



# Randomized Controlled Trials

## Methodological Concepts and Critique

Susan S. Salmond

Randomized controlled trials, also known as true experiments or intervention studies, are considered to be the gold standard research design for demonstrating a cause-and-effect relationship between an intervention and an outcome. This article will describe key methodological concepts that make the randomized controlled trial this gold standard. Practicing from an evidence-based perspective requires practitioners to be able to critique the strengths and weaknesses of a research study in order to make decisions about adoption of the intervention into one's practice area. Key components to the critique of a randomized controlled trial are defined and the process is illustrated by a critique of Gallo and colleagues article, *A Study of Naloxone Effect on Urinary Retention in the Patient Receiving Morphine PCA*, published in this issue.

**R**andomized controlled trials or true experiments constitute the most rigorous research design for intervention studies. In an intervention study, the investigator is interested in determining whether a cause-effect relation exists between treatment and outcome(s) (Sibbaid, & Roland, 1998; Stommel & Wills, 2004). To achieve this, there must be comparison groups where individuals in each group receive a different level of exposure to the intervention. The outcomes can then be compared to determine whether the intervention makes a difference.

This article will highlight the key principles underlying intervention studies and then examine components of methodological rigor with an accompanying critique of Gallo, DuRand, and Pshon's (2008) study of the effect of naloxone on urinary retention in the patient receiving morphine via patient-controlled analgesia (PCA) to illustrate the key components. In the study, Gallo et al. were interested in showing a cause-effect relation between administration of low-dose naloxone and complications of urinary retention. The study is a single factor, posttest-only experimental design as illustrated in Figure 1. The target population consisted of orthopaedic surgical patients who were randomly assigned to two groups: the experimental or intervention group and the control group. The intervention group received naloxone in conjunction with PCA morphine and the control group received PCA morphine with no

naloxone. It was hypothesized that administration of low-dose intravenous naloxone would decrease the incidence of urinary retention, need for catheterization, and hospital length of stay. The researchers identified appropriate outcome (dependent) variables and indicated objective approaches to measure these outcomes.

### Control: The Guiding Principle in Randomized Controlled Trial Design

Randomized controlled trials are considered the "gold standard" for intervention studies, as this design introduces the notion of *control*, thereby minimizing bias. Control is a systematic process of design that isolates the effect of the intervention being tested by ruling out other potential "causes" of the effects under study. Other causes, also called rival causes, are referred to as confounding variables in that they provide an alternative explanation for the effect on the dependent variable.

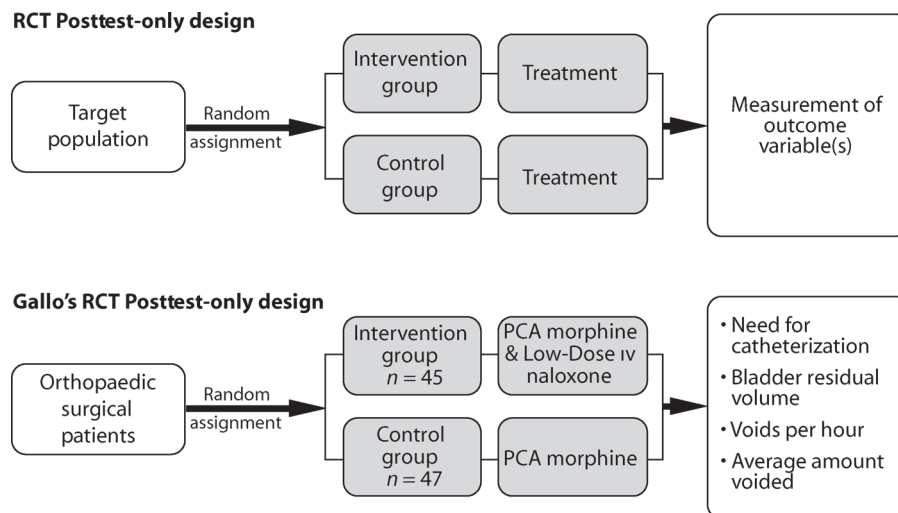
Planning for the design of a study requires identification of confounding factors that could create rival hypotheses so that these factors can be appropriately controlled for in the design. A thorough literature search on the selected dependent variable assists in identifying potential confounding factors. In the study by Gallo et al., a review of the literature on urinary retention and morphine PCA and low-dose naloxone would be called for. The results of this often appear in the review of the literature section of the manuscript or may not be evident except through examination of the research design.

Review of the empirical literature links postoperative urinary retention to various factors including a history of urinary problems, advanced age (especially over the age of 70), male gender, the total amount of fluid replacement over a 24-hr postoperative period, type of anesthesia (greater incidence with epidural analgesia), pain management medications and dosage (greater

---

Susan S. Salmond, EdD, RN, Interm Dean, School of Nursing, University of Medicine and Dentistry of New Jersey, Newark, NJ.

The author has no significant interest, financial or otherwise, to any company that might have an interest in the publication of this educational activity.



**FIGURE 1.** Randomized controlled trial (RCT) posttest-only design. *Note.* iv = intravenous; PCA = patient-controlled analgesia.

with narcotic analgesia), and administration with PCA (Fernandes, Carmo, Costa, & Saraiva, 2007; Koch, Grinberg, Gary & Farley, 2000; O’Riordan, Hopkins, Ravenscroft, & Stevens, 2000; Sarasin, Walton, Singh, & Clark, 2006; Wynd, Wallace, & Smith, 1996). Gallo and colleagues accounted for these variables in their research design. Sampling criteria excluded individuals over age 65 and those with a preexisting urinary condition. Random assignment was used which would likely distribute the other variables, and the researchers compared the groups for equivalence.

## Randomization: The Process for Achieving Control

A valid comparison of the intervention and control groups requires that the two groups be as alike as possible on the critical demographic and clinical variables that could pose alternative hypotheses for the outcome effect. Moreover, in all intervention studies, there may be confounding effects from factors that the researchers are not even aware of. Randomization is the preferred approach for equalizing conditions and differences of these potentially confounding factors across the different arms of treatment (Melnik & Fineout-Overholt, 2005). Through randomization, the two groups (experimental and control) are as similar as possible, with the major difference remaining between them being their exposure to different treatments (Stommel & Wills, 2004). Randomization eliminates both conscious bias and unconscious bias associated with the selection of a treatment for a given patient (Stanley, 2007). For example, without randomization, it could be possible to bias the results by assigning patients with anticipated “noncomplicated” surgical procedures to the treatment group. These patients would likely have shorter operative times and require less narcotic analgesia postoperatively; therefore fewer urinary retention complications. In this example, failure to randomize would predispose to imbalances in critical baseline characteristics between the

two study groups, thus introducing selection bias and generating rival hypotheses for the occurrence of decreased urinary retention.

Kerlinger and Lee (2000) emphasize the critical importance of control in intervention studies to ensure the equivalence of the comparison groups on all factors but the intervention variable itself. They indicate that the best approach to achieving this control is to *randomize whenever possible*. Nonrandomized studies have been reported to overestimate or underestimate treatment effects (Bhandari & Haynes, 2005).

Randomization can be accomplished by selecting participants at random or assigning participants to treatment groups at random. It is not always possible to select participants at random because researchers have access to limited populations and selecting random participants significantly increases the time needed to obtain an adequate sample size. When random selection is not used it is important to randomly assign participants to either the treatment or control group. If treatment groups receive the treatment at different times and with different investigators, times and investigators can be randomly assigned as well.

An additional protection against bias during randomization is *concealment of allocation*. Allocation is concealed when investigators enrolling subjects cannot determine in advance the treatment assignment of the next patient to be enrolled into the study. This prevents possible selection bias whereby the investigator can systematically allocate patients to the intervention or control group on the basis of some subjective perhaps unconscious bias. Use of a random-numbers table successfully conceals allocation and prevents biased assignment or response to assignment awareness.

## Blinding

Blinding is another methodological safeguard to ensure that individuals (participants, treatment team members, or data collectors/outcome assessors) are unaware of the treatment (intervention or control) that has been

received/administered/assessed until the study is complete so that differential treatment or a biased assessment of outcome (detection bias) is reduced or eliminated (Licciardone & Russo, 2006). Participants or the clinical team/research team that knows the group assignment may respond differently on the basis of group assignment, thereby biasing the results. Single blinding is when the participant is blinded to which treatment (intervention or control) is being administered. In double blinding, neither the participants nor the data collectors/observers know which subjects are assigned to treatment and control conditions. In triple-blind studies, none of the study participants, the healthcare professionals treating them, or the persons who observe or record the outcomes are aware of the treatment assignment (Stommel & Wills, 2004). Lack of blinding has been associated with increased or decreased magnitudes of observed treatment effects, especially when differences between treatments are small (Poolman et al., 2007). Gallo and colleagues did not use any blinding in their design, thus introducing the possibility of bias, especially at the point of outcomes measurement.

## Critique of Randomized Controlled Trials

Although experimental designs/randomized controlled trials are the most rigorous research design, it cannot be assumed to be of good quality. The reader of any intervention study needs to critique the study of interest to make conclusions as to the rigor of the particular study and its applicability for use in the reader's own clinical setting. Table 1 summarizes key components to be evaluated and critiques the study by Gallo et al. on these methodological components.

### CLEAR RESEARCH DESIGN

The first group of methodological factors relate to the clarity and appropriateness of the research design. The dependent/outcome variables should be clearly identified with reliable and valid approaches to measuring the variables specified. In critiquing the outcome variable, consider whether all clinically important outcomes were measured. Similarly, the treatment arms (in this case control group and experimental group) must be fully described not only to facilitate reproducibility in future studies but also to allow for comparison across studies. Melnyk and Fineout-Overholt (2005) point out that in some intervention studies, there is a significant difference in the time spent and the degree of patient/researcher interaction between the experimental and control groups, leading to the potential of an effect because of the differential attention factor. In critiquing the treatment arms, evaluate the control group treatment for its similarity with the intervention treatment. Sampling criteria and sample size are important to evaluate in the research design. What are the inclusion and exclusion criteria? Has a power analysis been done in advance to estimate the needed sample size? A small sample may lack the power needed to demonstrate a treatment effect and result in a false negative. This is known as a type II error—the researcher accepts the null hypothesis (that no difference

exists) when it should have been rejected (as a real treatment effect exists). Most treatments offer a small additional benefit and therefore require larger samples. Too small a sample could mean that no benefit is detected even though there is actually a small positive effect.

### APPROACHES FOR ENSURING VALIDITY OF THE FINDINGS

The key question in internal validity is whether the observed changes can be attributed to the intervention (i.e., the cause) and not to other possible causes ("alternative explanations" for the outcome). Processes used to control for alternative explanations include randomization, concealment of allocation, and blinding.

Ensuring validity requires attention to follow-up assessment, loss of participants during the study, the similarity and differences in the control and intervention group, and whether participants were in fact analyzed within the group to which they were originally assigned. Follow-up assessments need to be conducted long enough to study the effects of the intervention. An educational intervention that tests only after implementation of the education does not have an adequate follow-up to determine the extent of the effect of the educational intervention.

Examine the original sample size and compare it with the sample size used for data analysis to determine the mortality rate. The percentage of subjects who did not complete the study should not exceed 20%. In cases of mortality and noncompliance, a statistical approach to preserving randomizations is to ensure that participants are analyzed in the groups to which they were originally assigned—a process known as intention to treat. By specifying how to handle noncompliant patients and analyzing their data in the groups they were originally allocated to, randomization is preserved.

Group equivalence is important in assuring that the treatment, not difference among the groups, caused the effect. With small sample sizes, the intended effect of randomization may not be achieved. In addition, subject mortality may disrupt group equivalence. Examination of the statistical equivalence or distribution of key factors between groups is done to assure the effectiveness of randomization or the statistical equivalence across groups.

### DETERMINING THE MEANINGFULNESS OF RESULTS

Intervention studies should report both statistical significance and clinical relevance of the findings. This can be accomplished by reporting on the magnitude or size, the precision, and the clinical meaningfulness of the findings.

In an intervention study, the effect size indicates how effective the experimental treatment was. Effect size is calculated by subtracting the mean of the control group from the mean of the experimental group and dividing the resultant by the pooled standard deviation. Small, medium, and large effects are designated as 0.2, 0.5, and 0.8, respectively. As effect sizes are not dependent on sample size, they are a critical indicator of the magnitude of the experimental intervention (Melnyk & Fineout-Overholt, 2005).

**TABLE 1. KEY COMPONENTS TO THE CRITIQUE OF RANDOMIZED CONTROLLED TRIALS**

Critique components	Review of Gallo et al.'s study
<i>Clear and appropriate research design</i>	
1 Outcome/Dependent Variables	
a. The dependent or outcome variable(s) is clearly defined.	Primary outcome variable Urinary retention Secondary outcome variables Length of stay Pain levels
b. A <i>standardized</i> approach for measuring the outcome variable(s) is identified and allows for clear comparison of findings from the groups.	Urinary output # of patient voids Total voiding amount Need for catheterization Catheterization/no catheterization Catheterization time Amount residual urine Bladder distention Subjective feelings of bladder fullness Bladder ultrasound Length of stay Minutes from time of admission to time of discharge Pain/pain control 0–10 Numeric/visual analog scale Total morphine delivered
c. All clinically important outcomes are measured?	Yes. Clinically important outcomes were measured.
2 Experimental Treatment/Intervention	
a. The intervention(s) or treatment(s) that will be manipulated is clearly identified to ensure reproducibility.	Two groups: Control and intervention Control group: PCA morphine Intervention group: PCA morphine with 0.1 mg iv naloxone every 4 hr
b. The control group intervention is similar in time and interaction with the patient to control for effects secondary to time and attention.	The control group intervention was similar in time and interaction with the patient.
3 Sampling	
a. The target population/sample is clearly specified.	Target population: Orthopaedic patients from one acute care facility undergoing surgery and receiving postoperative morphine PCA.
b. Sampling criteria (inclusion and exclusion criteria) clearly depict the sample. Inclusion–exclusion criteria can be used to control for Confounding variables. For example, by excluding those with a history of urinary problems are controlling for a rival hypothesis.	Inclusion criteria/sample: Ninety-seven orthopaedic patients younger than 65 years scheduled for elective orthopaedic shoulder, hip, or knee surgery and agreeing to participate. Exclusion criteria: History of urinary problems (bladder or prostate cancer, renal dialysis), documented naloxone allergy, or opioid dependency.
c. Can adequate sample size is used. This can be done in advance by performing a power analysis, which guides the researcher on the sample size needed to demonstrate a treatment effect.	No power analysis was performed in advance to specify the sample size needed to detect a moderate difference. This author used G-Power 3 (Faul, Erdfelder, Lang, & Buchner, 2007) to calculate power on the basis of the findings of the study and found the study to be adequately powered for the number of voids per hour (effect size = 0.66, power = 0.92) and the average bladder scan amount (effect size = 0.516, power = 0.78) and underpowered for the average amount voided (effect size 0.280, power = 0.373) and the need for catheterization (effect size 0.3, total sample size suggested = 220).
4 <i>Methods for assuring internal validity of the study</i>	
a. Rigorous approaches to control confounding variables are used. Assess: Were participants randomized to the experimental and control groups? Was group assignment concealed from those enrolling subjects into the groups?	Random sampling was not used. The 97 participants in the original sample were randomly assigned using a computer-generated random-numbers table, thus concealing allocation. This resulted in subgroup sizes of: Control group: 45 Intervention group: 52

(continued)

**TABLE 1. KEY COMPONENTS TO THE CRITIQUE OF RANDOMIZED CONTROLLED TRIALS (Continued)**

Critique components	Review of Gallo et al.'s study
<p>The goal of this approach is to have equivalence across groups on all factors but the intervention variable itself. It does not mean that the groups must be equal in number.</p>	
<p>b. Blinding strategies are used to decrease detection bias. Assess: Were key groups blinded to treatment allocation? Specifically determine whether blinding occurred with:</p> <ul style="list-style-type: none"> <li>Patients</li> <li>Those administering intervention(s)</li> <li>Those scoring outcome(s)</li> <li>Data analysts</li> </ul>	<p>No blinding was used even though blinding would have been possible. It is important to assess to what degree detection bias could be present. No subjective patient measures were used; therefore, patient detection bias is not likely. However, scorer bias could be present, as the scorer is responsible for measuring urinary output and bladder distension and these reports could be skewed toward the observer's bias.</p>
<p>c. Were follow-up assessments are conducted long enough to study the effects of the intervention?</p>	<p>Follow-up assessments were effectively planned until normal urinary function returned</p> <p>Every 4 hr:</p> <ul style="list-style-type: none"> <li>Total morphine delivered</li> <li>No. of patient voids</li> <li>Total voiding amount</li> </ul> <p>Every 8 hr:</p> <ul style="list-style-type: none"> <li>Subjective feelings of bladder fullness</li> <li>Palpation for bladder distention</li> <li>Bladder ultrasound</li> </ul> <p>As indicated:</p> <ul style="list-style-type: none"> <li>Need for catheterization</li> <li>Amount of residual urine</li> </ul>
<p>d. Did all subjects complete the study? (The goal is for at least 80% of subjects to complete the study.)</p>	<p>A total of 97 participants consented to study and 96 completed the study (98%). One participant from the experimental group was lost because of increased pain after administration of naloxone. Data analysis was completed on 45 participants from the control group and 51 from the intervention group. No significant bias results from this.</p> <p>Two subjects from the control group and four from the experimental group did not have ongoing monitoring (no. of voids, amount voided, and bladder retention) because of need for indwelling Foley catheter placement. These subjects were included in the need for catheterization analysis.</p>
<p>e. Were participants analyzed in the groups to which they were randomized? Participants should be analyzed within the group to which they were allocated, irrespective of whether they experienced the intended intervention (intention to treat analysis).</p>	<p>There was one dropout in the intervention group that was not included in the final data analysis according to the assigned group. This one case should not significantly impact the results; however, preference would be to continue with data collection and included the subject in the experimental group even though they did not receive ongoing naloxone administration.</p>
<p>f. Were patients in the treatment and control groups similar with respect to known prognostic factors (key demographic and clinical factors)?</p>	<p>There were no significant differences across the intervention and control groups on gender, procedure type, age, hours on PCA, or hours in study.</p>
<p>5 <i>What are the results of the study and are they important?</i></p> <p>a. Analysis is focused on estimating the size of the difference in predefined outcomes between intervention groups—How large was the treatment effect? This is accomplished by determining the effect size or level of risk reduction.</p> <p><i>Effect size</i> is an estimate of the strength of the treatment. It is calculated by taking the mean of the control group minus the mean of the experimental group and dividing this amount by the pooled standard deviation.</p>	<p>Authors reported <i>p</i> values rather than risk reduction or effect size. Significance was found for average number of voids per hour (higher in the experimental group), average bladder scan amount per hour (lower in the experimental group), and percentage of patients who were catheterized.</p> <p>Calculation of effect size for number of voids, bladder scan amount, and amount voided was not reported by Gallo but was calculated by this author using the formula given on the left. Results showed a medium effect size for number of voids (0.66) and bladder scan amount (−0.516) and a small effect size for amount voided (0.28) and need for catheterization (0.30).</p>

(continued)

**TABLE 1. KEY COMPONENTS TO THE CRITIQUE OF RANDOMIZED CONTROLLED TRIALS (Continued)**

Critique components	Review of Gallo et al.'s study
<i>Relative risk reduction</i> measures how much the risk is reduced in the experimental group compared to the control group. Relative risk reduction would be calculated as Control Event Rate minus Experimental Event Rate divided by the Control Event Rate.	Calculation of risk was not reported by Gallo but was calculated for the variable "need for catheterization" by this author using the formula given on the left. <i>Relative risk reduction</i> : $(0.244 - 0.115)/0.244 = 0.528$ . This means that the need for catheterization was reduced by 52.8% in the treatment group (PCA morphine with naloxone) compared with the control group (PCA morphine only).
<i>Absolute risk reduction</i> is the absolute difference in outcome rates between the control and treatment groups.	<i>Absolute risk reduction</i> : $(0.244 - 0.115) = 0.129$ or 12.9%. This means that for every 100 patients enrolled in the treatment group, about 13 catheterizations would be averted. These calculations support the clinical significance of the findings.
b. Having established the clinical importance of the findings, one must then determine whether the probable treatment benefits are worth the effort. This is calculated by the "number needed to treat" (NNT) and is responsive to the question: Are the likely treatment benefits worth the potential harm and costs? This is calculated by taking the inverse of the absolute risk reduction.	Absolute risk reduction = 12.9%. $NNT = 1/12.9\% = 7.75$ . This means that for every eight patients treated with naloxone, one case of catheterization would be prevented.
c. How precise is the treatment effect? This is determined by confidence intervals of the outcome values. The narrower the confidence interval, the more confident one can be that the study result is the true result.	Confidence intervals were not calculated by the researchers.
6 <i>Relevance to practice</i>	
a. The applicability of the results is assessed by the question: Can the results be applied to my patient?	If the patients in your practice meet the inclusion criteria and do not violate any of the exclusion criteria, then you can apply the results to their care with confidence. In this case, exclusion criteria consisted of patients older than 65 years. So, the results may not be applicable for older patients.
Is the treatment feasible in my clinical setting?	Factors relevant to feasibility include costs of naloxone therapy, which has not been reported.
b. What are the patient's values and expectations regarding the outcome that is trying to be prevented and the treatment itself?	This is not known; however, it is likely that patients would prefer not to be catheterized because of the accompanying risks of such a procedure. The naloxone treatment is not accompanied by any significant discomfort.

Note. PCA = patient-controlled analgesia.

Confidence intervals are used to express the precision and uncertainty of the findings. When taking the findings from a sample, the ultimate goal is to be able to generalize to the larger population itself. No matter how carefully the sample has been selected to be a fair and unbiased representation of the population, relying on information from a sample will always lead to some level of uncertainty. A confidence interval is a range of values that tries to quantify this uncertainty. It can be viewed as a range of plausible values. A narrow confidence interval implies high precision; we can specify plausible values within a tiny range. A wide interval implies poor precision; we can only specify plausible values to a broad and uninformative range.

An important question needed to interpret the results of a clinical trial is whether the measured effect size is clinically important. Three commonly used measures of effect size are relative risk reduction, absolute risk reduction, and the number needed to treat to prevent one bad outcome. These findings put the results into a meaningful interpretation of clinical relevance.

Many studies report only the statistical significance of the findings, which is largely dependent on the sample size and statistical power. The larger the sample, the greater the power and probability of detecting significant differences between groups even when the effect size of the treatment is small. In the study by Gallo et al., the effect size of catheterization was small at 0.30 and statistical significance was not found. When risk and number needed to treat was calculated, it was found that even with a small effect size, the need for catheterization was reduced by 52.8% in the treatment group; or when looking at the absolute risk reduction, the findings indicated that for every 100 patients enrolled in the treatment group, about 13 catheterizations could be averted. With the inherent risks of catheterization, averting 13 catheterizations is clinically meaningful. If one considered only the *p* value, one would conclude that naloxone was not effective in reducing the number of complications. However, looking at the effect size, risk reduction, and number needed to treat, it is apparent that these findings are clinically meaningful.

## RELEVANCE TO PRACTICE

Evidence-based practice requires the integration of individual clinical expertise with the best available external clinical evidence from systematic research, available resources, and our patient's unique values and circumstances. Thus, the final area to assess when critically appraising an intervention study is its relevance to one's own clinical population. Answering questions of clinical relevance assists the practitioner in determining whether the proposed intervention should be adopted in their setting with their specific population. To determine this, compare your practice population with the population in the study. Determine whether the treatment is feasible in your own clinical setting. Finally, determine your patient's values and expectations and whether they are congruent with the proposed treatment and outcome effect.

## Summary

Gallo and colleagues have studied an important topic to orthopaedic nurses—urinary retention in the postoperative orthopaedic surgical patient receiving morphine PCA. A randomized controlled design was the appropriate selection for an intervention study. The research design was clearly described and the outcome measures were objective and clinically relevant. Random assignment and allocation concealment were effectively accomplished and the intervention and control groups were similar on important demographic and clinical characteristics. Blinding was not done, which decreases the validity of the findings.

The authors reported only on statistical significance with findings that naloxone was effective in influencing the number of voids per hour and the average bladder scan amount. Average amount voided and need for catheterization showed differences; however, they were not statistically significant. As the sample size was small, a post hoc power analysis showed the study to be underpowered which would account for the lack of statistical significance. Further analysis by this author in calculating clinical meaningfulness found a moderate effect size for number of voids per hour and the average bladder scan amount, and a small effect size for average amount voided and the need for catheterization. Calculation of risk reduction and number needed to treat revealed the findings to be clinically meaningful in reducing need for catheterization.

Study findings showed that low-dose naloxone can be safely used without compromising pain management and that it may be efficacious in reducing the negative outcome of urinary retention and need for catheterization. Replication of this study should consider three methodological changes: (1) using a preimplementation power analysis to determine necessary sample size; (2) blinding the outcome assessor; (3) and using statistics associated with clinical relevance. Moreover,

patients over the age of 65 should be included, as this is the group that is most vulnerable to urinary retention and it is important to establish the effectiveness of naloxone in this age group.

## REFERENCES

- Bhandari, M., & Haynes, R. B. (2005). How to appraise the effectiveness of treatment. *World Journal of Surgery, 29*(5), 570–575.
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G\*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175–191.
- Fernandes, B., Carmo, M., Costa, V., & Saraiva, R. A. (2007). Postoperative urinary retention: Evaluation of patients using opioid analgesic. *Revista Latino-Americana de Enfermagem, 15*(2), 318–322.
- Gallo, S., DuRand, J., & Pshon, N. (2008). A study of naloxone effect on urinary retention in the patient receiving morphine PCA. *Journal of Orthopaedic Nursing, 27*(2), 111–115.
- Kerlinger, F. N., & Lee, H. B. (2000). *Foundations of behavioral research*. Belmont, CA: Wadsworth.
- Koch, Grinberg, Gary, & Farley. (2000).
- Licciardone, J. C., & Russo, D. P. (2006). Blinding protocols, treatment credibility, and expectancy: Methodological issues in clinical trials of osteopathic manipulative treatment. *Journal of the American Osteopathic Association, 106*(8), 457–463.
- Melnyk, B. M., & Fineout-Overholt, E. (2005). Rapid critical appraisal of randomized controlled trials (RCTs): An essential skill for evidence-based practice (EBP). *Pediatric Nursing, 31*(1), 50–52.
- O'Riordan, J. A., Hopkins, P. M., Ravenscroft, A., & Stevens, J. D. (2000). Patient-controlled analgesia and urinary retention following lower limb joint replacement: Prospective audit and logistic regression analysis. *European Journal of Anaesthesiology, 17*, 431–435.
- Poolman, R. W., Struijs, A. A., Krips, R., Sierevelt, I. N., Marti, R. K., Farrokhyar, F., et al. (2007). Reporting of outcomes in orthopaedic randomized trials: Does blinding of outcome assessors matter? *Journal of Bone & Joint Surgery, 89-A*(3), 550–558.
- Sarasin, S. M., Walton, M. J., Singh, H. P., & Clark, D. I. (2006). Can a urinary tract symptom score predict the development of postoperative urinary retention in patients undergoing lower limb arthroplasty under spinal anaesthesia? A prospective study. *Annals of the Royal College of Surgeons of England, 88*(4), 394–398.
- Sibbald, B., & Roland, M. (1998). Understanding controlled trials: Why are randomised controlled trials important? *British Medical Journal, 316*, 201.
- Stanley, K. (2007). Design of randomized controlled trials. *Circulation, 115*(3), 1164–1169.
- Stommel, M., & Wills, C. E. (2004). *Clinical research: Concepts and principles for advanced practice nurses*. Philadelphia: Lippincott, Williams & Wilkins.
- Wynd, C. A., Wallace, M., & Smith, K. M. (1996). Factors influencing postoperative urinary retention following orthopaedic surgical procedures. *Orthopaedic Nursing, 15*(1), 43–50.