

Comparative Analysis of External Validity Reporting in Non-randomized Intervention Studies

**Noeman A. Mirza, Noori Akhtar-Danesh,
Eric Staples, Lynn Martin, Charlotte Noesgaard**

This article describes a comparative analysis of external validity reporting in non-randomized behavioural and public health intervention studies that used and did not use the TREND (Transparent Reporting of Evaluations with Non-randomized Designs) statement. The search resulted in 14 non-randomized intervention studies that were rated based on Green and Glasgow's criteria for external validity reporting. Studies that used the TREND statement demonstrated improved external validity reporting when compared with studies that did not use the TREND statement. The implication is that the TREND statement and Green and Glasgow's criteria can improve external validity reporting of non-randomized behavioural and public health interventions.

Keywords: TREND statement, external validity, comparative analysis, public health, non-randomized interventions

Analyse comparative d'établissement de rapports de validité externe dans le cadre d'études d'interventions non aléatoires

**Noeman A. Mirza, Noori Akhtar-Danesh,
Eric Staples, Lynn Martin, Charlotte Noesgaard**

Cet article présente une analyse comparative d'établissements de rapports de validité externe dans le cadre d'études d'interventions non aléatoires en matière de comportements et de santé publique faisant usage et ne faisant pas usage de l'énoncé TREND (Transparent Reporting of Evaluations with Non-randomized Designs). La recherche a relevé 14 études d'interventions non aléatoires, lesquelles ont été évaluées selon les critères de Green et Glasgow quant à l'établissement de rapports de validité externe. Les études ayant utilisé l'énoncé TREND ont démontré la présence de rapports de validité externe améliorés, en comparaison avec les études n'ayant pas fait usage de l'énoncé TREND. En conclusion, l'énoncé TREND et les critères de Green et Glasgow peuvent améliorer les rapports de validité externe d'interventions non aléatoires en matière de comportements et de santé publique.

Mots clés : énoncé TREND, validité externe, analyse comparative, santé publique, interventions non aléatoires

Introduction

The past two decades have seen the emergence of several guidelines aimed at enhancing the quality of reports of randomized controlled trials (RCTs), non-randomized experiments, systematic reviews, and meta-analyses. The TREND (Transparent Reporting of Evaluations with Non-randomized Designs) statement is used to enhance the quality of reporting in non-randomized intervention studies. However, its impact on external validity reporting is unclear. Therefore, this comparative analysis is intended to determine whether the use of the TREND statement enhances external validity reporting in non-randomized intervention studies. Both TREND and non-TREND studies are evaluated by external validity criteria recommended by Green and Glasgow (2006). Findings and implications for nurse researchers who are engaged in conducting, reporting, and evaluating studies involving non-randomized interventions are discussed.

Background

Investigators concerned with health promotion engage in clinical research in order to draw inferences from study findings about the nature of their surroundings. To interpret study findings, two sets of inferences are commonly used. The first, known as *internal validity*, is the extent to which correct conclusions are drawn about what actually happened in an experiment, while the second, *external validity* (i.e., *generalizability*), is the extent to which the findings can be applied to situations outside the experiment (Hulley, Cummings, Browner, Grady, & Newman, 2006). To enable an accurate interpretation of findings, a study must first have strong internal validity, which is achieved through a strong relationship between its research operations built upon good choice of study design, outcome measurement, and representative sampling. It is for this reason that researchers and journals give precedence to internal validity and scientific rigour instead of generalizability of findings (Ferguson, 2004). This practice jeopardizes translation of research into practice in applied disciplines such as medicine, public health, and nursing, which are concerned with health promotion and improving the health of the public (Steckler & McLeroy, 2008).

Balas and Boren (2000) claim that it takes several years to translate even small amounts of original research into interventions that enhance patient care. They attribute this delay partly to the inadequacy of how health-care providers are assisted in assessing the strengths of study results and applying them to practice. Over the past decade, since the introduction of the CONSORT (Consolidated Standards of Reporting Trials)

statement, aimed at improving the quality of reporting of RCTs (Begg et al., 1996), there has been an increased focus on the methodological quality of research reports (Moher et al., 2010). However, reporting criteria of the CONSORT statement emphasize internal validity while they do not address external validity in its entirety (Glasgow et al., 2006). Reviews show that lack of discussion on external validity disadvantages judgement around the potential effectiveness of interventions and their applicability to practice (Glasgow, Klesges, Dzewaltowski, Bull, & Estabrooks, 2004). Therefore, there is a need to strengthen the reporting of generalizability of research findings (Ferguson, 2004).

Given that RCTs are not always feasible and may not be ethical within public health (Victora, Habicht, & Bryce, 2004), the TREND statement was developed to improve the quality of reporting of non-randomized evaluations of behavioural and public health interventions (Des Jarlais, Lyles, Crepaz, & TREND Group, 2004). After its publication, the statement drew immediate praise from the editors of several journals (Caetano, 2004; Kirkwood, 2004; Ross, Elford, Sherr, & Hart, 2004; Treasure, 2004). However, it was criticized for its limited external validity criteria, which were viewed as insufficient for reporting and evaluating the generalizability of study results (Dzewaltowski, Estabrooks, Klesges, & Glasgow, 2004). The critics insisted on additional criteria related to external validity. Green and Glasgow (2006) later addressed this concern by proposing criteria for external validity reporting (Table 1).

Purpose

Since its introduction in 2004, the TREND statement has been used by several researchers as a guideline for reporting of studies involving non-randomized designs. To the best of our knowledge, the impact of the use of TREND statement guidelines on external validity reporting of non-randomized intervention studies has not been reported in the literature. Therefore, the purpose of this comparative analysis was to fill the gap, with three objectives:

- (1) review selected reports claiming to have used the TREND statement as a guideline (i.e., TREND studies) and evaluate the extent to which these studies report external validity
- (2) review selected recent reports that did not use the TREND statement as a guideline (i.e., non-TREND studies) and evaluate the extent to which these studies report external validity
- (3) offer a comparative overview of external validity reporting of both TREND and non-TREND studies.

| Criteria for External Validity Reporting^b | TREND | | Non-TREND | |
|-------------------------------------------------------------|--------------|-----------|------------------|-----------|
| | <i>n</i> | % | <i>n</i> | % |
| <i>Reach and representativeness</i> | | | | |
| Participation | 7 | 100 | 6 | 86 |
| Target audience | 7 | 100 | 7 | 100 |
| Representativeness – settings | 6 | 86 | 6 | 86 |
| Representativeness – individuals | 7 | 100 | 7 | 100 |
| <i>Implementation and adaptation</i> | | | | |
| Consistent implementation | 5 | 71 | 2 | 29 |
| Staff expertise | 5.5 | 79 | 4 | 57 |
| Program adaptation | 5 | 71 | 3 | 43 |
| Mechanisms | 2.5 | 36 | 0 | 0 |
| <i>Outcomes for decision-making</i> | | | | |
| Significance | 7 | 100 | 5.5 | 79 |
| Adverse consequences | 4.5 | 64 | 1.5 | 21 |
| Moderators | 5 | 71 | 0.5 | 7 |
| Sensitivity | 7 | 100 | 4.5 | 64 |
| Costs | 4 | 57 | 2.5 | 36 |
| <i>Maintenance and institutionalization</i> | | | | |
| Long-term effects | 2 | 29 | 2 | 29 |
| Institutionalization/sustainability | 5.5 | 79 | 4 | 57 |
| Attrition | 6.5 | 93 | 4.5 | 64 |
| <i>Mean</i> | 5.4 | 77 | 3.8 | 54 |

^a Scores based on mean of two raters, who independently rated all studies listed in Table 2.
^b Criteria for external validity reporting based on recommendations of Green and Glasgow (2006).

Literature Search

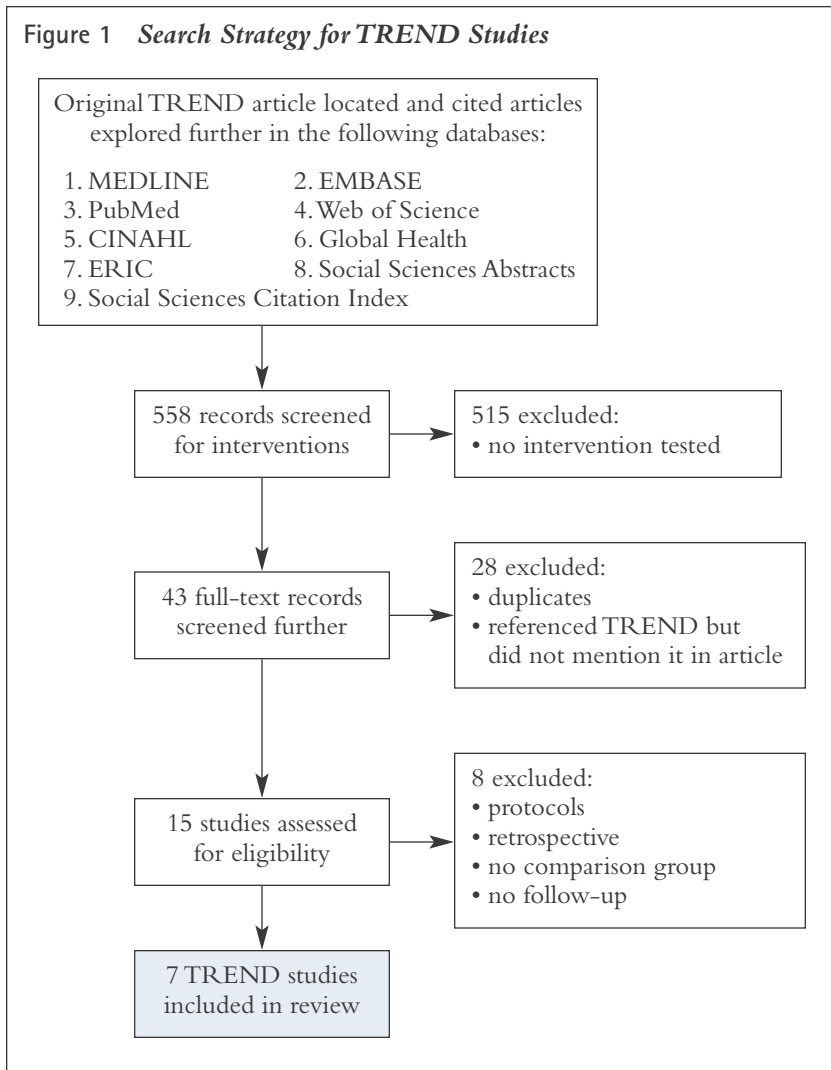
TREND Studies

Before the literature search was carried out, it was decided that the analysis would focus on prospective non-randomized intervention studies with a comparison group and a follow-up. No date limitations were set when searching for TREND studies, because the TREND guideline was published in 2004. The original TREND article by Des Jarlais et al. (2004) was sought in several databases (Figure 1), after which its citations (i.e., articles citing the original TREND article) in each database were examined. The combined search resulted in 558 records, 515 of which were excluded because the studies were not intervention studies. Of the

TREND Statement and External Validity Reporting

Noeman A. Mirza, Noori Akhtar-Danesh, Eric Staples, Lynn Martin, Charlotte Noesgaard

remaining 43 records, 28 were excluded because either they were duplicates or they cited the TREND guidelines without mentioning, discussing, or declaring whether or not the TREND guidelines were used for reporting. This resulted in 15 records, eight of which were excluded because they either were study protocols, used retrospective study designs, had no comparison group, lacked follow-up, or referred to their post-test as follow-up. The result was a total of seven TREND studies being included in the analysis.

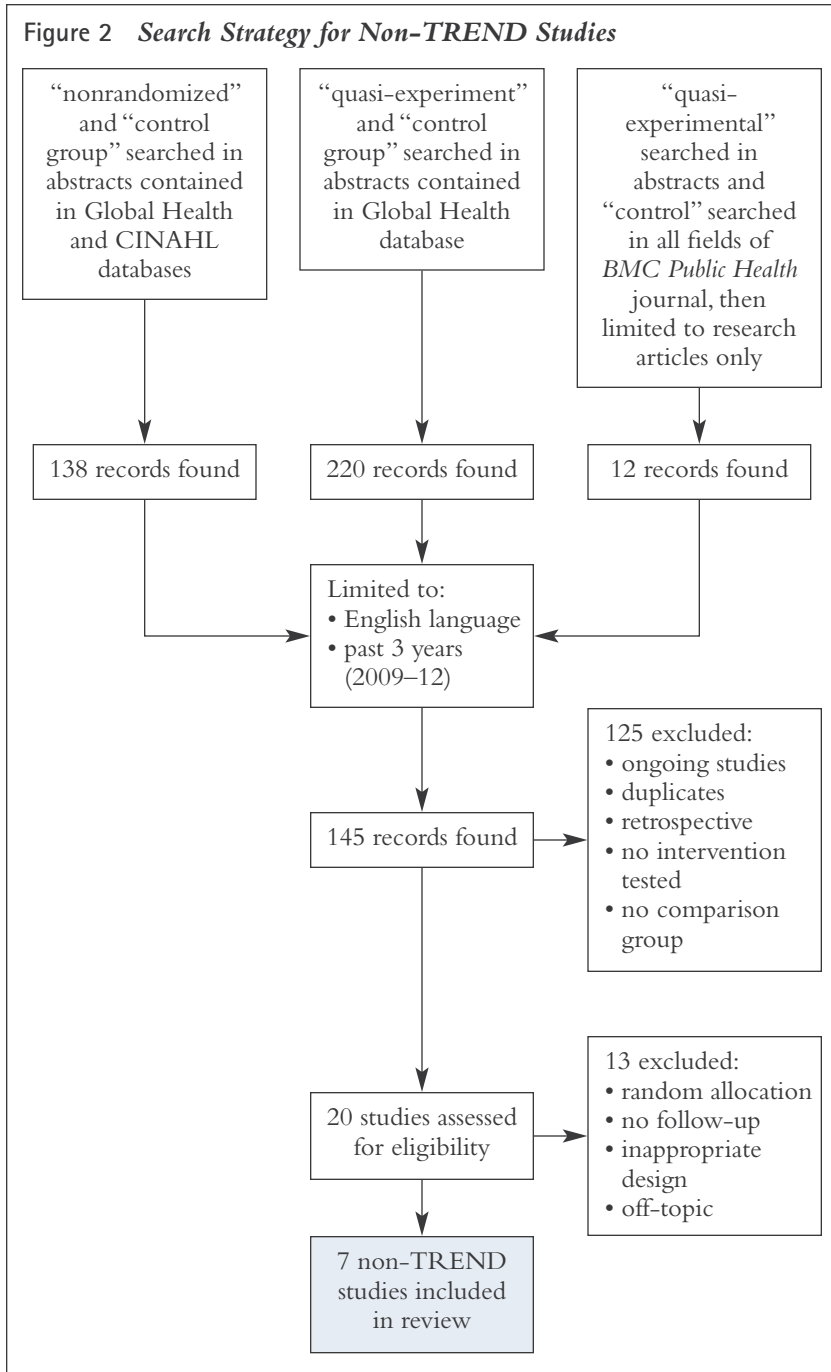


Non-TREND Studies

To compare the seven TREND studies with non-TREND studies, seven non-TREND studies were sought. Selection criteria for non-TREND studies were similar to those for TREND studies (i.e., non-randomized intervention studies with control group and follow-up). While the search for TREND studies permitted studies dating back to 2004 (when the TREND guidelines were published), only recent non-TREND studies were obtained. This was done by examining the most recent studies first and then moving back in time until seven non-TREND studies were obtained. The reason for using recent studies was based on two assumptions: that reporting of research studies will improve over time and recent study reports will represent improved trends in reporting; and that the 5-year span from 2004 (when TREND guidelines were published) to 2009 was sufficient to allow for the uptake of such guidelines by the research community. Therefore, recent studies were limited to those published in 2009 or later.

Since the TREND statement was developed initially for behavioural and public health interventions, popular nursing and public health databases were selected (i.e., Global Health, CINAHL). Key terms such as *nonrandomized* and *control group* were used in study abstracts contained in CINAHL and Global Health databases (Figure 2). This resulted in 138 records. The term *quasi-experiment* was also used in the Global Health database, which generated 220 records. Furthermore, the term *quasi-experiment* was also used in the *BMC Public Health* journal since this was a common journal among the selected TREND studies. This search generated 12 records. All 370 records were limited to the English language and to publication as early as 2009. This resulted in 145 records, 125 of which were excluded because they were ongoing studies (incomplete), used retrospective design, did not test an intervention (e.g., survey), had no comparison group, or were duplicate records. Of the 20 remaining records, 13 were excluded because they had randomized allocation, lacked follow-up, had an inappropriate design (e.g., mentioned quasi-experiment but were cross-sectional studies), or strayed from the theme of behavioural and public health interventions. Coincidentally, this search also resulted in seven non-TREND studies published from 2009 to 2011. If there had been more or fewer than seven non-TREND studies, the year of publication would have been adjusted to 2010 or 2008, respectively, in order to yield a comparable number of TREND and non-TREND studies for the analysis.

Figure 2 Search Strategy for Non-TREND Studies



Data Evaluation

To assess the external validity of studies used in the analysis, Green and Glasgow's (2006) criteria for external validity reporting were utilized by two raters, who rated all studies independently. Both raters were nurses. One had a doctorate and the other was completing a doctorate. Each rater read each study twice. During the first reading, raters scored the studies based on Green and Glasgow's criteria (Table 1). To score all these studies, a simple dichotomous scale (0 = *unreported*; 1 = *reported*) similar to the TREND checklist was employed. Studies were then read for the second time to double-check initial ratings and to seek any necessary clarification.

The choice of Green and Glasgow's (2006) proposed criteria for external validity was based on recommendations by the TREND Group (personal communication, 2012). These criteria have previously been used as a gold standard (by Klesges, Dzewaltowski, & Glasgow, 2008). As outlined in Table 1, Green and Glasgow's criteria consist of (a) *reach and representativeness*, (b) *implementation and adaptation*, (c) *outcomes for decision-making*, and (d) *maintenance and institutionalization*. Each of these four criteria comprises several attributes that a research study must include. For the purpose of the comparative analysis, a checklist with a dichotomous rating scale was developed based on all of the 16 attributes of Green and Glasgow's four criteria for external validity reporting. The rating scale was then used to rate all TREND and non-TREND studies.

Results and Analysis

All 14 studies included in the analysis reported non-randomized evaluations of behavioural and public health interventions. Both TREND and non-TREND studies were conducted in different parts of the world, with the majority originating in the United States. Most studies evaluated an intervention comprising some form of education aimed at promoting healthy behaviours (e.g., smoking cessation). Target populations ranged from children to older adults, with both males and females represented. Study participants were often allocated geographically (e.g., comparing participants in two different cities), while alternating allocation techniques were also employed (e.g., comparing participants in one setting but during different periods).

Overall, all reports based on the TREND guidelines made reference to the TREND statement but offered no further discussion about its usefulness or how each of its dimensions was addressed. Of the seven TREND studies, two recruited control and intervention participants in different years and one used non-participant controls. The remaining four were similar to all of the seven non-TREND studies in that they con-

Table 2 External Validity Scores of Reviewed Studies^a

| | Evaluation Score | |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------|-----------|
| | <i>n</i> | % |
| TREND Studies | | |
| Ciliberto et al. (2005) | 12 | 75 |
| Fisher, Wynter, & Rowe (2010) | 11.5 | 72 |
| Giangregorio et al. (2009) | 13 | 81 |
| Oupra, Griffiths, Pryor, & Mott (2010) | 12.5 | 78 |
| Sorensen et al. (2010) | 12.5 | 78 |
| Storro, Oien, Dotterud, Jenssen, & Johnson (2010) | 13.5 | 84 |
| Taylor et al. (2008) | 11.5 | 72 |
| Mean | 12.4 | 77 |
| SD | 0.75 | – |
| Non-TREND Studies | | |
| Cardarelli et al. (2011) | 9 | 56 |
| Chan et al. (2011) | 11 | 69 |
| Elmasri (2011) | 5.5 | 34 |
| Kwak, Kremers, Visscher, van Baak, & Brug (2009) | 9.5 | 59 |
| Lv & Brown (2011) | 5.5 | 34 |
| Ma et al. (2009) | 8.5 | 53 |
| Wolfers, de Wit, Hospers, Richardus, & de Zwart (2009) | 11 | 69 |
| Mean | 8.6 | 54 |
| SD | 2.30 | – |
| ^a Scores based on mean of two raters, who independently rated studies based on Green and Glasgow's (2006) criteria for external validity reporting outlined in Table 1. | | |

sisted of parallel control and intervention groups that progressed simultaneously. Table 1 summarizes the scores and percentages of external validity reporting of both TREND and non-TREND studies based on Green and Glasgow's (2006) criteria, while Table 2 summarizes the extent to which TREND and non-TREND studies addressed Green and Glasgow's criteria for external validity reporting.

Overall, all studies lacked full reporting of external validity criteria and presented limited discussion on generalizability. Across all 16 external validity criteria, mean reporting for TREND and non-TREND studies was 12.4 (*SD* = 0.75) and 8.6 (*SD* = 2.30), respectively. A non-parametric test (i.e., Mann-Whitney) indicated that this difference was statistically significant (*p* = 0.0017). To check for agreement between the scores of the two raters who independently rated each study, the Intraclass Correlation Coefficient (ICC) was calculated to be 0.86 (95% CI: 0.69, 0.97). This indicated strong interrater reliability.

Criterion 1: Reach and Representativeness

All TREND and non-TREND studies described the target audience and compared study participants with the target population, while one TREND and one non-TREND study did not report on the intended settings nor compare them with those settings that declined participation. Furthermore, while all TREND studies discussed participation rates of eligible persons, one non-TREND study did not report participation rate.

Criterion 2: Implementation and Adaptation

While five TREND studies (71%) reported on the consistency of implementation of the various intervention components and the extent to which study settings adapted the intervention program to fit their settings, only two non-TREND studies (29%) reported this information. None of the non-TREND studies (0%) reported the mechanisms through which the intervention achieved its effect. This, however, was reported by a few TREND studies (36%). Moreover, most TREND studies (79%) and several non-TREND studies (57%) presented data on staff expertise (i.e., level of training, level of expertise, quality of implementation, etc.). In relation to program adaptation, five TREND studies (71%) and three non-TREND studies (43%) reported on adaptation.

Criterion 3: Outcomes for Decision-Making

While at least four TREND studies (57%) reported on all attributes of *outcomes for decision-making*, only a few non-TREND studies (21%) reported one attribute, such as adverse consequences and moderator effects. Information on two attributes (i.e., sensitivity and significance) was provided by all TREND studies (100%). These two attributes were reported in several (> 64%) non-TREND studies. In the TREND group, there was fair reporting of cost, moderator effects, and adverse consequences by more than half of the studies (> 57%). However, these attributes were poorly reported in the non-TREND studies, with fewer than three reporting cost (36%), adverse consequences (21%), and moderator effects (7%).

Criterion 4: Maintenance and Institutionalization

Although all studies consisted of a follow-up, only two TREND studies (29%) and two non-TREND studies (29%) conducted a 12-month follow-up. In the TREND studies, follow-up ranged from 8 weeks to 2 years (8 weeks = 1 study, 3 months = 2 studies, 6 months = 2 studies, 2 years = 2 studies) with an average of 9.7 months. In non-TREND studies, this range was from 4 weeks to 2 years (4 weeks = 1 study,

3 months = 1 study, 4 months = 1 study, 6 months = 2 studies, 1 year = 1 study, 2 years = 1 study) with an average of 8 months. Furthermore, several TREND studies (79%) and non-TREND studies (57%) reported on sustainability or evolution of the program implemented as part of the intervention. Lastly, most TREND studies (93%) and several non-TREND studies (64%) reported attrition and presented some basic discussion on reasons why participants dropped out.

Discussion

A comparison of seven TREND and seven non-TREND studies is an encouraging step towards promoting external validity reporting. In this analysis, we discovered that the majority of TREND and non-TREND studies did not address Green and Glasgow's (2006) criteria for external validity reporting. The TREND Group (personal communication, 2012) views these criteria as crucial for future policy decisions and knowledge translation efforts. This analysis also highlights the lack of external validity reporting in recent non-randomized intervention studies, which could limit appropriate translation of interventions to real-life situations.

In an attempt to compare external validity reporting, this analysis shows that, compared with non-TREND studies (54%), TREND studies (77%) scored significantly higher on Green and Glasgow's (2006) criteria for external validity reporting. This illustrates that the use of the TREND statement promotes increased external validity reporting. Table 1 indicates that this difference could be due to the TREND statement's ability to draw researchers' attention towards specific external validity criteria that are important for generalizing study findings. Hence, the TREND statement, as a leap towards a systematic method of reporting, appears to promote external validity reporting in non-randomized intervention studies.

While TREND studies succeeded in reporting several criteria for external validity, there are a few areas that were not reported by several studies. These include (a) mechanisms, (b) adverse consequences, (c) costs, and (d) long-term effects.

Although the TREND statement has received criticism regarding its external validity criteria, it is important to note that the statement includes several internal validity criteria that, if reported, would also strengthen a study's external validity reporting. This is visible in Table 1, which shows that non-TREND studies performed poorly in reporting the external validity criterion of *outcomes for decision-making* while TREND studies succeeded in addressing this criterion. While many attributes under *outcomes for decision-making* (e.g., significance, adverse consequences, and moderator effects) do not correspond to criteria under

the “generalizability” section of the TREND checklist, they can be found elsewhere within the checklist, under headings intended to strengthen internal validity reporting.

This illustrates that complete use of the TREND statement can encourage the reporting of many internal validity components (e.g., significance, adverse events, implementation of intervention, moderator effects, expertise, participants, setting, cost) that can directly address Green and Glasgow’s (2006) criteria for external validity reporting. While all TREND studies made the claim that they used the TREND guidelines, they did not indicate how or to what extent. Improved external validity reporting among TREND studies could have resulted from the focus on several previously discussed internal validity criteria that directly influence Green and Glasgow’s criteria for external validity reporting.

Implications

This comparative analysis of TREND and non-TREND studies has several implications for the research community. Nurse researchers considering the TREND guidelines are encouraged to thoroughly discuss how and to what extent they used the TREND guidelines, and to pay particular attention to each of its criteria for external validity reporting. Nurse researchers must also be aware of and address these criteria when preparing study protocols before actual research is conducted. In addition to the TREND guidelines, Green and Glasgow’s (2006) external validity criteria should be considered in reports. This approach will enhance external validity reporting in journal articles and will promote subsequent knowledge translation efforts.

Dzewaltowski et al. (2004) and Steckler and McLeroy (2008) advocate for a greater emphasis on external validity reporting in journals of applied disciplines that aim to improve the health of the public. The various characteristics of external validity recommended by these authors resemble Green and Glasgow’s (2006) external validity criteria, which, along with the TREND guidelines, should be considered by researchers of all health disciplines when conducting, reporting, or evaluating non-randomized intervention studies. It is important for health researchers to report on all criteria or state that information is unavailable on criteria that may not be applicable to their research study. This can help nurses, other health-care practitioners, and policy and administrative decision-makers to determine whether or not a given study’s findings are generalizable and applicable to their local population and setting.

Findings of this comparative analysis can be used by the TREND Group to make revisions to the original TREND statement to reflect external validity criteria that emphasize and strengthen generalizability of study findings and the use of research findings in real-life situations.

Since clinicians in the public health sector often conduct non-randomized research that evaluates behavioural and public health interventions, it is important that they be aware of the usefulness of the TREND statement (Des Jarlais et al., 2004) and the external validity criteria proposed by Green and Glasgow (2006). Nurse researchers are encouraged to build partnerships with nurses and policy developers in order to address real-life problems and facilitate appropriate knowledge translation efforts.

Although Green and Glasgow's (2006) criteria do not focus on the type of intervention reported, they do focus on whether the treatment was consistently administered, whether there were any adverse reactions, what the cost was, what the long-term effects were, the attrition rate, and how the intervention was sustained. All of these factors are important when reporting intervention studies, because they allow readers to determine whether the findings can be generalized to their environment. When reports do not address such criteria, it is difficult for readers to decide whether the study intervention is suitable for their environment. Therefore, researchers are encouraged to use the TREND guidelines when reporting non-randomized intervention studies.

While use of the TREND guidelines promotes external validity reporting, single-study results must be used with extreme caution. Should nurses and policy and administrative decision-makers discover that a study report is applicable to their population and setting, they must still explore and rely on synthesized results of several research studies prior to disseminating findings in the practice setting. The use of one study and its findings is insufficient and the combined results of several well-conducted studies must be considered when making decisions around the usefulness of research and its possible effectiveness in the practice setting.

Although tools to evaluate external validity reporting are useful, the final decisions around the translation of research into practice are based on judgements of health professionals and policy and administrative personnel who understand the characteristics of people and settings outside the study experiment and are able to make accurate judgements about the applicability of research findings and their sustainable potential. Therefore, there is a need for creative solutions aimed at expanding evidence in certain areas. This analysis indicates that two areas where evidence must be expanded are the long-term follow-up of studies and the sustainability of intervention programs in institutions.

Nursing and health journals that welcome reports on intervention studies are encouraged to request that authors consider the TREND statement and relevant external validity criteria before submitting reports of non-randomized intervention studies for publication. This requirement could be incorporated into author guidelines published by journals.

Often, limited funding and urgency to publish research results prevent the conduct of studies with long-term follow-up to evaluate intervention sustainability. For this reason, nursing and health journals are also encouraged to offer researchers a venue for publishing follow-up reports on studies of interventions after an initial study has been reported. Furthermore, funding agencies are encouraged to consider providing increased support for long-term follow-up studies that allow researchers to evaluate the institutionalization and sustainability of interventions (Klesges et al., 2008).

This analysis is the first attempt to compare TREND and non-TREND studies reporting non-randomized intervention studies with a control group and follow-up. Increased utilization of the TREND guidelines is encouraged. This will increase the number of TREND studies, which could then be used in a future analysis similar to the one presented here. With more studies using the TREND guidelines, future reviewers will have more reports from which to choose and more opportunities to set further inclusion and exclusion criteria, which could ensure that selected TREND and non-TREND studies are more comparable in terms of treatment, setting, and population.

Limitations

There were several limitations in this analysis. First, there were other studies that utilized the TREND statement that were not included in the analysis, mainly because they employed a cross-sectional or time-series study design or were pre-post study designs without comparison or control groups. Therefore, this analysis lacks discussion on external validity reporting of these other studies that also used the TREND guidelines.

Second, all TREND studies briefly mentioned, in a sentence, that the TREND statement was used as a guide. However, they failed to provide further discussion on the TREND statement, which made it difficult to determine whether or not the TREND statement was useful in promoting external validity reporting of non-randomized intervention studies.

Third, although the focus of the TREND statement is not to test education interventions, some studies used the TREND guidelines for health education interventions. While this association was not explained, it is possible that use of the TREND guidelines could be expanded beyond health-related disciplines to include research studies from the social sciences and humanities.

Fourth, studies were evaluated based on a dichotomous scale, while a Likert-type scale would have been more appropriate for studies that partially met criteria for external validity reporting.

Finally, the number of studies used also influenced the analysis since small changes caused large fluctuations in percentages. Therefore, caution

is advised when generalizing the findings of this analysis. With increasing use of the TREND statement, it is recommended that a similar analysis be conducted in a few years with a larger number of TREND and non-TREND studies.

Conclusion

This comparative analysis highlights the lack of external validity reporting among non-randomized intervention study reports in the medical, nursing, and public health literature. Findings from this analysis demonstrate that use of the TREND guidelines improves external validity reporting of studies that do not use these guidelines. As a result, nurse researchers are encouraged to consider the TREND guidelines when reporting non-randomized intervention studies. It is also recommended that criteria for external validity reporting based on the work of Green and Glasgow (2006) be added to the TREND statement in order to promote external validity reporting by nurse researchers. Future non-randomized intervention study reports that succeed in addressing these external validity criteria will not only enhance generalizability, but also enrich evidence-informed decision-making and facilitate more appropriate translation of research findings into clinical practice.

References

- Balas, E. A., & Boren, S. A. (2000). Managing clinical knowledge for health care improvement. In J. Bemmel & A. T. McCray (Eds.), *Yearbook of medical informatics* (pp. 65–70). Stuttgart: Schattauer.
- Begg, C., Cho, M., Eastwood, S., Horton, R., Moher, D., Olkin, I., . . . Stroup, D. F. (1996). Improving the quality of reporting of randomized controlled trials: The CONSORT statement. *Journal of the American Medical Association*, 276(8), 637–639.
- Caetano, R. (2004). Standards for reporting non-randomized evaluations of behavioral and public health interventions: The TREND statement. *Addiction*, 99(9), 1075–1080.
- Cardarelli, K., Jackson, R., Martin, M., Linnear, K., Lopez, R., Senteio, C., . . . Johnson, E. S. (2011). Community-based participatory approach to reduce breast cancer disparities in South Dallas. *Progress in Community Health Partnerships: Research, Education, and Action*, 5(4), 375–385.
- Chan, W. L., Hui, E., Chan, C., Cheung, D., Wong, S., Wong, R., . . . Woo, J. (2011). Evaluation of Chronic Disease Self-Management Programme (CDSMP) for older adults in Hong Kong. *Journal of Nutrition, Health and Aging*, 15(3), 209–214.
- Ciliberto, M. A., Sandige, H., Ndekha, M. J., Ashorn, P., Briend, A., Ciliberto, H. M., Manary, M. J. (2005). Comparison of home-based therapy with ready-to-use therapeutic food with standard therapy in the treatment of malnourished Malawian children: A controlled, clinical effectiveness trial. *American Journal of Clinical Nutrition*, 81(4), 864–870.

TREND Statement and External Validity Reporting

Noeman A. Mirza, Noori Akhtar-Danesh, Eric Staples, Lynn Martin, Charlotte Noesgaard

- Des Jarlais, D. C., Lyles, C., Crepaz, N., & TREND Group. (2004). Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: The TREND statement. *American Journal of Public Health, 94*(3), 361–366.
- Dzewaltowski, D. A., Estarooks, P. A., Klesges, L. M., & Glasgow, R. E. (2004). TREND: An important step, but not enough. *American Journal of Public Health, 94*(9), 1474–1475.
- Elmasri, Y. M. (2011). Effect of educational program on insight into illness and attitudes toward medications among schizophrenic patients. *Journal of American Science, 7*(3), 269–277.
- Ferguson, L. (2004). External validity, generalizability, and knowledge utilization. *Journal of Nursing Scholarship, 36*(1), 16–22.
- Fisher, J. R. W., Wynter, K. H., & Rowe, H. J. (2010). Innovative psycho-educational program to prevent common postpartum mental disorders in primiparous women: A before and after controlled study. *BMC Public Health, 10*, 432.
- Gianguregorio, L. M., Thabane, L., deBeer, J., Farrauto, L., McCartney, N., Adachi, J. D., & Papaioannou, A. (2009). Body weight-supported treadmill training for patients with hip fracture: A feasibility study. *Archives of Physical Medicine and Rehabilitation, 90*(12), 2125–2130.
- Glasgow, R. E., Green, L. W., Klesges, L. M., Abrams, D. B., Fisher, E. B., Goldstein, M. G., . . . Orleans, C. T. (2006). External validity: We need to do more. *Annals of Behavioral Medicine, 31*(2), 105–108.
- Glasgow, R. E., Klesges, L. M., Dzewaltowski, D. A., Bull, S. S., & Estabrooks, P. (2004). The future of health behavior change research: What is needed to improve translation of research into health promotion practice? *Annals of Behavioral Medicine, 27*(1), 3–12.
- Green, L. W., & Glasgow, R. E. (2006). Evaluating the relevance, generalization, and applicability of research: Issues in external validation and translation methodology. *Evaluation and the Health Professions, 29*(1), 126–153.
- Hulley, S. B., Cummings, S. R., Browner, W. S., Grady, D. G., & Newman, T. B. (2007). *Designing clinical research* (3rd ed.). Philadelphia: Lippincott Williams & Wilkins.
- Kirkwood, B. (2004). Making public health interventions more evidence based: TREND statement for non-randomised designs will make a difference. *British Medical Journal, 328*(7446), 996–997.
- Klesges, L. M., Dzewaltowski, D. A., & Glasgow, R. E. (2008). Review of external validity reporting in childhood obesity prevention research. *American Journal of Preventative Medicine, 34*(3), 216–223.
- Kwak, L., Kremers, S. P., Visscher, T. L., van Baak, M. A., & Brug, J. (2009). Behavioral and cognitive effects of a worksite-based weight gain prevention program: The NHF-NRG In-Balance project. *Journal of Occupational and Environmental Medicine, 51*(12), 1437–1446.
- Lv, N., & Brown, L. (2011). Impact of a nutrition education program to increase intake of calcium-rich foods by Chinese-American women. *Journal of the American Dietetic Association, 111*(1), 143–149.

TREND Statement and External Validity Reporting

Noeman A. Mirza, Noori Akhtar-Danesh, Eric Staples, Lynn Martin, Charlotte Noesgaard

- Ma, G. X., Shive, S., Tan, Y., Gao, W., Rhee, J., Park, M., . . . Toubbeh, J. I. (2009). Community-based colorectal cancer intervention in underserved Korean Americans. *Cancer Epidemiology*, *33*(5), 381–386.
- Moher, D., Hopewell, S., Schulz, K. F., Montori, V., Gotzsche, P. C., Devereaux, P. J., . . . Altman, D. G. (2010). CONSORT 2010 explanation and elaboration: Updated guidelines for reporting parallel group randomised trials. *British Medical Journal*, *340*, c869.
- Oupra, R., Griffiths, R., Pryor, J., & Mott, S. (2010). Effectiveness of supportive educative learning program on the level of strain experienced by caregivers of stroke patients in Thailand. *Health and Social Care in the Community*, *18*(1), 10–20.
- Ross, M. W., Elford, J., Sherr, L., & Hart, G. (2004). The TREND statement: Social science, communication and HIV/AIDS. *AIDS Care*, *16*(6), 667–668.
- Sorensen, G., Stoddard, A., Quintiliani, L., Ebbeling, C., Nagler, E., Yang, M., . . . Wallace, N. (2010). Tobacco use cessation and weight management among motor freight workers: Results of the Gear Up for Health study. *Cancer Causes and Control*, *21*(12), 2113–2122.
- Steckler, A., & McLeroy, K. R. (2008). The importance of external validity. *American Journal of Public Health*, *98*(1), 9–10.
- Storro, O., Oien, T., Dotterud, C. K., Jenssen, J. A., & Johnsen, R. (2010). A primary health-care intervention on pre- and postnatal risk factor behavior to prevent childhood allergy: The Prevention of Allergy among Children in Trondheim (PACT) study. *BMC Public Health*, *10*, 443.
- Taylor, R. W., McAuley, K. A., Barbezat, W., Farmer, V. L., Williams, S. M., & Mann, J. I. (2008). Two-year follow-up of an obesity prevention initiative in children: The APPLE project. *American Journal of Clinical Nutrition*, *88*(5), 1371–1377.
- Treasure, E. (2004). The TREND statement. *Evidence-Based Dentistry*, *5*(4), 88–91.
- Victoria, C. G., Habicht, J. P., & Bryce, J. (2004). Evidence-based public health: Moving beyond randomized trials. *American Journal of Public Health*, *94*(3), 400–405.
- Wolfers, M. E., de Wit, J. B., Hospers, H. J., Richardus, J. H., & de Zwart, O. (2009). Effects of a short individually tailored counselling session for HIV prevention in gay and bisexual men receiving hepatitis B vaccination. *BMC Public Health*, *9*, 255.

Noeman A. Mirza, RN, BScN, is a PhD student in the School of Nursing, McMaster University, Hamilton, Ontario, Canada. Noori Akhtar-Danesh, PhD, is Associate Professor, School of Nursing, McMaster University. Eric Staples, RN, DNP, is a Nursing Consultant. Lynn Martin, RN, PhD, is Assistant Professor, School of Nursing, McMaster University. Charlotte Noesgaard, RN, MScN, is Associate Professor, School of Nursing, McMaster University.