Running head: INTENTION-TO-TREAT IN NURSING CLINICAL TRIALS

The Use of the Intention-to-Treat Principle in Nursing Clinical Trials

Denise F. Polit, Ph.D.

President, Humanalysis, Inc., Saratoga Springs, NY and

Adjunct Professor, Research Centre for Clinical & Community Practice Innovation

Griffith University, Gold Coast, Australia

Brigid M. Gillespie, RN, BHSc (Hons), PhD

Lecturer & Research Ethics Adviser

Research Centre for Clinical & Community Practice Innovation

Griffith University, Gold Coast, Australia

Abstract

Background: In randomized controlled trials (RCTs), the intention-to-treat (ITT) principle, which involves maintaining study participants in the treatment groups to which they were randomized regardless of post-randomization withdrawal, is the recommended analytic approach for preserving the integrity of randomization; yet, little is known about the use of ITT in nursing RCTs.

Objectives: The purposes of this study were to describe the extent to which nurse researchers who conduct RCTs state that they have used ITT, the extent to which they actually adhere to ITT principles, and the methods they use to handle missing data.

Method: Data regarding ITT analysis, participant flow, rates of attrition, and methods of handling missing data were extracted and coded from a consecutive sample of 124 RCTs published in 16 nursing journals in 2007 and 2008.

Results: ITT was declared in only 15.3% of the nursing RCTs, and fewer than half of these studies offered a definition of ITT. Based on authors' description of analytic procedures, we concluded that 10.5% of those claiming ITT use had actually used a per protocol rather than an ITT analysis. Overall, we classified 46.8% of the RCTs as having used either a classic or modified ITT analysis, indicating that many nurse researchers are not stating their actual adherence to ITT, despite advice to do so in the CONSORT guidelines.

Conclusions: Nurse researchers conducting RCTs should be more diligent in following CONSORT guidelines about ITT, documenting ITT use in their reports, clarifying their definition of ITT, and presenting flowcharts that describe subject flow. Readers of nursing reports, in evaluating evidence from RCTs, should not rely on stated use of ITT, but should examine how analyses were actually conducted.

Keywords: randomized controlled trials; intention-to-treat; attrition

In studies that test the effects of an intervention, a key objective is to provide an unbiased comparison of outcomes among groups exposed to different treatment conditions. Randomized controlled trials (RCTs) are considered the "gold standard" design strategy for achieving this objective (Hulley, Cummings, Browner, Grady, & Newman, 2001; Polit & Beck, 2008), and an intention-to-treat approach is the "gold standard" analytic strategy for preserving the integrity of randomization (Altman et al., 2002), but little is known about adherence to ITT principles in nursing clinical trials.

RCTs are widely considered to yield the highest quality evidence about the effects of an intervention because randomization to different treatment groups serves to equalize groups prior to treatment exposure with regard to an infinite number of characteristics. With randomization, post-intervention group differences in outcomes can be inferred as having been caused by the intervention, given that groups were equivalent at the outset (Polit & Beck, 2008).

Despite their avoidance of selection bias, RCTs can be undermined by other types of biases, such as those than can arise from the loss or removal of study participants after randomization. The principle referred to as intention-to-treat or intent-to-treat (ITT) is specifically designed to guard against the risk of bias that can occur when subjects who were randomized are not included in the analysis of outcomes.

Background

The literature generally attributes the first written description of ITT to the renowned methodologist Sir Austin Bradford-Hill (1961), who noted that post-randomization exclusion of subjects could affect the validity that randomization sought to achieve. Yet, although the term ITT has been used for nearly 50 years, there is no clearcut consensus on what intention-to-treat means (DeMets, 2004; Gravel, Opartney, & Shapiro, 2007; Hollis & Campbell, 1999). The strict definition of ITT involves a "once randomized, always analysed" philosophy—that is, that analyses of outcomes

must include all subjects who were randomized in the group to which they were

assigned, regardless of treatments actually received, deviations from the protocols, and

withdrawals from the study (Gravel et al., 2007; Lachin, 2004; Whittaker, Sutton, &

Burton, 2006). A less restrictive definition of ITT involves including all subjects in the

groups to which they were randomized, making efforts to obtain outcome data for all

subjects (including those who may not have gotten the intervention), and analyzing data

for those with follow-up outcome data, disregarding any missing data (Gravel et al.,

2007).

Although the ITT principle has had some controversy within the medical

community (Lachin, 2000), it is now widely championed by both professional

organizations and regulatory agencies, such as the U. S. Food and Drug Administration

(U. S. FDA, 1998). The Cochrane Collaboration, in its handbook for systematic reviews

of interventions, describes the risk of bias in studies not using ITT, and further states

that "analyses of randomized trials that do not include all randomized participants are

not intention-to-treat analyses" (Higgins & Green, 2008, p. 483).

Of particular importance, ITT has been advocated in the influential

Consolidated Standards of Reporting Trials (CONSORT) guidelines (Altman et al.,

2002). These guidelines, which have been adopted by dozens of medical journals and

several nursing journals, state the importance of including in RCT reports "information

about whether the investigators included in the analysis all participants who underwent

randomization, in the groups to which they were originally allocated (intention-to-treat

analysis)" (p. 677). The CONSORT guidelines recommend including a diagram to

show participant flow into and out of the study, and one purpose of such flowcharts is

to document whether ITT was adopted.

In the years following the issuance of the CONSORT guidelines, many studies

were undertaken to document the extent to which reports of RCTs in medical journals

adhered to the guidelines, and so there is considerable information in the medical literature about stated adherence to ITT. In generalist medical journals, declared use of ITT has ranged from 48% in studies published in the late 1990s (Hollis & Campbell, 1999; Ruiz-Canela, 2000) to 87% published in 2002-2003 (Mills, Wu, Gagnier, & Devereux, 2005), leading some to conclude that the CONSORT guidelines have had a large effect on researchers' use of—or at least on *reporting* use of—ITT.

Factors associated with stated ITT use have also been explored. For example, it has been found that ITT adherence is lower in medical specialty journals than in top-tier general journals—for example 12% in dermatology (Adetugbo & Williams, 2000) and 18% in endocrinology (Rios, Adueyungbo, Moitri, Rahman, & Thabane, 2008). Researchers have also found that stated use of ITT tends to be higher in studies with larger samples (Rios et al., 2008; Ruiz-Canela, 2000), in journals that adhere to CONSORT guidelines and that have higher impact factors (Gravel, *et al.*, 2007), and in studies that have been rated as having higher overall methodologic rigor (Huwiler-Muntener, Juni, Junker, & Egger, 2002; Ruiz-Canela, 2000).

Use of ITT within nursing RCTs has received little attention. Recently, however, a group of nurse researchers explored the extent to which reports of RCTs published from 2002 to 2005 in four major nursing research journals adhered to the CONSORT guidelines (Smith, Lee, Lee, Choi, Jones, Bausell, & Broome, 2008). These researchers found that only 11% of the 96 nursing reports in their sample of studies explicitly stated that ITT was used.

Several investigators who have scrutinized medical RCTs more closely have found that stated use of ITT does not mean that an ITT analysis was actually pursued. Indeed, the Cochrane handbook cautions that "it is generally unwise to accept study authors' descriptions of an analysis as ITT; such a judgment should be based on the detailed information provided" (Higgins & Greene, 2008, p. 489).

Hollis and Campbell (1999) were the first to compare stated and actual use of ITT in medical clinical trials. They noted that most reports that claimed use of ITT did not indicate how missing outcomes or deviations from protocols were handled; they found that several studies in their sample clearly violated ITT principles. Kruse and colleagues (2002) found that only 42 studies, out of 100 that professed use of ITT, included all randomized subjects in the analysis. Baron and colleagues (2005) found that ITT use was reported in 66.7% of the RCTs on rheumatic disease, but that a classic ITT analysis was actually performed in only 7.4% of the studies.

Most recently, Gravel and colleagues (2007) examined 403 studies published in 10 medical journals in 2002. They found that the use of ITT was reported in 62% of the studies in their sample, but that only 39% of the studies claiming ITT actually used it. They further examined an issue of vital importance to ITT, missing data. Among the studies with declared use of ITT, more than 60% had some attrition, with 21% having attrition in excess of 10% of those randomized. In most cases (59%), participants with missing outcome data were simply removed from the analysis. Imputations of missing data were reported in 12% of the studies with attrition, and the most frequently used method was an imputation strategy no longer considered optimal, namely last observation carried forward (i.e., using outcome information from the previous round of data collection to fill in a missing value in a later round). Gravel's findings are similar to those reported by Wood, White, and Thompson (2004), who found in their analysis of 71 trials in top-tier medical journals that missing outcome data was a widespread problem in RCTs and that missing values are often inadequately handled in the analysis of intervention effects.

In summary, ITT adherence is increasingly being reported in RCTs in the medical literature, but stated use cannot always be trusted to reveal how the analysis was actually performed. There is limited information about ITT in nursing studies, but

one recent investigation suggests that stated adherence to ITT is very low. To the best of our knowledge, there is no information on the *actual* use of ITT in RCTs conducted by nurse researchers.

Purpose

The purpose of this study was to describe the stated use and actual use of the ITT principle in RCTs reported in the nursing literature. Related purposes were to examine the extent of attrition in nursing RCTs, to understand how missing outcome values are handled in the analyses, and to describe the extent to which researchers do an analysis of attrition bias. Finally, we sought to identify factors that could influence stated and actual use of ITT analysis, including methodologic characteristics such as sample size, journal characteristics such as impact factor value, and other study characteristics such as having funding. On a more global level, our goal is to raise the consciousness of nurse researchers, users of evidence from nursing studies, and peer reviewers and editors of nursing journals to the issue of ITT and to its importance in drawing conclusions about the quality of evidence from nursing interventions studies.

Method

Sample

A consecutive sample of RCTs published in 16 English-language nursing research journals in the years 2007 and 2008 comprised the study sample. Journals were selected if they were English-language nursing journals that regularly published reports of RCTs. This criterion was operationized as journals that had published at least 5 studies listed as randomized controlled trials under "Type of Article" in PubMed in 2007-2008, and were classified in the nursing subset. Articles themselves, however, were not selected electronically. Rather, RCTs were identified by hand-searching all issues of the 16 journals, a process that has been used in other similar studies because of miscodings of article type within PubMed (Adetugbo & Williams, 2000).

All issues in the 16 journals in 2007-2008 (276 issues) were hand-searched, yielding 2,916 articles. Table 1 shows the journal names and the number of articles published in them over the 2-year period. The abstracts of these articles were perused to identify intervention studies. Articles that simply described an intervention model, reported analyses of baseline data only, described secondary analyses not involving intervention effects, or were systematic reviews of intervention studies were excluded from further consideration. Articles that reported on the effects of an intervention were further scrutinized for possible inclusion in the sample of RCTs. As shown in Figure 1, a total of 266 studies that reported the effects of an intervention were identified, but only 124 of these were included in our analysis, primarily because about half of the intervention studies were not RCTs. An RCT was defined as a study in which study participants were randomly allocated to 2 or more groups for the purpose of testing intervention effects. Intervention studies were excluded if they used a quasi-experimental design, if no author was an RN, if the randomization unit was not a human (e.g., animal studies), or if the article involved a secondary analysis of RCT data.

Variables and Data Extraction

Full reports for the 124 studies in our sample were retrieved and reviewed. Relevant information from each article was extracted, coded, and entered onto a coding protocol, which can be requested from the corresponding author. This section describes variables for which data were extracted.

Attrition. We recorded the number of study participants randomly assigned to various treatment groups, as well as the number in each group at the final post-random assignment follow-up. These numbers were used to compute the percent of attrition at the end of the study. We also coded, for studies with attrition, whether the researchers did an analysis of attrition bias. For studies with no attrition, we classified a study as

having a "captive audience" if follow-up data were collected essentially immediately after a short intervention (e.g., a massage), leaving virtually no opportunity for subject loss. Finally, we coded whether or not the article mentioned any efforts to minimize attrition, such as using telephone or email reminders or incentive payments.

Handling of Missing Values. We recorded how missing values were handled in the analyses. The coding categories were: no missing values, listwise or pairwise deletion of missing cases, and imputation of missing values or multi-level modeling. Method of imputation was recorded, into the following categories: last observation carried forward, worst case/best case imputation, mean substitution, regression imputation, expectation-maximization imputation, multiple imputation, and use of mixed models that accommodate missing values. Finally, we coded whether the robustness of assumptions about missing values were tested using sensitivity analyses.

Intention to Treat. We recorded what researchers said about having used ITT, according to the following categories: No mention of ITT or any specific analytic approach; intention to treat; modified intention to treat; and *not* ITT (per protocol). We also coded whether a definition of ITT was provided in studies that reported having used ITT.

Based on information we could glean from the report regarding attrition, subject flow, and handling of missing values, we assessed the approach that was actually used. We classified a study as having used "true ITT" according to the strictest definition of ITT—that is, if all subjects randomly assigned to different treatment groups were accounted for in the final analysis of outcomes. A true ITT can be accomplished either by having no post-randomization attrition, or by using a statistical method that does not remove a case with missing values from the analysis, namely by means of imputation or mixed modelling within a repeated measures framework. We classified a study as having used a "modified ITT" if there was evidence that the researchers attempted to

obtain follow-up outcome data from everyone who was randomized, regardless of whether they received the full or any treatment, but then analysed only cases for whom follow-up data were obtained. We classified studies as using a "per protocol" analysis using the Cochrane Collaboration definition: analyses based on people who were kept in the treatment group to which they had been randomly assigned *and* who completed the trial. We coded the analysis as "unclear" if no determination of analytic approach could be made.

Other study characteristics. Additional characteristics of study methods were coded, including whether the study was described as a pilot study, and whether the report included a CONSORT-type flow chart. We also recorded the number and timing of post-randomization points of data collection, and the number of study sites.

Intervention characteristics. The specialty area of each intervention was coded, with up to two codes allowed for studies that cut across specialty boundaries, such as an intervention for children with cancer. The specialty areas were pediatrics; geriatrics; obstetrics/gynecology; oncology; cardiovascular; critical care; anesthesia; other medical/surgical; psychiatry; community health; health promotion; and nursing education/clinical practice. The length of the intended intervention was also recorded.

Researcher and participant characteristics. We recorded how many authors were listed on the report, and coded whether or not the research team had received funding for the study, either from a government sponsor or from another source. We categorized intervention recipients in terms of whether they were patients or clients, caregivers or family members of patients or clients, or nurses or other health care staff. We also recorded the participants' mean age and their country of residence.

Journal characteristics. We retrieved information about the journal's 2007 impact factor from Thomson's *Journal Citation Reports.* Journals in our sample whose impact score was not calculated were assigned an impact factor value of zero, under the

assumption that their score would be lower than the lowest-rated journal in the nursing subset, which in 2007 was .216. Journals were coded as to whether they had adopted the CONSORT guidelines. This information was obtained through scrutiny of the journal's guidelines to authors and through email communication with journal editors.

Intercoder Reliability

A detailed codebook was developed to enhance reliability of coding. The two authors independently coded 15 articles, and then met to discuss their coding decisions, after which the codebook was further refined. An additional 20 randomly selected studies were coded by both researchers. Interrater agreement on the 75 coded variables ranged from 85% to 100%, with a median agreement of 95%. Coding discrepancies were resolved, and in subsequent coding by a single author, second opinions were sought if there were ambiguities.

Data Analysis

The Statistical Package for the Social Sciences (SPSS) version 16.0 software was used for all data analysis. Descriptive statistics (primarily means, medians, and percentages) were used to describe characteristics of the studies in the sample. Crosstabulations were done to compare stated use of ITT against our assessments of the actual use of ITT. For indicators of special importance, such as percentage of studies using ITT, 95% confidence intervals around the estimate were constructed.

Logistic regression analyses were performed to assess whether factors that have been found to be predictive of ITT use in medical studies were related to the use or declared use of ITT in the nursing literature. The predictors in the models were: the impact factor of the journal in which the study was published, whether the journal was a specialty or generalist journal, whether the journal had endorsed the CONSORT guidelines, whether the study had received funding, and sample size. We also included a variable not previously studied, the number of authors on the research team.

Results

Sample Characteristics

In the sample of 124 RCT studies, just over half (54.0%) were published in journals that had adopted the CONSORT guidelines (6 of the 16 journals had done so), and the median journal impact factor was 1.30. The studies were undertaken in 19 countries, with the highest percentages done in the United States (41.9%), Taiwan (12.1%), Canada (8.1%), and the UK (6.5%). Nursing specialties that were especially well represented included cardiovascular nursing (20.2%), oncology (19.4%), gerontology (15.3%), and pediatrics (11.3%).

The length of the nursing interventions ranged from less than one day to a full year, with the median length of time being 28 days. Nearly half (47.5%) were interventions that were longer than one month duration.

Patients/clients were the intervention recipients in the vast majority of the studies (89.5%). Five interventions (4.0%) were designed for caregivers or family members of patients, and 8 (6.5%) were for nurses, nursing students, or other health care staff. The median age of participants was 57.0 years.

Funding of some type was reported by 71.8% of the studies, and about half of the funded studies had grants from government agencies. Table 2 describes other characteristics of the studies, separately for those with different funding profiles. In this sample of nursing studies, a total of 16,773 people had been randomly assigned to different groups. The median baseline sample size was 76.0, ranging from a median of 66.0 for studies without government funding to 100.0 for those with such funding. Only 37.4% of all studies had an initial sample size of more than 100 people. The length of time between baseline and final follow-up ranged from 1 day to 1 year, with the median substantially longer in government-funded studies (84.0) than in other-

funded studies (24.5) and unfunded studies (21.0). Overall, 12.1% of the studies were described as pilot or feasibility studies.

Attrition

For the sample as a whole, the percentage of cases lost between random assignment and the final follow-up ranged from 0.0% to 62.6%, with the median being 9.2% missing cases. As shown in Table 3, 26.3% of the studies had no attrition, but most of these studies (90.6%) were situations we described as having "captive audiences" of participants—that is, situations in which there was virtually no opportunity to leave the study. In a full 25.4% of these RCTs, the rate of attrition exceeded 20% of those randomized.

As a result of attrition, the median overall sample size at the end of the study was 66.0, and the median per group was 30.0 for both experimental and control groups. The rate of attrition was similar in experimental groups (median = 8.1%) and control groups (median = 9.4%).

Not surprisingly, the rate of attrition was correlated with length of time between random assignment and the final collection of outcome data ($r = .28$, $p = .002$). Among studies with 0% attrition, the median length of time between randomization and final follow-up was 1 day, whereas among those with more than 20% attrition, the median time to final follow-up was 92 days.

Efforts to minimize attrition were infrequently mentioned, although because of page constraints in journals, the absence of mentioning such efforts does not necessarily mean that they were not made. Among the 90 studies with any attrition, 16.7% ($n = 15$) reported strategies designed to keep participants in the study.

Among the same subset of 90 studies with some attrition, only 36 (40.0%) included a CONSORT-type flowchart that documented how and when participants were lost. Inclusion of such a flowchart was substantially more likely in reports

published in journals that had endorsed the CONSORT guidelines (52.1%) than in those published in other journals (26.2%). Nevertheless, 23 studies that had some attrition lacked a flowchart despite having been reported in a CONSORT-endorsing journal. In all, 17 studies with greater than 20% attrition were missing a flowchart.

Only 2 studies in this sample tested whether rates of missingness were significantly higher in the experimental group or in the control group. An attrition bias analysis, that is, a comparison of baseline characteristics of dropouts versus completers, was reported in 18 studies, which is 20.2% of the studies with attrition. None of the articles mentioned that the researchers had examined *patterns* of missingness, such as missing completely at random (MCAR) or missing at random (MAR) (McCleary, 2002; Polit, 2010).

Methods of Handling Missing Outcome Data

Among the 90 studies with some attrition, 76 (84.4%) used either listwise or pairwise deletion of cases in analyzing program effects on outcomes. In other words, in these 76 studies, cases with missing outcome values were simply dropped and ignored. In 11 studies (12.2%), missing values were either imputed or addressed through multi-level modelling. Five of the studies used last observation carried forward, and 5 used modelling. The method used could not be determined in one study. None of the studies imputed missing values using state-of-the-art missing values approaches such as expectation-maximization imputation or multiple imputation (Polit, 2010; Wood et al., 2004). Only one study reported that the researchers had done a sensitivity analysis to assess the impact of imputations on their conclusions.

Intention to Treat

In the full sample of 124 studies, 18 articles (14.5%) reported that an ITT approach was used in the analysis, and one reported using a modified ITT; the 95% CI around the 15.3% of studies that mentioned ITT is 9.1% to 21.8%. No researchers

stated that they used a per protocol analysis, although two noted that they did *not* use ITT. The vast majority of reports (82.3%) were silent with regard to the issue of ITT.

Among the 19 studies that stated they had used ITT or modified ITT, only 8 explained their definition of ITT. For example, Artinian and colleagues (2007) stated, "Analysis of intention to treat was conducted to preserve the baseline comparability between groups achieved by random assignment. Because deviation from the original randomized groups can contaminate the intervention comparison, participants were analysed according to the assigned intervention (thus ignoring nonadherence to intervention protocol and withdrawal), not the actual intervention received" (p. 317). More typically (N = 11), reports simply declared adherence to ITT. For example, the report by Perry and colleagues (2007) noted that, "Intention-to-treat analysis was used" (p. 307). Of the studies that reported using ITT, 8 of them (42.1%) handled missing values by listwise or pairwise deletion.

According to our own assessments of the type of analysis used, far more nurse researchers used ITT than claimed its use. We classified 46.8% of the nursing RCTs as using either true or modified ITT (95% CI, 38.2% to 55.8%). "True" ITT was used in 35.5% of the studies; 25.8% of the studies were classified as using true ITT because there was no attrition, and another 9.7% were classified as true ITT because missing cases were accounted for in the analysis, through imputation or modelling. An additional 11.3% of the studies met our definition for a modified ITT.

Table 4 shows a crosstabulation between researchers' stated use of ITT and our categorization. Of particular note, two studies (10.5% of those claiming an ITT analysis) said they used ITT when, in fact, their approach was more appropriately described as a per protocol analysis. Yet, researchers in 41 studies who did not declare the use of ITT (39.0%) could have done so legitimately.

The journal with the highest percentage of studies that actually used ITT was *Nursing Research*—5 of the 6 RCT studies in that journal (83.3%) actually used ITT, and 4 of the 6 reported adherence to ITT principles.

Factors Associated with Stated and Actual ITT

Logistic regression analyses were performed to assess the effect of journal characteristics, funding, number of authors, and sample size on the declared and actual use of ITT. The overall logistic model for predicting stated use of ITT or modified ITT was significant (chi square goodness of fit = 36.22, *df* = 7, *p* < .001), and the Hosmer-Lemeshow test was nonsignificant (*p* = .294). Two predictors were significant—the journal's impact factor (OR = 5.23, 95% CI = 1.39, 19.64) and the number of authors on the paper (OR = 1.49, 95% CI = 1.08, 2.06). The greater the number of authors, and the higher the impact factor, the more likely the researchers were to state that they had used ITT. In sharp contrast, the model for predicting *actual* use of ITT was nonsignificant (chi square goodness of fit = 2.27, *df* = 7, *p* = .94), and none of the predictors was significant or even approached significance.

Discussion

Gravel and colleagues (2007), in their analysis of ITT usage in medical journals, asked questions that are relevant to our analysis: Are authors saying what they do, and are they doing what they say? In contemporary nursing literature, the answer is "no" to both questions. A few researchers are taking inappropriate credit for using ITT, and many who have done an ITT analysis are failing to inform readers that they pursued a "gold standard" approach to analysis. We think that researchers who use a strong randomized design for testing intervention effects should use ITT analysis, should state that they have used it, and should define what they mean when they claim ITT adherence, given alternative ways of defining the term.

We also encourage readers to scrutinize RCT reports carefully in drawing conclusions about the study's evidence, because it is risky to rely on authors' declaration of ITT adherence. Readers should examine CONSORT-type flowcharts, when they are provided, to learn how and when subjects withdrew from a study. Unfortunately, many articles in our analysis, including ones published in journals that have adopted the CONSORT guidelines, did not present such a flowchart. Hopefully, reviewers and editors will increasingly come to demand such flowcharts whenever there is attrition.

The development of the CONSORT guidelines reflected concern about inadequacies in reporting key elements of study design, a concern with special significance in an environment that is increasingly focused on the use of high quality evidence in clinical practice. The CONSORT guidelines advocate an ITT analytic approach because removal of subjects post-randomization can threaten the internal validity of an RCT. Indeed, there is ample evidence that non-use of ITT leads to biased estimates of treatment effectiveness, generally in the direction of Type I errors and overestimates of effect size (Lachin, 2000; Tierney & Stewart, 2005; Porta, Bonet, & Cobo, 2007). Our finding that ITT was used in fewer than half of nursing RCTs thus suggests that effects in many nursing intervention trials may be inflated.

We found fairly high rates of attrition in this sample of RCTs, with 46% of the studies having greater than 10% attrition. Missing outcome data are likely inevitable, but clearly it is a problem to which greater attention needs to be paid in designing and implementing studies. The use of state-of-the-art analytic strategies for addressing missing values was used in only a handful of studies. The topic of missing values has had tremendous conceptual and mathematical advances in the past few decades. Powerful software to diagnose patterns of missingness and to impute missing values is

now available in popular user-friendly software, such as in the Missing Values Analysis program of SPSS, and its use should be considered in nursing trials.

The logistic regression analyses were more predictive of stated use of ITT than of actual use, which may suggest that there is greater awareness of the desirability of *stating* ITT use than of how ITT is actually defined. Researchers publishing in journals with high impact factors were more likely to state that ITT was used, perhaps because those writing and reviewing for these high-impact journals are more knowledgeable about standards of rigor, such as those expressed in the CONSORT guidelines. We are unsure about how to interpret the finding that research teams with more authors had a greater likelihood of stating ITT use, but one possibility is that large teams were more likely than smaller ones to include statisticians. Yet, number of team members did not predict *actual* use of ITT, perhaps because many studies that we categorized as true ITTs were ones without attrition that did not require a sophisticated missing values strategy.

This study should be considered an early benchmark of the status of ITT in nursing intervention research. As understanding of and adherence to the CONSORT guidelines becomes more widespread, the use of ITT is likely to grow.

Study Limitations

Our sample of studies is likely not to be representative of all RCTs by nurse researchers published in 2007-2008. In particular, nurse researchers often publish papers in medical and interdisciplinary journals, in part because of collaboration with researchers from other disciplines, and in part because of a desire to publish in journals with higher impact factors than those in the nursing subset. It is possible, and perhaps likely, that the rate of stated and actual use of ITT by nurse researchers who publish in such journals is higher than what we found.

Within the English-language nursing literature, however, our sample of RCTs was large and broadly drawn. Our approach to selecting a sample of RCT studies was not exhaustive, but we hand-searched a larger number of journals than has typically been done in similar studies. This includes the work of Gravel and colleagues (2007, 10 journals), Wood and colleagues (2004, 4 journals), Hollis and colleagues (1999, 4 journals), Adetugmo and Williams (2000, 1 journal), and Smith and colleagues' (2008), who analyzed RCT reports from four nursing journals. Given that our sample of journals included every major generalist research journal in nursing, we are reasonably confident that our findings do not exaggerate the relatively low use of ITT in the nursing literature, nor the need for improvement in addressing the risk of bias resulting from attrition in RCTs.

References

Adetugbo, K., & Williams, H. (2000). How well are randomized controlled trials reported in the dermatology literature? *Archives of Dermatology, 136* (3), 381-385.

Altman, D. G., Schulz, K. F., Moher, D., Egger, M., Davidoff, F., Elbourne, D., et al. (2001). The revised CONSORT statement for reporting randomized trials: Explanation and elaboration. *Annals of Internal Medicine, 134*(8), 663-694.

Artinian, N., Flack, J., Nordstrom, C., Hockman, E., Washington, O., Jen, K., et al. (2007). Effects of nurse-managed telemonitoring on blood pressure at 12-month follow-up among urban African Americans. *Nursing Research, 56*(5), 312-322.

Baron, G., Boutron, I., Giraudeau, B., & Ravaud, P. (2005). Violation of the intent-to-treat principle and rate of missing data in superiority trials assessing structural outcomes in rheumatic diseases. *Arthritis & Rheumatism, 52* (6), 1858-1865.

Bradford-Hill, A. (1961). *Principles of medical statistics.* New York: Oxford University Press.

DeMets, D. L. (2004). Statistical issues in interpreting clinical trials *Journal of Internal Medicine, 255,* 529-537.

Gravel, J., Opartny, L., & Shapiro, S. (2007). The intention-to-treat approach in randomized trials: Are authors saying what they do and doing what they say? *Clinical Trials, 4*, 350-356.

Higgins, J. P. T., & Green, S., eds. (2008). *Cochrane handbook for systematic reviews of interventions,* version 5.0.1. London: John Wiley & Sons, Ltd.

Hollis, S. & Campbell, F. (1999). What is meant by intention to treat analysis? Survey of published randomised controlled trials. *BMJ, 319*, 670-674.

Hulley, S., Cummings, S., Browner, W., Grady, D., & Newman, T. (2001). *Designing clinical research* (2nd ed.). Philadelphia: Lippincott Williams & Wilkins.

Huwiler-Muntener, K., Juni, P., Junker, C., & Egger, M. (2002). Quality of reporting of randomized trials as a measure of methodologic quality. *Journal of the American Medical Association, 287* (21), 2801-2804.

Kruse, R., Alper, B., Reust, C., Stevermer, J., Shannon, S., & Williams, R. (2002). Intention-to-treat analysis: Who is in? Who is out? *Journal of Family Practice, 51*(11), 969-973.

Lachin, J. M. (2000). Statistical considerations in the intent-to-treat principle. *Controlled Clinical Trials, 21* (3), 167-189.

McCleary, L. (2002). Using multiple imputation for analysis of incomplete data in clinical research. *Nursing Research, 51,* 339-343.

Mills, E., Wu, P., Gagnier, J., & Devereux, P. (2005). The quality of randomized trial reporting in leading medical journals since the revised CONSORT statement. *Contemporary Clinical Trials, 26,* 480-487.

Perry, C. K., Rosenfeld, A., Bennett, J., & Potempa, K. (2007). Promoting walking in rural women through motivational interviewing and group support. *Journal of Cardiovascular Nursing, 22* (3), 304-312.

Polit, D. F. (2010). *Statistics and data analysis for nursing research* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.

Polit, D. F., & Beck, C. T. (2008). *Nursing research: Generating and assessing evidence for nursing practice* (8th ed.). Philadelphia: Lippincott Williams & Wilkins.

Porta, N., Bonet, C., & Cobo, E. (2007). Discordance between reported intention-to-treat and per protocol analysis. *Journal of Clinical Epidemiology, 60*, 663-669.

Rios, L., Odueyungbo, A., Moitri, M., Rahman, M., & Thabane, L. (2008). Quality of
reporting of randomized controlled trials in general endocrinology literature.
Journal of Clinical Endocrinology & Metabolism, 93(10), 3810-3816.

Ruiz-Canela, M. (2000). Intention to treat analysis is related to methodological quality.
BMJ, 320, 1007.

Smith, B., Lee, H., Lee, J., Choi, M., Jones, D., Bausell, R., et al. (2008). Quality of
reporting randomized controlled trials (RCTs) in the nursing literature:
Application of the Consolidated Standards of Reporting Trials (CONSORT).
Nursing Outlook, 56, 31-37.

Tierney. J. F., & Stewart, L. (2005). Investigating patient exclusion bias in meta-
analysis. *International Journal of Epidemiology, 34*(1), 79-87.

United States Food and Drug Administration (1998). International conference on
harmonization: Guidelines on statistical principles for clinical trials. *Federal
Register, 62* (179), 49583-49598.

Whittaker, K., Sutton, C., & Burton, C. (2006). Pragmatic randomised controlled trials
in parenting research: The issue of intention to treat. *Journal of Epidemiologic
& Community Health, 60,* 858-864.

Wood, A. M., White, I. R., & Thompson, S. (2004). Are missing outcome data
adequately handled? A review of published randomized controlled trials in
major medical journals. *Clinical Trials, 1,* 368-376.

Figure 1
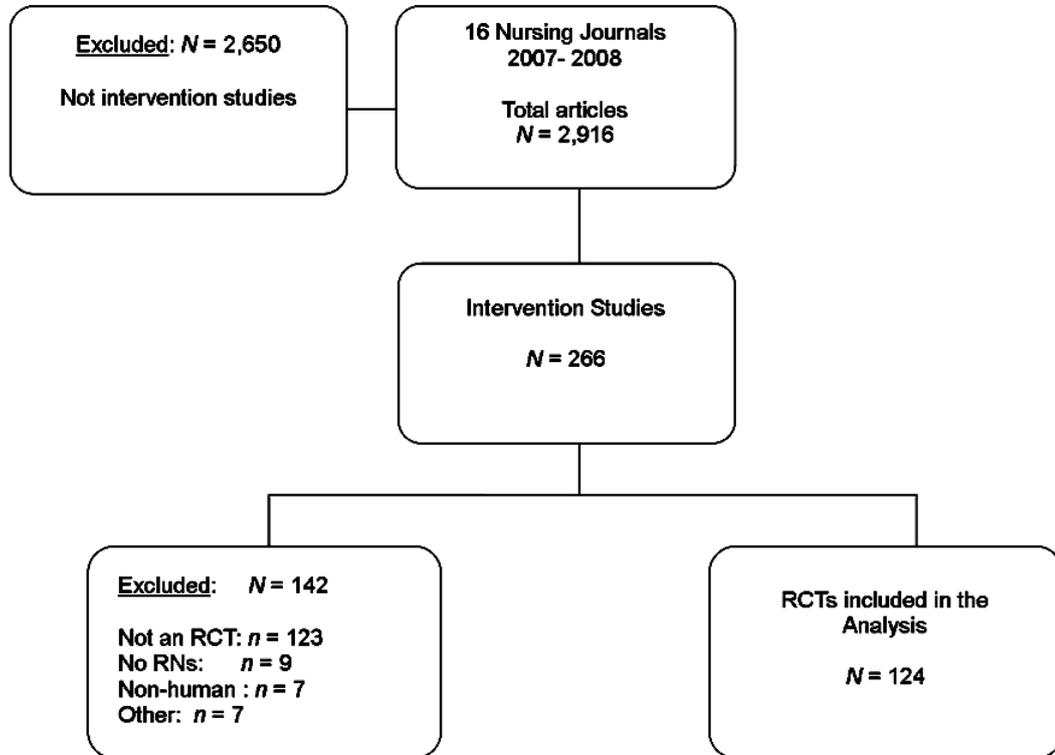**Flow Chart of Studies in the Sample**

Excluded: *N* = 2,650

Not intervention studies

16 Nursing Journals
2007- 2008

Total articles
N = 2,916

Intervention Studies

N = 266

Excluded: *N* = 142

Not an RCT: *n* = 123
No RNs: *n* = 9
Non-human : *n* = 7
Other: *n* = 7

RCTs included in the
Analysis

N = 124

Table 1
RCTs in the Study Sample, by Journal

Journal	Number of Articles, 2007-2008	Number of Intervention Studies	RCTs Included in Study Sample N (%)	Impact Factor, 2007[a]
AANA Journal	79	17	10 (8.1)	--
American Journal of Critical Care	87	6	3 (2.4)	1.078
Applied Nursing Research	63	9	6 (4.8)	0.774
Cancer Nursing	164	16	8 (6.5)	1.262
European Journal of Cardiovascular Nursing	90	9	4 (3.2)	--
International Journal of Nursing Studies	305	28	11 (8.9)	2.115
Journal of Advanced Nursing	537	28	14 (11.3)	1.442
Journal of Cardiovascular Nursing	116	8	6 (4.8)	--
Journal of Clinical Nursing	644	63	23 (18.5)	1.301
Journal of Gerontological Nursing	133	9	4 (3.2)	--
Journal of Pediatric Nursing	94	11	4 (3.2)	--
Nursing Research	119	8	6 (4.8)	1.748
Oncology Nursing Forum	148	16	11 (8.9)	1.438
Public Health Nursing	126	16	3 (2.4)	0.559
Research in Nursing & Health	111	7	4 (3.2)	1.000
Western Journal of Nursing Research	100	15	7 (5.6)	0.848
Total	2,916	266	124 (100%)	

[a]**Source: Thomson's *Journal Citation Reports***

Table 2

Characteristics of the RCT Studies in the Sample, by Funding Category

Study Characteristic	Received government funding (*N* = 43)	Received other funding (*N* = 46)	No declared funding (*N* = 35)	All studies (*N* = 124)
Number of participants, median	100.0	66.0	66.0	76.0
$\leq$ 50 participants, N (%)	11 (25.6)	14 (30.4)	10 (29.4)	35 (28.5)
51-100 participants, N (%)	11 (25.6)	15 (35.7)	16 (47.1)	42 (34.1)
> 100 participants, N (%)	21 (48.8)	17 (37.0)	8 (23.5)	46 (37.4)
Number of study sites, range	1 – 19	1 – 31	1 – 14	1 – 31
More than 1 study site, N (%)	15 (34.9)	13 (28.3)	9 (25.7)	37 (29.8)
No. of post-RA measurements, range	1 – 5	1 – 14	1 – 5	1 – 14
More than 1 measurement, N (%)	20 (46.5)	21 (45.7)	10 (28.6)	51 (41.1)
Time to last follow-up (days), median	84.0	24.5	21.0	38.5
1 day, N (%)	4 (9.3)	10 (21.7)	11 (31.4)	25 (20.2)
2 – 30 days, N (%)	9 (20.9)	17 (37.0)	9 (25.7)	35 (28.2)
31 – 90 days, N (%)	12 (27.9)	6 (13.0)	8 (22.9)	26 (21.0)
> 90 days, N (%)	18 (41.9)	13 (28.2)	7 (20.0)	38 (30.7)
Described as a pilot study, N (%)	7 (16.3)	5 (10.9)	3 (8.6)	15 (12.1)

Table 3

Selected Attrition-Related Findings in Nursing RCTs

Attrition-Related Finding	*N*	Percent
Amount of attrition (subject loss between randomization and final follow-up)		
0 % missing, "captive audience"[a]	29	23.8%
0% missing, not captive audience	3	2.5%
1% - 10% missing	34	27.9%
11% – 20% missing	25	20.5%
21% - 30% missing	15	12.3%
More than 30% missing	16	13.1%
Median percentage missing, by time between random assignment and final follow-up		
≤ 1 day	25	0.0%
2 to 30 days	35	8.1%
31 to 90 days	35	12.5%
3 months to 6 months	27	15.0%
More than 6 months	10	18.0%
Article reported effort to minimize attrition		
All studies (*N* = 124)	17	13.7%
Studies with any attrition (*N* = 90)	15	16.7%
Included a CONSORT-type flow chart		
All studies (*N* = 124)	40	32.3%
Studies with any attrition (*N* = 90)	36	40.0%
Reported an attrition bias analysis comparing completers and dropouts, studies with any attrition (*N* = 90)	18	20.2%

[a]We classified a study as having a "captive audience" of subjects if participants had essentially no opportunity to leave the study; most were studies conducted in hospital settings or involved data collection immediately following a brief intervention.

Table 4

Stated Use of ITT Versus Actual Use in Nursing RCTs

Classification of actual approach	Approach declared in article		Total
	Declared as ITT or modified ITT (N = 19)	Not declared as ITT (N = 105)	(N = 124)
True ITT, no missing values	1 (5.3%)	31 (29.5%)	32 (25.8%)
True ITT, imputation or modelling	10 (52.6%)	2 (1.9%)	12 (9.7%)
Modified ITT	6 (31.6%)	8 (7.6%)	14 (11.3%)
Per protocol	2 (10.5%)	50 (47.6%)	52 (41.9%)
Could not determine if modified ITT or per protocol	0 (0.0%)	14 (13.3%)	14 (11.3%)