Davis (2006) found that many caregivers were in their 50s whereas the inclusion criteria stated that they had to be 60 or older. For the full-scale RCT, modifications were made to allow for the enrolment of caregivers in their 50s, in the knowledge that the intervention would be relevant for younger as well as older caregivers.

An important feature of RCTs is the blinding of participants, research staff, or clinicians — meaning that one is unaware of whether a participant has been assigned to the intervention or the control arm of the study (DiCenso, Guyatt, & Ciliska, 2005). A double-blind study design (i.e., both the participants and the outcome assessors are blind) is not usually possible in nursing intervention studies because in such studies participants are often provided with information or are assisted in changing their behaviour (Sidani, Epstein, & Moritz, 2003). However, single-blind studies, in which the assessors of outcomes are blind to group assignment, are often possible in nursing research. Pilot work can be an opportunity to determine the feasibility of blinding outcome assessors, if this is part of the protocol. For example, research staff who assess the primary outcome can be asked if they have any knowledge of the participants' group assignment. The same approach can be used, if applicable, to determine whether the clinicians caring for participants are aware of group assignment. If the pilot study reveals that assessors or clinicians have learned of participants' group assignment, then strategies can be devised to correct this problem in a full-scale study.

Reduction of bias in an RCT hinges on successful randomization. Randomization consists of two elements: (a) the generation of a random allocation sequence, and (b) the implementation of that sequence in such

a way that it is concealed until the participant is assigned to a group — that is, allocation concealment (Schulz & Grimes, 2002). Allocation concealment prevents research staff members who enrol participants and others from knowing the subsequent group assignment. A variety of allocation concealment methods are considered adequate. They include opaque, sequentially numbered, sealed envelopes; centralized randomization via a telephone service; and randomization via a Web site (Schulz & Grimes, 2002). When the pilot study includes random assignment of participants to a group, information can be gleaned as to the effectiveness of the allocation concealment method as well as how randomization will be received by potential participants (Lancaster et al., 2004).

Some participants may be reluctant to agree to random assignment, a difficulty that has been noted in nursing intervention studies (Fogg & Gross, 2000; Gross & Fogg, 2001). Some people may not wish to participate because they have a clear preference for one intervention and wish to choose which intervention they receive (Miranda, 2004). It is critical for the researchers to determine whether this might be the case before conducting a full-scale RCT.

If the pilot reveals that too few participants will agree to be randomized, then investigators may consider other design options, such as a partial RCT, random assignment by site, or the Zelen design. The partial RCT design allows participants who have a treatment preference and do not wish to be randomized to choose their group assignment, while those with no preference are randomly assigned (Sidani et al., 2003). The partial RCT design is prone to selection bias, of course, due to the lack of random assignment. If random assignment by site is utilized and the sites are not comparable, there is the threat of selection bias. With the Zelen design, participants are randomly assigned to groups before they are approached and give their consent (Homer, 2002). This option is controversial due to ethical concerns and the potential for loss of power if a large portion of participants decide to withdraw after randomization.

Patients' reluctance to take part in an RCT may indicate a lack of equipoise regarding the benefits and risks of the intervention. First proposed by Freedman (1987), clinical equipoise means that there is no consensus within the scientific community about the comparative merits of the intervention being tested in an RCT. There is evidence demonstrating that the acceptability of clinical equipoise can be crucial in determining whether participants will consent to randomization and accept their group allocation (Mills et al., 2003). For example, if potential participants feel strongly that the experimental intervention is more effective than the control, they may not agree to accept an equal chance of being assigned to the control group. It is important that willingness to accept random-

ization be established in a pilot study, before the investigators embark on a full-scale RCT.

Two particular issues that may warrant careful appraisal in a pilot study are contamination and co-intervention. Contamination occurs when participants in either group receive the intervention intended for those in the other group (DiCenso et al., 2005). When participants in the control group are exposed to the intervention, even in part, they may experience change in the desired outcomes (Becker et al., 2003). In a pilot study, researchers can devise ways to assess contamination and determine whether it needs to be addressed. For example, in the pilot study with mothers of VLBW infants, the research staff asked participants not to discuss the information they were acquiring in the experimental program with other mothers of infants in the NICU (Feeley et al., 2008). Post-intervention, mothers were asked if, and with whom, they had shared the information they acquired during the intervention. If contamination is a concern, the researcher might consider revising the design to randomize sites rather than individuals (Gross & Fogg, 2001); alternatively, randomization could be limited to one participant per patient care room.

Co-intervention occurs when interventions other than the experimental intervention, but affecting the target outcome, are administered differentially to both the experimental and the control group (DiCenso et al., 2005). Co-intervention can be an issue in nursing intervention studies, and a pilot study can determine whether this is problematic. Data can be gathered from participants about the use of services that might constitute co-intervention. For example, a pilot study for an intervention to promote maternal and infant sleep in the early postpartum period assessed mothers' sources of sleep information by asking them if they sought additional sleep-related advice (Stremler et al., 2006). This allowed the investigators to assess the pervasiveness of co-intervention and to determine whether there was an imbalance between groups in the amount of co-intervention. Excessive use of co-interventions or imbalance between groups may reflect a lack of confidence among participants in the potential usefulness of their assigned condition and indicate a need for future trial investigators to tailor the intervention to participants.

A pilot RCT can also be useful for scrutinizing the appropriateness, timing, and sequencing of study measures. For example, in a pilot test for an intervention for caregivers of Parkinson's disease patients, researchers assessed whether it was more effective to administer a questionnaire before or after an in-depth interview with participants (Haberman & Davis, 2006). The findings indicated that the questionnaire should be administered before the interview, as this would allow participants to develop rapport with the interviewer. The feasibility of observational

measures can be a particularly important consideration. In the pilot study with mothers of VLBW infants, the plan was to observe mother-infant interaction in the NICU before the intervention (Feeley et al., 2008). However, it was unclear whether it would be possible to observe such interaction so soon after preterm birth. In fact, the pilot study revealed that there was very little interaction, due to infant sleep or health status. Based on this finding, it was decided not to measure pre-intervention interaction in the RCT. As a result, a change from baseline to post-intervention could not be examined.

A pilot study can also examine the possibility of unexpected outcomes, which may prompt the investigators to include additional outcomes in the full-scale RCT — ones that initially had not been considered. Conversely, the pilot study may identify outcome measures that are inappropriate, lacking in sensitivity, or unfeasible (Hinds & Gattuso, 1991; van Teijlingen & Hundley, 2002). Pilot work also allows the researcher to determine which outcomes may be amenable to change, and at what time points (Campbell et al., 2007), facilitating the selection of the best outcome measures. For example, in their pilot study Stremler and colleagues (2006) found that, at 6 months postpartum, mothers who received the experimental intervention slept more than mothers in the control group. Given the evidence of the short-term efficacy of the intervention, a future RCT will examine longer-term outcomes.

Acceptability

A pilot study also provides an opportunity to assess the acceptability of data-collection procedures (Hinds & Gattuso, 1991), such as gauging participants' willingness to complete the study procedures or determining whether the data can be collected in a reasonable span of time. It is particularly important that the burden to participants be assessed in the pilot RCT study, as both the intervention and the study procedures place demands on participants. Moreover, in nursing studies participants are often ill or in the midst of dealing with a health crisis. Reviewers of the grant application for the full-scale RCT may have concerns about these issues that could well be assuaged by the pilot data. If the pilot study reveals that the procedures are unacceptable to participants, then the researcher can explore means of reducing the burden, such as shorter instruments, shorter data-collection sessions, or telephone interviews. For example, in the pilot study with mothers of VLBW infants, mothers completed post-intervention questionnaires, infant development was assessed, and mother-infant interaction was videotaped (Feeley et al., 2008). Mothers were willing to participate in a 1-hour home visit, and the pilot study found that data collection could be completed within this time-frame if carried out by two research assistants.

Determining Effect Sizes

Researchers have used pilot RCT estimates of outcome variance or proportion to determine the sample size for the full-scale RCT (van Teijlingen & Hundley, 2002). This is a particularly useful option when there is scant literature on which to base effect-size estimates and when the pilot RCT uses the same design and outcome measures as the full-scale RCT (Hertzog, 2008). However, it should be kept in mind that the pilot RCT may lead to an over- or underestimation of effect size and the possibility of over- or underpowering the full-scale trial (Kraemer, Mintz, Noda, Tinklenberg, & Yesavage, 2006). Ideally, effect-size estimates or outcome variance estimates from pilot studies are supported by the literature as well as by estimates of clinically meaningful differences between groups. Alternatively, a conservative approach to the use of such estimates in sample-size calculations should be used, in order to ensure adequate power for the future trial (see Hertzog [2008] for a detailed discussion of these methods). Since pilot studies typically have small sample sizes and are often underpowered to reveal statistically significant differences between groups, investigators and funding agencies should resist viewing small effect sizes in pilot studies as reason to reject moving forward to the full-scale RCT.

Regardless of the outcome estimate used in calculating sample size for a future RCT, attention must be paid to the issue of clinical significance. While an intervention may lead to statistically significant differences between groups, it will be useful in practice only if that translates to a clinically meaningful difference in health outcomes. Ideally, clinical consensus will indicate what difference in participant outcomes merits the time, expense, and effort of the proposed intervention. Although estimates of clinically significant differences in outcomes should be elicited from expert clinicians, a pilot RCT offers an excellent opportunity to elicit participants' opinions about what they would consider meaningful. This information can be invaluable in interpreting outcomes from the pilot study (Oakley et al., 2006) and can also be helpful in planning future RCTs. Sample-size calculations are more difficult to complete when the relationship between statistically significant and clinically meaningful differences is unknown. Pilot work can be useful in this regard.

Since the pilot study is not expected to be powered to detect differences between groups, there is no universally accepted calculation for pilot study sample size. However, Hertzog (2008) provides a statistical approach to determining sample size for pilot RCTs. This approach is aimed at obtaining estimates of variance in an outcome when an important difference between groups has already been identified. In such cases,

it is suggested that 10 to 20 participants per group will suffice. However, if a meaningful group difference is unknown and the pilot study is intended to establish an effect size for sample-size calculation, then 30 to 40 participants per group is suggested. Ultimately, the decision regarding sample size for a pilot RCT must also take into account the research timeline, human and financial resources, and the research objectives.

Publication of Pilot RCT Findings

Researchers wishing to publish the findings of their pilot study may encounter difficulties, since there is currently much controversy over the usefulness of such publications. Opinions vary regarding whether the report should focus on the process of implementing the study (i.e., feasibility and acceptability objectives, as discussed in this article), the study outcomes (i.e., hypothesis testing as to whether the intervention was efficacious), or both implementation and outcomes. We agree with those who argue that pilot RCT findings should be published, but with the stipulation that they should contribute meaningfully to the literature (Watson, Atkinson, & Rose, 2007). The publication of feasibility and acceptability findings of a pilot RCT serves to inform other researchers about the methodological or practical challenges of designing such studies (Gardner et al., 2003) and constitutes a meaningful contribution to nursing knowledge. However, we would also agree with van Teijlingen and Hundley (2002), who argue that the outcomes of pilot RCTs should be published only if they are interpreted with caution. For example, in the pilot RCT for the mother-infant sleep intervention, statistically significant differences in amount of maternal night-time sleep and number of infant awakenings were found between the experimental and control groups at 6 weeks postpartum (Stremler et al., 2006). The investigators published the results, reporting on both the process and the outcomes, but were careful to ensure that any conclusions about the efficacy of the intervention acknowledged both the small sample size and the limited generalizability. Furthermore, for these reasons they indicated that they would proceed to a full-scale RCT to enrol a more diverse sample.

A last issue regarding the publication of pilot RCT outcomes relates to knowledge transfer. Lancaster and colleagues (2004) reviewed pilot studies published in seven major journals and found that only a few reported that their purpose was to prepare for a future RCT. Thus clinicians may take the positive outcomes of a pilot RCT to mean that the intervention should be applied in practice, particularly if this is the best published evidence. Although various groups, such as the Centre for Evidence-Based Medicine at Oxford (InfoPOEMS, 2007), have developed systems to grade evidence, they do not mention pilot studies, thus

leaving it unclear how these studies should be considered. In any publication of a pilot RCT, the researcher should explicitly state that the purpose of the study was to prepare for a future full-scale RCT, so that clinicians do not conclude that the pilot RCT results can be interpreted as final.

Conclusion

In this article we have outlined the key objectives of a pilot RCT in order to assist investigators with the design of their own pilot studies. As noted by Hinds and Gattuso (1991), pilot studies are different from full-scale studies not in sample size but in purpose. Pilot studies can address a variety of methodological and practical questions and can be an invaluable first step in conducting feasible, scientifically sound, full-scale RCTs that will provide high-quality evidence on which to base practice. The pilot RCT study represents an excellent opportunity to assess the acceptability of an intervention and fine-tune the content and format, through feedback from participants and others. The dissemination of pilot findings can contribute to knowledge by furthering researchers' understanding of the methodological and practical challenges of designing and conducting intervention studies.

References

- Acceptableness. (2008). In *Wiktionary: A Wiki-based open content dictionary*. Retrieved August 15, 2008, from <http://en.wiktionary.org/wiki/acceptableness>.
- Becker, H., Roberts, G., & Voelmeck, W. (2003). Explanations for improvement in both experimental and control groups. *Western Journal of Nursing Research*, 25, 746–755.
- Bennett, J. A. (2005). The consolidated standards of reporting trials (CONSORT): Guidelines for reporting randomized trials. *Nursing Research*, 54, 128–132.
- Bruckenthal, P., & Broderick, J. E. (2007). Assessing treatment fidelity in pilot studies assists in designing clinical trials: An illustration from a nurse practitioner community-based intervention for pain. *Advances in Nursing Science*, 30, E72–E84.
- Campbell, N. C., Murray, E., Darbyshire, J., Emery, J., Farmer, A., Griffiths, F., et al. (2007). Designing and evaluating complex interventions to improve health care. *British Medical Journal*, 334, 455–459.
- Conn, V. S., Rantz, M. J., Wipke-Tevis, D. D., & Maas, M. L. (2001). Focus on research methods: Designing effective nursing interventions. *Research in Nursing and Health*, 24, 433–442.
- Dennis, C.-L. (2003). The effect of peer support on postpartum depression: A pilot randomized controlled trial. *Canadian Journal of Psychiatry*, 48, 115–124.

- DiCenso, A., Guyatt, G., & Ciliska, D. (2005). *Evidence-based nursing: A guide to clinical practice*. Toronto: Elsevier Mosby.
- Feeley, N., Zelkowitz, P., Charbonneau, L., Cormier, C., Lacroix, A., Ste. Marie, C., et al. (2008). Promoting mothers' ability to interact sensitively with their very-low birthweight infant: A pilot study. *Advances in Neonatal Care*, 8(5), 276–284.
- Fogg, L., & Gross, D. (2000). Threats to validity in randomized clinical trials. *Research in Nursing and Health*, 23, 79–87.
- Freedman, B. (1987). Equipoise and the ethics of clinical research. *New England Journal of Medicine*, 317, 141–145.
- Friedman, L. M., Furberg, C. D., & DeMets, D. L. (1998). Recruitment of study participants. In *Fundamentals of Clinical Trials*, 3rd ed. (pp. 141–155). New York: Springer.
- Gardner, G., Gardner, A., MacLellan, L., & Osborne, S. (2003). Reconceptualising the objectives of a pilot study for clinical research. *International Journal of Nursing Studies*, 40, 719–724.
- Gross, D., & Fogg, L. (2001). Clinical trials in the 21st century: The case for participant-centered research. *Research in Nursing and Health*, 24, 530–539.
- Habermann, B., & Davis, L. L. (2006). Lessons learned from a Parkinson's disease caregiver intervention pilot study. *Applied Nursing Research*, 19, 212–215.
- Hertzog, M. A. (2008). Considerations in determining sample size for pilot studies. *Research in Nursing and Health*, 31, 180–191.
- Hinds, P. S., & Gattuso, J. S. (1991). From pilot work to a major study in cancer nursing research. *Cancer Nursing*, 14, 132–135.
- Homer, C. S. (2002). Using the Zelen design in randomized controlled trials: Debates and controversies. *Journal of Advanced Nursing*, 38, 200–207.
- InfoPOEMS. (2007). *Levels of evidence: From the Centre for Evidence-Based Medicine (POEM)*. InfoPOEMS, the Clinical Awareness System. [Online.] Available: http://www.infopoems.com/concept/ebm_loe.cfm.
- Jairath, N., Hogerney, M., & Parsons, C. (2000). The role of the pilot study: A case illustration from cardiac nursing research. *Applied Nursing Research*, 13, 92–96.
- Kraemer, H. C., Mintz, J., Noda, A., Tinklenberg, J., & Yesavage, J. A. (2006). Caution regarding the use of pilot studies to guide power calculations for study proposals. *Archives of General Psychiatry*, 63, 484–489.
- Lancaster, G. A., Dodd, S., & Williamson, P. R. (2004). Design and analysis of pilot studies: Recommendations for good practice. *Journal of Evaluation in Clinical Practice*, 10, 307–312.
- Melnyk, B. M., & Fineout-Overholt, E. (2005). *Evidence-based practice in nursing and healthcare: A guide to best practice*. New York: Lippincott Williams & Wilkins.
- Mills, N., Donovan, J. L., Smith, M., Jacoby, A., Neal, D. E., & Hamdy, F. C. (2003). Perceptions of equipoise are crucial to trial participation: A qualitative study of men in the ProtecT study. *Controlled Clinical Trials*, 24, 272–282.
- Miranda, J. (2004). An exploration of participants' treatment preferences in a partial RCT. *CJNR*, 36, 100–114.

- Oakley, A., Strange, V., Bonell, C., Allen, E., Stephenson, J., & RIPPLE study team. (2006). Process evaluation in randomized control trials of complex interventions. *British Medical Journal*, *332*, 413–416.
- Perry, S. E. (2001). Appropriate use of pilot studies. *Journal of Nursing Scholarship*, *33*, 107.
- Pridham, K. A., Krolikowski, M. M., Limbo, R. K., Paradowski, J., Rudd, N., Meurer, J. R., et al. (2006). Guiding mothers' management of health problems of very low birth-weight infants. *Public Health Nursing*, *23*, 205–215.
- Santacroce, S. J., Maccarelli, L. M., & Grey, M. (2004). Intervention fidelity. *Nursing Research*, *53*, 63–66.
- Schulz, K. F., & Grimes, D. A. (2002). Generation of allocation sequences in randomised trials: Chance, not choice. *Lancet*, *359*, 515–519.
- Sidani, S., & Braden, C. J. (1998). *Evaluating nursing interventions: A theory-driven approach*. Thousand Oaks, CA: Sage.
- Sidani, S., Epstein, D. R., & Moritz, P. (2003). An alternative paradigm for clinical nursing research: An exemplar. *Research in Nursing and Health*, *26*, 244–255.
- Soanes, C., & Stevenson, A. (Eds.). (2005). *The Oxford dictionary of English*, rev. ed. Oxford Reference Online. Retrieved August 15, 2008, from <http://www.oxfordreference.com/views/ENTRY.html?subview=Main&entry=t140.e27054>.
- Stremler, R., Hodnett, E., Lee, K., MacMillan, S., Mill, C., Ongcangco, L., et al. (2006). A behavioral-educational intervention to promote maternal and infant sleep: A pilot randomized, controlled trial. *Sleep*, *29*, 1609–1615.
- Vandelanotte, C., & De Bourdeaudhuij, I. (2003). Acceptability and feasibility of a computer-tailored physical activity intervention using stages of change: Project FAITH. *Health Education Research*, *18*, 304–317.
- van Meijel, B., Gamel, C., van Swieten-Duijfjes, B., & Grypdonck, M. H. (2004). The development of evidence-based nursing interventions: Methodological considerations. *Journal of Advanced Nursing*, *48*, 84–92.
- van Teijlingen, E., & Hundley, V. (2002). The importance of pilot studies. *Nursing Standard*, *16*, 33–36.
- Watson, R., Atkinson, I., & Rose, K. (2007). Pilot studies: To publish or not? *Journal of Clinical Nursing*, *16*, 619–620.

Acknowledgements

This study was funded by the Quebec Interuniversity Nursing Intervention Research Group (GRISIIQ). Dr. Feeley is supported by a Clinical Research Scholar Award from Fonds de la recherche en santé du Québec (FRSQ).

We wish to thank Christine Saykaly and Robin Canuel for their assistance with the preparation of this manuscript.

Drs. Cossette, Côté, and Stremler and Ms. Héon contributed equally to the preparation of the manuscript and are listed below in alphabetical order.

Nancy Feeley, RN, PhD, is Senior Researcher, Centre for Nursing Research, and Project Director, Sir Mortimer B. Davis Jewish General Hospital; and Assistant Professor, School of Nursing, McGill University, Montreal, Quebec, Canada. Sylvie Cossette, RN, PhD, is Researcher, Montreal Heart Institute Research Centre, and Associate Professor, Faculté des sciences infirmières, Université de Montréal. José Côté, RN, PhD, is Associate Professor, Faculté des sciences infirmières, Université de Montréal, and Researcher, Centre de recherche du Centre hospitalier de l'Université de Montréal. Marjolaine Héon, RN, is a PhD candidate and Lecturer, Faculté des sciences infirmières, Université de Montréal. Robyn Stremler, RN, PhD, is Assistant Professor, Lawrence S. Bloomberg Faculty of Nursing, University of Toronto, and Adjunct Scientist, Hospital for Sick Children, Toronto, Ontario, Canada. Geraldine Martorella, RN, is a PhD candidate and Lecturer, Faculté des sciences infirmières, Université de Montréal. Margaret Purden, RN, PhD, is Scientific Director, Centre for Nursing Research, and Project Director, Sir Mortimer B. Davis Jewish General Hospital; and Assistant Professor, School of Nursing, McGill University.