Annals of Work Exposures and Health, 2018, Vol. 62, No. 3, 259–268
doi: 10.1093/annweh/wxx112
Advance Access publication 30 January 2018
Original Article

Original Article

The Effect of Workforce Mobility on Intervention Effectiveness Estimates

Justin Manjourides[1*], Emily H. Sparer[2,3], Cassandra A. Okechukwu[2], Jack T. Dennerlein[3,4]

[1]Department of Health Sciences, Bouvé College of Health Sciences, Northeastern University, 360 Huntington Avenue, Boston, MA 02115, USA; [2]Department of Social and Behavioral Sciences, Harvard T.H. Chan School of Public Health, 677 Huntington Avenue, Boston, MA 02115 , USA; [3]Center for Community-Based Research, Dana-Farber Cancer Institute, 450 Brookline Avenue LW703, Boston, MA 02215, USA; [4]Department of Physical Therapy, Movement, and Rehabilitation Sciences, Bouvé College of Health Sciences, Northeastern University, 360 Huntington Avenue, Boston, MA 02115, USA

*Author to whom correspondence should be addressed. Tel: +1-617-373-3666; fax: +1-617-373-2968; e-mail: j.manjourides@northeastern.edu

Submitted 30 May 2017; revised 14 November 2017; editorial decision 19 November, 2017; revised version accepted 15 December 2017.

Abstract

Background: Little is known about how mobile populations of workers may influence the ability to implement, measure, and evaluate health and safety interventions delivered at worksites.
Methods: A simulation study is used to objectively measure both precision and relative bias of six different analytic methods as a function of the amount of mobility observed in the workforce. Those six methods are then used to reanalyze a previously conducted cluster-randomized control trial involving a highly mobile workforce in the construction industry.
Results: As workforce mobility increases, relative bias in treatment effects derived from standard models to analyze cluster-randomized trials also increases. Controlling for amount of time exposed to the intervention can greatly reduce this bias. Analyzing only subsets of workers who exhibit the least amount of mobility can result in decreased precision of treatment effect estimates. We demonstrate a 59% increase in the treatment effect size from the reanalysis of the previously conducted trial.
Conclusions: When evaluating organizational interventions implemented at specific worksites by measuring perceptions and outcomes of workers present at those sites, researchers should consider the effects that the mobility of the workforce may have on the estimated treatment effects. The choice of analytic method can greatly affect both precision and accuracy of estimates.

Keywords: cluster-randomized control trials; dynamic workforce; linear modeling

© The Author(s) 2018. Published by Oxford University Press on behalf of the British Occupational Hygiene Society.

Downloaded from https://academic.oup.com/annweh/article-abstract/62/3/259/4830164
by Centers For Disease Control user
on 13 March 2018

Introduction

To affect change in workers' health and safety, worksite-based interventions usually target a multilevel host of factors that may include the organizational, physical, and/or social environment of workers (OSHA, 2012). Literature reviews demonstrate that worksite-based interventions that address changes in the work environment in addition to targeting specific attributes or behaviors of workers are associated with greater success (Sorensen and Barbeau, 2005; Cherniack *et al.*, 2011).

The activities involved in worksite-based intervention programs usually take less than 6 months in duration, and most workers are at their worksites for longer than 6 months (Sorensen *et al.*, 2005; Barbeau *et al.*, 2007; Kennedy *et al.*, 2010); however, there are some exceptions as some workplaces are highly dynamic. Consider worksites in the construction industry that feature a constantly changing array of sub-contractors specializing in different aspect of the construction processes working under the umbrella of a general contractor and owner. The length of time that different workers are physically present at construction site varies depending on the phase of the construction project and the workers' trades (Carlan *et al.*, 2012). In practice, we see workers such as carpenters and laborers on a worksite for long periods of time, perhaps even the entire duration of a project, whereas others, such as ironworkers and floorlayers, may only be on-site for discrete amounts of time during certain phases of a project (Sparer *et al.*, 2015).

There is a dearth of research on how dynamic population of workers, those that move on and off of jobsites frequently, may influence the ability to implement, measure, and evaluate interventions at worksites. In one group-randomized controlled intervention trial targeting safety practices that took place over 5 months, approximately 50% of the construction workers who were on-site at the start of the intervention moved on within 1 month (Sparer *et al.*, 2015). A steady flow of new workers replaced those who exited; therefore, the number of the workers on-site remained approximately the same, but the individuals composing the workforce changed constantly. This dynamic poses several statistical and experimental design challenges to the measurement and evaluation of worksite-based interventions. Employing intent-to-treat analysis is one strategy for handling the variability in dosage (Little and Yau, 1996). However, in such a dynamic setting, employing intent-to-treat analysis may result in too great a loss in sample size and overly pull estimated treatment effects towards the null hypothesis.

In this article, we aimed to estimate the effects that varying levels of worker mobility have on the ability to accurately and precisely measure the efficacy of worksite-based health interventions. To accomplish this aim, we designed a flexible, construction-site level, person-based model that simulates different worker mobility patterns to mirror these dynamic work environments.

We have previously seen the effects that the mobility of study participants can have on the ability to locate clusters of both communicable and non-communicable diseases (Manjourides and Pagano, 2011; Manjourides *et al.*, 2012). Here, we will analyze a previously conducted study to demonstrate the effect of a mobile workforce on treatment effect estimates. We will then use a simulation to compare the estimation properties of several different analytic methods applied to an intervention study conducted at a construction worksite. Finally, we will apply these different analytic methods to the original study. The goal of this work is to inform the design, implementation, and evaluation of future worksite-based intervention programs targeting workers with dynamic worksites.

Materials and methods

Motivating intervention study

The current analysis uses data from a cluster-randomized control trial on the effectiveness of a safety communication and recognition program (which we will refer to as the C-RCT Program) implemented on construction worksites (Sparer *et al.*, 2015, 2016). This intervention aimed to improve organizational policies, programs, and practices regarding worker safety. It was not an individual, behavior-based safety program; rather, it focused on the communication infrastructure on the construction site. While the conceptual framework of the intervention was based at the organizational level, we focused our data collection for intervention effectiveness evaluation at the worker level. We used worker surveys measuring antecedents of injury, which are a common method of evaluations in occupational safety and health research due to the fact that injuries are fairly rare, and directly measuring changes in organizational factors can be difficult. Worker surveys on the other hand can capture a range of constructs and reflect the worker experience in perceiving safety climate, safety leadership, and other related outcomes.

In this work, we will focus on the safety climate and practices outcome. Safety climate and practices were measured using a nine-item index (scored from 0 to 90) (Dedobbeleer and Beland, 1991). Safety practices have strong associations with safety- and injury-related outcomes in many different industries (Huang *et al.*, 2006; Probst *et al.*, 2008) and are often used as a proxy for

Downloaded from https://academic.oup.com/annweh/article-abstract/62/3/259/4830164
by Centers For Disease Control user
on 13 March 2018

changes in health and safety outcomes. The full nine-item safety practices index was measured on workers from four pairs of medium to large (between 50 and 300 workers at any given time) commercial construction sites in the Greater Boston area between August 2011 and December 2013. Each pair consisted of one intervention and one control site. All procedures and methods were reviewed and approved by the Harvard School of Public Health's Office of Regulatory Affairs and Research Compliance and Northeastern University's Institutional Review Board.

Despite the fact that over 50% of workers surveyed remained on a given site for less than 1 month, this study found a positive increase in safety practices, over a 4–6 month time frame (Sparer *et al.*, 2016). However, post-intervention safety practices were only assessed from those workers who provided names and mobile phone numbers during collection of the baseline survey, resulting in response rates of 71% on the intervention sites and 81% on the control sites. There were no differences in average duration of time-on-site between the intervention and control sites. Additionally, it was determined that time on-site was significantly associated with change in safety climate and that adjusting for time-on-site increased the observed effect size of the intervention from 2.06 (95% CI: 0.39, 3.70) to 2.29 (95% CI: 0.51, 4.07) when adjusting for the same parameters (Sparer *et al.*, 2016). The final models account for time-on-site and presented a relative difference in effect estimate between the control and intervention sites.

Given the high level of mobility among workers (Sparer *et al.*, 2015), we hypothesize that the observed improvement in the safety practice index may actually be an underestimate of the true effect of the intervention. The simulation study described below is designed to determine the degree to which this effect may have been underestimated and how different choices of analytic method may impact estimation.

Simulation model structure

We simulated data modeled to match the C-RCT Program population by creating worksites composed of individual workers that would come and go throughout a 6-month period. The primary variables for each worker included their starting month, duration of stay on the site, and demographic variables. Since the data were collected in 1-month intervals in the original study, we used the same time resolution for describing this cohort's mobility pattern. Based on previous results, presence of baseline pain was also simulated and used to model time on-site (Sparer *et al.*, 2015). This simulation structure allows for flexibility to create cohorts of workers with different mobility patterns.

Each worker was assigned baseline and follow-up values for the primary outcome (safety practice index). Baseline scores were randomly sampled from a distribution mirroring the baseline measures in the C-RCT Program (Fig. 1).

To determine outcomes at the follow-up time points, we assumed: (i) each worker had a true safety practice score that can be influenced by the intervention until they were no longer on-site; (ii) the intervention runs for 6 months; and (iii) there is a linear dose–response relationship with respect to time on treatment. Specifically, we calculated the safety practice score (SP_{ij}) for person *i* (i=1,…, n) at follow-up month *j* (j=0,…,F) as

$$SP_{ij} = SP_{i0} * (1+\beta)^{(j/D)} + N(0,\sigma^2), \qquad (1)$$

where SP_{i0} was the baseline safety practice score, D was the duration of the intervention, β was the intervention effect, and σ^2 was an error term estimated from the variance of the observed safety practice scores. We denoted the final safety practice score for person *i* as $SP_{iF.}$

In this work, we assumed a balanced study, with 600 workers at an intervention site and 600 workers at a control site. This design simulates a perfectly matched and balanced study with six pairs of intervention and controls sites, each with 100 workers. Each of these simulated workers is identical to each other with regard to all potential confounders, with the exception of variables that were found to be predictive of duration, including age, tenure, and baseline pain, though these are equally distributed across worksites. Treatment effects were simulated to only be functions of baseline safety practice and duration. Through this design, no other worker characteristics can confound the relationships between the change in safety practice score and duration. The intervention program in the simulation model was assumed to run for 6 months. We allowed workers who were not present at baseline to receive the full dosage of the intervention by assuming that the effect of the intervention on a given worker was maximized after 4 months of exposure (i.e. D in Equation 1 is equal to 4 months).

We estimated treatment effects across four different worker mobility patterns (Fig. 2). The first pattern matched the observed data in the C-RCT Program. The second pattern assumed a uniform distribution on the durations, given the starting month, where any possible duration is equally likely to be assigned to each worker. The third mobility pattern was created empirically to be slightly less mobile than the C-RCT Program. And the fourth pattern corresponded to a static worksite where all participants remain on-site for the entire study duration. The last duration distribution is particularly important as it represents the assumption

Downloaded from https://academic.oup.com/annweh/article-abstract/62/3/259/4830164
by Centers For Disease Control user
on 13 March 2018

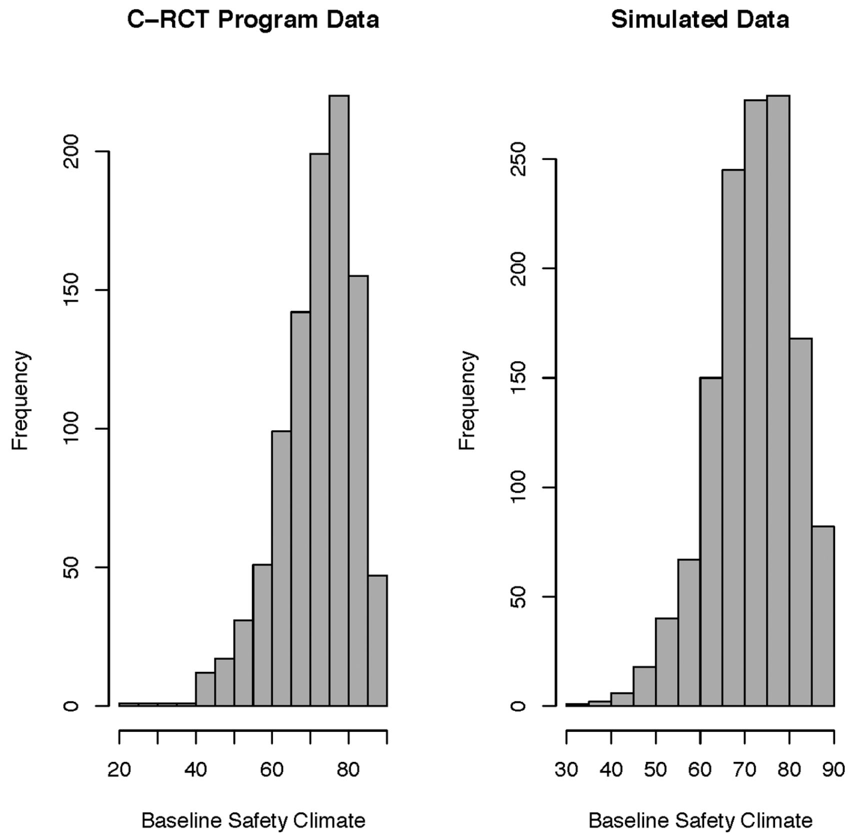

Figure 1. (Left) The distribution of baseline safety practice scores observed in the C-RCT Program. The scores were calculated from the nine-item index developed for construction workers by Dedobbeleer and Béland (1991). (Right) The distribution of simulated baseline safety practice scores.

made in the extant literature where adjustments are currently not made for worker mobility.

We examined six different treatment effects for the intervention site ranging from no change to a 50% increase in the outcome measure. These effect sizes ($\beta = 0, 0.05, 0.1, 0.15, 0.2, 0.5$, in Equation 1) represent a reasonable range of intervention effects on safety practice and include an intervention effect equal to zero allowing us to verify that the Type 1 error rate of our procedure is 0.05. We simulated 1000 complete cohorts for each combination of mobility pattern and treatment effect.

Six different intervention evaluation methods were used to estimate the treatment effect: (i) a standard intent-to-treat analysis linear model (LM); (ii) an intent-to-treat analysis, including an interaction between duration on worksite and treatment (INT); (iii) a LM including only the subset of workers on-site for a consecutive 4-month period (Sub-1); (iv) a LM using only those workers on-site at the conclusion of the study (Sub-2); (v) a LM with a duration interaction including only those workers on-site at the conclusion of the study (Sub-INT); (vi) subset analysis of only those workers who both were on-site for 4 months and were present at the end of the study (Sub-3). The LM and INT methods treat the last observed outcome for each worker as a surrogate for the final measure. Sub-1 reduces the sample to only those workers who have complete outcome data. The Sub-2, Sub-INT, and Sub-3 analyses treat those workers that leave the worksite prior to the end of the study as missing data and exclude them from the analysis.

The standard LM analysis is essentially the same as a last observation carried forward (LOCF) analysis, in that we assumed the final safety practice observation for each subject represents their actual post-intervention safety practice score (Little and Rubin, 2014). This situation can be reframed as an observation that has been lost to follow-up. LOCF methods have been previously shown to underperform in situations where there is loss to follow-up (Molnar *et al.*, 2008; Streiner, 2008; Molnar *et al.*, 2009; Saha and Jones, 2009). The LM, Sub-1, Sub-2, and Sub-3 are the same model:

Downloaded from https://academic.oup.com/annweh/article-abstract/62/3/259/4830164
by Centers For Disease Control user
on 13 March 2018

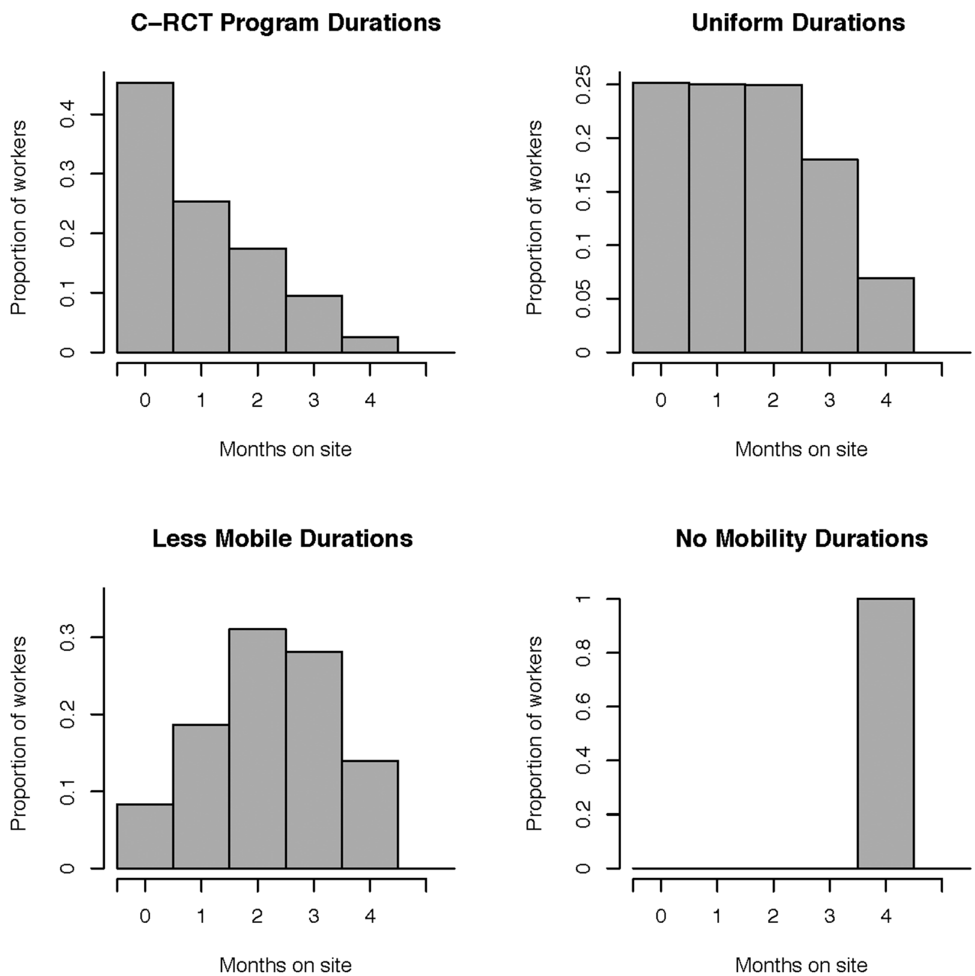

Figure 2. Distributions of the duration of time each worker spends on a worksite assuming differing mobility models. (Top left) Observed in C-RCT Program. (Top right) Uniform durations. (Bottom right) No mobility. (Bottom left) A duration distribution empirically generated to be less mobile than the observed C-RCT Program distribution.

$$\left(\mathrm{SP}_{\mathrm{iF}} - \mathrm{SP}_{\mathrm{i0}}\right) = \beta_0 + \beta_1 * \mathrm{Trt}_\mathrm{i} + \varepsilon_\mathrm{i}, \quad (2)$$

where $\varepsilon_i \sim N(0,\sigma^2)$, applied to different subsets of the full data. From Model (1), we were interested in estimating the treatment effect, β_1.

The LM-INT and Sub-INT analyses acknowledge that the effect of the treatment may be modified by the amount of time exposed to the intervention. This model is specified as

$$\left(\mathrm{SP}_{\mathrm{iF}} - \mathrm{SP}_{\mathrm{i0}}\right) = \beta_0 + \beta_1 \mathrm{Trt}_\mathrm{i} + \beta_2 * \mathrm{D}_\mathrm{i} + \beta_3 * \mathrm{D}_\mathrm{i}\mathrm{Trt}_\mathrm{i}. + \varepsilon_\mathrm{i}. \quad (3)$$

The estimate of the treatment effect from this model can be calculated for any specified duration of exposure to the intervention. As we were attempting to determine the effect of the total intervention, we calculated the estimated treatment effect when the duration on-site equals 4 months as $(\beta_1 + 4\beta_3)$.

Evaluation of analytic methods

The C-RCT Program data were then reanalyzed using the INT and Sub-1 methods. The original analysis of the effect of the C-RCT Program intervention on change in safety practice consisted of three mixed effects linear regression models: (i) a base model including only the worksite as a random effect; (ii) an expanded model including a fixed effect for the matched pair of the block-randomized study design; and (iii) a fully adjusted model controlling for worker start month, total time on-site, job title, trade, race, and age. Here, we replicate these three models and evaluate the effect of controlling for worker mobility on the intervention effect estimates.

We applied the six methods to the 1000 simulated cohorts, under each mobility distribution. When there was no worker mobility, the lack of variability in the participant durations did not allow us to fit either of the

Downloaded from https://academic.oup.com/annweh/article-abstract/62/3/259/4830164
by Centers For Disease Control user
on 13 March 2018

LMs with the interaction terms. Methods that worked with subsets of workers who were present for all months of the intervention were by definition equivalent to the complete data methods. A total of 120 combinations of mobility pattern, treatment effect, and analytic method were examined, each with 1000 simulations.

We compared the performance of each method with respect to accuracy and precision. The accuracy of each method was evaluated based on the relative bias, calculated as the difference between the estimated treatment effect and the true treatment effect, divided by the true treatment effect. When no treatment effect was present ($\beta_2 = 0$), we calculated the absolute bias of each method, as the difference between the mean of the estimates of β_2 from the simulated data and true parameter value of zero. Precision of each of the six methods was evaluated by calculating the standard deviations of the 1000 estimates, summarizing the variability of each method.

Results

The relative bias of the estimated intervention effects varied across the four mobility patterns for traditional LM evaluation method (Fig. 3), though was consistent across the five levels of treatment effect. As hypothesized, the LM and Sub-2 methods consistently underestimated the true treatment effect across all scenarios involving worker mobility. The most severe underestimation from the LM and Sub-2 methods, relative biases of up to −0.77 and −0.53 respectively, corresponded to the most mobile duration distribution examined (the observed distribution of durations from the C-RCT Program). This observed bias towards the null hypothesis of no treatment effect was consistent across all non-zero effect sizes. When there was no treatment effect ($\beta = 0$), the absolute bias, $(\hat{\beta} - \beta)$, was approximately zero across all scenarios. In each scenario, the amount of bias associated with these standard LMs increased as average worker durations decreased.

Accounting for differential durations by including an interaction term removed much of the bias that is introduced by the workforce mobility (INT and Sub-INT in Fig. 3). This reduction in bias was consistent across duration distributions as well as treatment effect sizes. Using the C-RCT Program duration distribution, the maximum relative bias of the INT model was −0.059, observed when the treatment effect was largest ($\beta = 0.50$).

Using an LM on the subset of workers remaining on-site for the entirety of the intervention, Sub-1, produced the least biased estimates across all duration distributions. However, creating this subset increased the standard error of the estimate, resulting in a reduction in precision (Fig. 4). As the workforce became less mobile, the increase in the standard error was less pronounced, a result of an increased number of workers on-site for the entire duration of the study.

The two LMs that did not include an interaction term, LM and Sub-2, appear slightly more precise across most effect sizes in each duration distribution; nevertheless, the magnitude of the bias in these settings outweighed any potential variance reductions. Strictly comparing the estimation bias, there does not appear to be an appreciable difference between the interaction model and the models that analyzed subsets of the least mobile members of the population. After considering the precision, the method that includes both an interaction term and the full duration of the intervention performs best.

Reanalyzing the C-RCT Program data

Compared to the standard procedure, both the INT and Sub-1 methods produced increases in effect estimates (Table 1). The biggest gains for the INT method were seen in the base and expanded models (31.98% and 29.64% increases in effect estimates, respectively). The fully adjusted model showed an increase of 1.11%. Though, the Sub-1 method produced larger increases in the effect estimates (gains of 47.5% for the base model and 58.3% for the adjusted model), these results should be considered along with the necessary reduction in sample size and the complete exclusion of one site.

Because of a reduction from 615 to 93 workers when we examined the subset of workers who spent at least 4 months on-site, we did not fit the fully adjusted LM included in the original analysis, which included random-site effects and adjusted for 15 demographic covariates. Additionally, only two subjects from Site 2 and no subjects from Site 4 were exposed the full 4 months of treatment, adding further uncertainty to the estimated treatment effects from the Sub-1 model results.

Discussion

The goal of this article was to estimate the effects of varying levels of worker mobility on the ability to accurately and precisely measure the efficacy of a worksite-based health and safety intervention. In practice, we rely on surveys of worker perceptions of safety as a proximal measure to evaluate organizational interventions due to the rarity of observing injuries during the span of an intervention program. However, evaluating the effectiveness of an organizational intervention by measuring these individual outcomes on a dynamic workforce may result in a tendency to underestimate the true effect. Over the long term, as more organizations,

Downloaded from https://academic.oup.com/annweh/article-abstract/62/3/259/4830164
by Centers For Disease Control user
on 13 March 2018

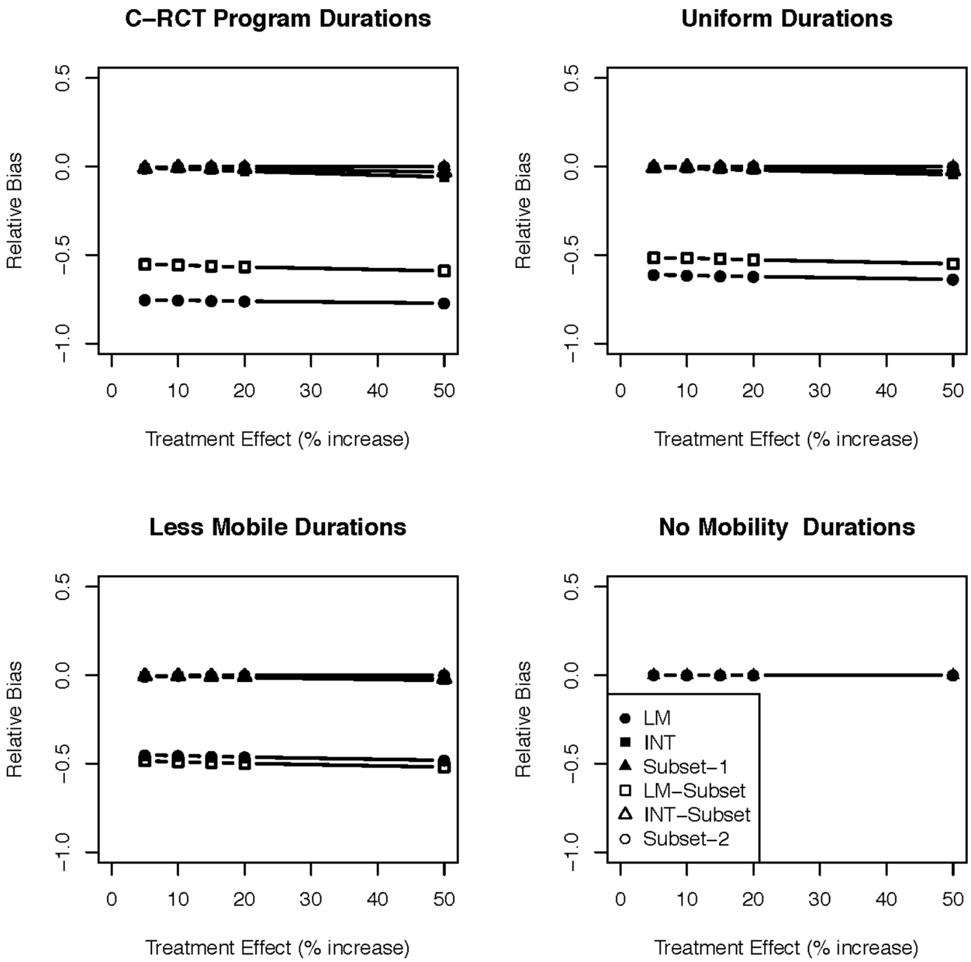

Figure 3. Simulated relative bias in six analytic approaches under the four different mobility distributions, across five different treatment effect sizes. Bias was calculated as the mean of the difference between the effect estimates from 1000 simulated datasets and the true treatment effect. The lines marked with filled circles represent the bias from the standard intent-to-treat LM, solid squares denote the LM with an interaction between workplace duration and treatment, the filled triangles represent the standard LM only including those workers on-site for the entire duration of the study, empty squares represent the LM subset analysis only including those workers present at the conclusion of the study period, empty triangles represent the LM with an interaction between workplace duration and treatment applied only to those workers present at the conclusion of the study, and the empty circles represent the standard LM applied only to those workers who were both present at the conclusion of the study and received the full 4 months of the intervention.

and worksites throughout each individual organization, implement these types of programs, we would expect a general improvement in the health and well-being of the workforce, and it is this effect that may be underestimated by not considering worker mobility.

We have demonstrated the impact of workforce mobility when designing cluster-randomized controlled trials of individual level treatments at dynamic worksites. By investigating simple, yet flexible simulated scenarios that were informed by real data, this study showed that standard analysis of a cluster-based randomized trial, in the presence of increasing worker mobility, resulted in biased estimates of the true intervention treatment effects. The amount of bias present, relative to the true effect size, seemed to be mostly dependent on the amount of mobility among the individual subjects, with greater relative bias associated with shorter worker durations on each site. Analyses that accounted for the time spent on the worksite site mitigated some of these biases, while including an interaction between time on-site and exposure to the intervention appeared to have the biggest bias reduction.

When mobility among subjects increased, the effects were 2-fold. First, as workers are more mobile, they are exposed to the intervention for shorter periods of

Downloaded from https://academic.oup.com/annweh/article-abstract/62/3/259/4830164
by Centers For Disease Control user
on 13 March 2018

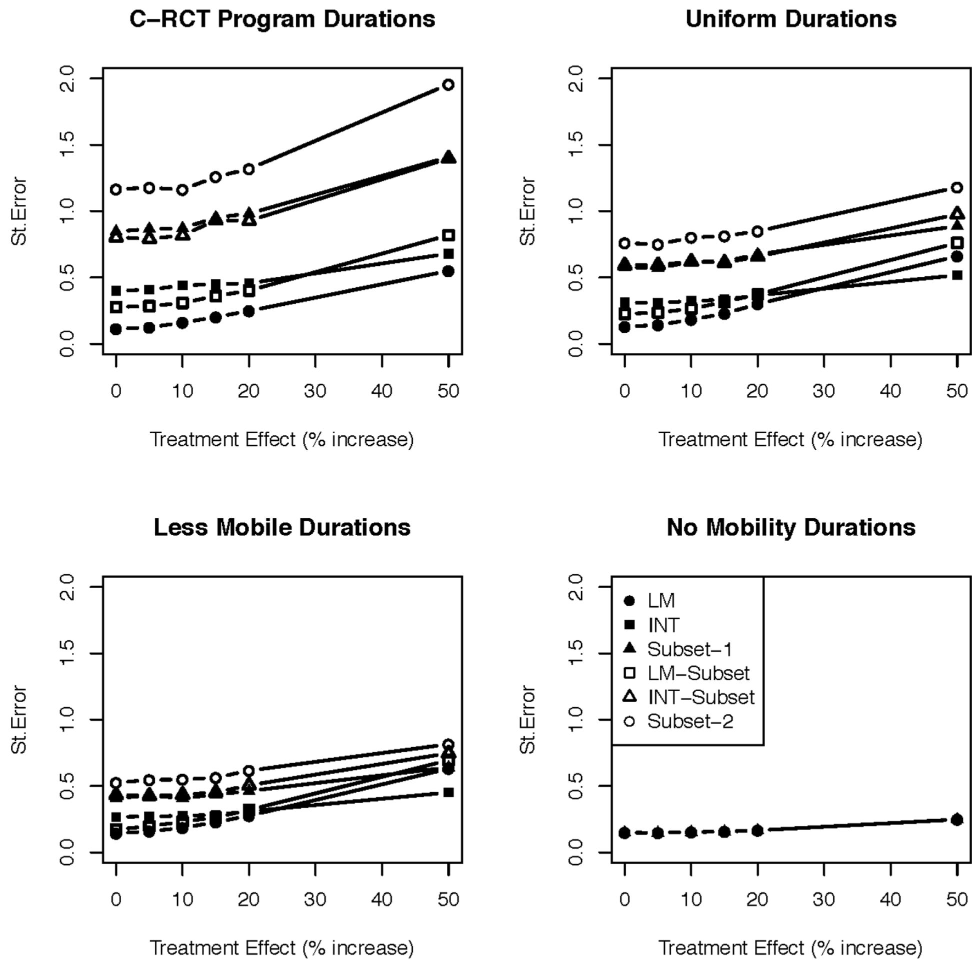

Figure 4. Simulated standard errors of estimates from six analytic approaches assuming four different mobility distributions, across six different treatment effect sizes. Standard error is the standard deviation of the treatment effect estimates from the 1000 simulations. The lines marked with filled circles represent the bias from the standard intent-to-treat LM, solid squares denote the LM with an interaction between workplace duration and treatment, the filled triangles represent the standard LM only including those workers on-site for the entire duration of the study, empty squares represent the LM subset analysis only including those workers present at the conclusion of the study period, empty triangles represent the LM with an interaction between workplace duration and treatment applied only to those workers present at the conclusion of the study, and the empty circles represent the standard LM applied only to those workers who were both present at the conclusion of the study and received the full 4 months of the intervention.

time. Assuming a linear dose–response model of the treatment means that the intervention effect size for each worker is a fraction of the intended dose. Second, as workers who have been exposed to the intervention leave the worksite, they are replaced by workers who have not been exposed. In any given month, there is a non-negligible proportion of workers on the treatment sites that have not received the intervention. These consequences of workplace mobility consistently bias estimates of treatment efficacy towards the null hypothesis of ‘no association’ when workplace interventions are analyzed using standard methods. In some cases, these biases were as much as 77% of the true treatment effect.

When we applied the INT and Sub-1 methods to a previously analyzed study, we saw increased treatment effect estimates when compared with standard modeling procedures. The mobility of the workforce made fitting fully adjusted models on subsets of the population that were exposed to the full duration of the intervention inappropriate due to number of independent variables compared with the number of observations.

This simulation study, though based on an actual study and informed by real data, made several

Downloaded from https://academic.oup.com/annweh/article-abstract/62/3/259/4830164
by Centers For Disease Control user
on 13 March 2018

Table 1. Intervention effect estimates.[a]

	Standard approach	INT method	Increase (%)	Sub-1 method	Increase (%)
Base model	1.98	2.61	31.98	2.92[b]	47.47
Adjusted model[c]	2.06	2.67	29.64	3.26[b]	58.25
Full model[d]	2.29	2.32	1.11	[e]	[e]

[a]Change in safety practice index associated with exposure to intervention.

[b]Missing one site (Site D) completely. Only two observations from Site B.

[c]Dependent variable is the change in pre- and post-safety climate scores from the C-RCT Program. Independent variables are worksite treatment status (control or intervention); standard model includes a random effect for site and adjusts for worksite pair.

[d]Same parameters as Model 2. Also adjusted for worker trade, title, age, race/ethnicity, month started on-site, total amount of time on-site.

[e]Model failed to converge due to reduced sample size.

simplifying assumptions, resulting in several limitations. For interventions where treatment requires only a single dose, we would expect to see slightly less biased results. The portion of the bias due to the individual workers not receiving the entire treatment should be reduced; however the contamination bias due to treated subjects leaving the intervened upon worksite and being replaced by untreated workers would still likely be present. While the model used in this study was specifically chosen to isolate the impacts of mobility on one's ability to accurately measure the effect of a specific treatment, if there were factors that influence a workers mobility (such as union status, tenure, or baseline pain), and treatment effects differ across levels of that additional covariate, then one would certainly need to account for these relationships to more accurately estimate model parameters. Further, our model only examined impact on the estimation of a continuous outcome (safety practice). The applicability of our conclusions to binary or time-to-event outcomes should be examined in future work.

These results suggest that if an effective organizational intervention were run on a long-term (multi-year) construction project, with similar levels of worker movement on- and off-site, we would expect to see a bigger improvement on proximal worker outcomes. In practice, because most construction projects are shorter in duration, we do not have the ability to observe the full effects of these organizational changes.

Conclusion

Our study demonstrated the potential underestimation of a true treatment effect that can occur as a result of implementing a cluster-randomized organizational intervention in a dynamic environment that is evaluated by measuring the perceptions and outcomes of those workers present on specific worksites. We found that this bias was directly related to the level of mobility among the participants of the study and was consistently in the direction of the null hypothesis of no treatment effect. We determined that a reasonable method to reduce this bias, while maintaining precision, is to account for worksite duration and include an interaction term between this duration and exposure to treatment. These results underscore the need for intervention implementation and analysis to account for worker mobility.

Acknowledgments

The authors would like to thank Benjamin Smith from Northeastern University for assistance in programming the simulation. Research reported in this publication was partially supported by a pilot study grant from the Harvard T. H. Chan School of Public Health Center for Work, Health and Wellbeing, a National Institute for Occupational Safety and Health Center of Excellence to Promote a Healthier Workforce (grant U19 OH008861). Additional support was provided by a grant from the Center for Construction Research and Training (CPWR) under grant U60OH009762, by National Institutes of Health (NIH) grant number 3R25CA057711, and by the National Institute of Arthritis and Musculoskeletal and Skin Diseases of the NIH under award number T32AR055885.

Conflict of Interest

The authors declare no competing interests. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the CPWR or the NIH.

References

Barbeau EM, Delaurier G, Kelder G *et al.* (2007) A decade of work on organized labor and tobacco control: reflections on research and coalition building in the United States. *J Public Health Policy*; **28**: 118–35.

Downloaded from https://academic.oup.com/annweh/article-abstract/62/3/259/4830164
by Centers For Disease Control user
on 13 March 2018

Carlan NA, Kramer DM, Bigelow P *et al*. (2012) Digging into construction: social networks and their potential impact on knowledge transfer. *Work*; **42**: 223–32.

Cherniack M, Henning R, Merchant JA *et al*. (2011) Statement on national worklife priorities. *Am J Ind Med*; **54**: 10–20.

Dedobbeleer N, Beland F. (1991) A safety climate measure for construction sites. *J Safety Res*; **22**: 97–103.

Huang YH, Ho M, Smith GS *et al*. (2006) Safety climate and self-reported injury: assessing the mediating role of employee safety control. *Accid Anal Prev*; **38**: 425–33.

Kennedy CA, Amick BC III, Dennerlein JT *et al*. (2010) Systematic review of the role of occupational health and safety interventions in the prevention of upper extremity musculoskeletal symptoms, signs, disorders, injuries, claims and lost time. *J Occup Rehabil*; **20**: 127–62.

Little RJ, Rubin DB. (2014). *Statistical analysis with missing data*. Hoboken, New Jersey, USA: John Wiley & Sons.

Little R, Yau L. (1996) Intent-to-treat analysis for longitudinal studies with drop-outs. *Biometrics*; **52**: 1324–33.

Manjourides J, Pagano M. (2011) Improving the power of chronic disease surveillance by incorporating residential history. *Stat Med*; **30**: 2222–33.

Manjourides J, Lin HH, Shin S *et al*. (2012) Identifying multi-drug resistant tuberculosis transmission hotspots using routinely collected data. *Tuberculosis (Edinb)*; **92**: 273–9.

Molnar FJ, Hutton B, Fergusson D. (2008) Does analysis using 'last observation carried forward' introduce bias in dementia research? *CMAJ*; **179**: 751–3.

Molnar FJ, Man-Son-Hing M, Hutton B *et al*. (2009) Have last-observation-carried-forward analyses caused us to favour more toxic dementia therapies over less toxic alternatives? A systematic review. *Open Med*; **3**: e31–50.

OSHA. (2012) *Injury and illness prevention program: White paper*. From OSHA.gov. Available at: https://www.osha.gov/dsg/topics/safetyhealth/OSHAwhite-paper-january2012sm.pdf. Retrieved December 7, 2014.

Probst TM, Brubaker TL, Barsotti A. (2008) Organizational injury rate underreporting: the moderating effect of organizational safety climate. *J Appl Psychol*; **93**: 1147–54.

Saha C, Jones MP. (2009) Bias in the last observation carried forward method under informative dropout. *J Stat Plan Inference*; **139**: 246–55.

Sorensen G, Barbeau EM. (2006) Integrating occupational health, safety and worksite health promotion: opportunities for research and practice. *Med Lav*; **97**: 240–57.

Sorensen G, Barbeau E, Stoddard AM *et al*. (2005) Promoting behavior change among working-class, multiethnic workers: results of the healthy directions–small business study. *Am J Public Health*; **95**: 1389–95.

Sparer EH, Catalano PJ, Herrick RF *et al*. (2016) Improving safety climate through a communication and recognition program for construction: a mixed methods study. *Scand J Work Environ Health*; **42**: 329–37.

Sparer EH, Okechukwu CA, Manjourides J *et al*. (2015) Length of time spent working on a commercial construction site and the associations with worker characteristics. *Am J Ind Med*; **58**: 964–73.

Streiner DL. (2008) Missing data and the trouble with LOCF. *Evid Based Ment Health*; **11**: 3–5.

Downloaded from https://academic.oup.com/annweh/article-abstract/62/3/259/4830164
by Centers For Disease Control user
on 13 March 2018