

TECHNICAL ADVANCE

Open Access



# How to analyze work productivity loss due to health problems in randomized controlled trials? A simulation study

Wei Zhang<sup>1,2\*</sup>  and Huiying Sun<sup>2</sup>

## Abstract

**Background:** An increasing number of randomized controlled trials (RCTs) have measured the impact of interventions on work productivity loss. Productivity loss outcome is inflated at zero and max loss values. Our study was to compare the performance of five commonly used methods in analysis of productivity loss outcomes in RCTs.

**Methods:** We conducted a simulation study to compare Ordinary Least Squares (OLS), Negative Binomial (NB), two-part models (the non-zero part following truncated NB distribution or gamma distribution) and three-part model (the middle part between zero and max values following Beta distribution). The main number of observations each arm,  $N_{obs}$ , that we considered were 50, 100 and 200. Baseline productivity loss was included as a covariate.

**Results:** All models performed similarly well when baseline productivity loss was set at the mean value. When baseline productivity loss was set at other values and  $N_{obs} = 50$  with  $\leq 5$  subjects having max loss, two-part models performed best if the proportion of zero loss  $> 50\%$  in at least one arm and otherwise, OLS performed best. When  $N_{obs} = 100$  or 200, the three-part model performed best if the two arms had equal scale parameters for their productivity loss outcome distributions between zero and max values.

**Conclusions:** Our findings suggest that when treatment effect at any given values of one single covariate is of interest, the model selection depends on the sample size, the proportions of zero loss and max loss, and the scale parameter for the productivity loss outcome distribution between zero and max loss in each arm of RCTs.

**Keywords:** Productivity loss, Absenteeism, Presenteeism, Zero-inflated data, Simulation studies, Randomized controlled trial, Two-part model, Three-part model

## Background

In addition to direct medical care costs, indirect cost (i.e., work productivity loss due to health problems) is an important component when estimating burden of illness [1–4] or conducting economic evaluations of health care interventions from a societal perspective [5, 6]. Correspondingly, in recent years, an increasing number of randomized controlled trials (RCTs) have measured the

impact of health care interventions on work productivity loss either considering it as one patient-centered outcome [7–14] or as one component for economic evaluations [15–18].

Complete components of work productivity loss include 1) absenteeism; 2) presenteeism; 3) employment status changes including reduced routine work hours and work stoppage due to illness [19]. Before transformed into monetary amount, work productivity loss due to health is usually first expressed as work time loss, i.e., counting the days missed from work (absenteeism), the hours lost due to reduced productivity while working

\* Correspondence: [wzhang@cheos.ubc.ca](mailto:wzhang@cheos.ubc.ca)

<sup>1</sup>School of Population and Public Health, University of British Columbia, Vancouver, British Columbia, Canada

<sup>2</sup>Centre for Health Evaluation and Outcome Sciences, 588-1081 Burrard Street, Vancouver, British Columbia V6Z 1Y6, Canada



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

(presenteeism), or stopped workdays. Therefore, productivity loss data could be non-negative count data. When presenteeism is first measured using percentage of loss (e.g., from the Work Productivity and Activity Impairment Questionnaire (WPAI) [20, 21] and the Valuation of Lost Productivity questionnaire [22, 23]) and then transformed into work time loss by multiplying the percentage of loss by the actual working time, the estimate may also be non-negative continuous data.

When estimating work time loss among a study population, people can be divided into three groups: I) those with no time loss, II) those with some time loss, and III) those who have lost all work time. Studies have shown that the proportion of Group I is very high, i.e., zero loss [24–27]. Such a distribution applies to the sum of time loss from all the three subcomponents and the time loss from absenteeism and presenteeism, respectively. For example, among patients with arthritis, a study found that the frequency of ‘0’ value varied by the measurement questionnaires: 61% if presenteeism was measured using the Health Labour Questionnaire, 5% using Work Limitations Questionnaire, 16% using the World Health Organization Health and Work Performance Questionnaire, and 27% using WPAI [25]. A clinical trial among employed patients with early rheumatoid arthritis showed that about 50–70% did not have any paid work productivity loss (sum of all three subcomponents) depending on the follow-up time point, and that about 5–10% stopped working due to health problems at Week 13 or cumulatively at Week 52 [27], which contributed to a high proportion of Group I and Group III, i.e., inflated zero value and maximum value.

Various statistical models have been used for analyzing productivity loss. Ordinary Least Squares (OLS) regression was commonly used in previous studies [7–9, 28, 29]. Poisson and Negative Binomial (NB) models for count data were also used [11, 12]. Some studies avoided estimating the mean time loss by using logistic models where productivity loss was treated as binary or categorical variables [10, 30, 31]. These methods may be problematic. For example, the logistic models did not take full use of the continuous data information. Various models have been suggested to deal with zero-inflated and bound-inflated data, e.g., two-part models, zero-inflated models, and other mixture models [32–36]. Kleinman et al. utilized a two-part model to estimate annual lost days due to absenteeism and presenteeism [37]. Also, zero-inflated models were used to analyze productivity loss outcomes [13, 27]. Each method has its own assumptions and its estimation could be biased if these assumptions are not satisfied. Thus, it is important to compare which analytic method works best for analyzing work productivity loss. This study’s objective was to compare the performance of these commonly used

methods for analyzing work time loss data in RCTs using simulations.

**Methods**

Our simulation methods followed the published guidance by Morris et al. [38] on using simulation studies to evaluate statistical methods.

**Data-generating mechanisms**

**Distributions of productivity loss outcome  $Y$**

We assumed that the productivity (time) loss outcome  $Y$  in an RCT depends on the treatment  $arm$  ( $arm = 0, 1$ ) and a covariate  $x$ , denoted by  $Y(arm, x)$ . For given  $arm$  and a value of  $x$ , the probabilities of  $Y(arm, x)$  being zero (Group I above), all loss  $Max$  (Group III), and in  $(0, Max)$  (Group II) are denoted by  $P_1(arm, x)$ ,  $P_3(arm, x)$ , and  $P_2(arm, x)$ , respectively, where  $P_1(arm, x) + P_2(arm, x) + P_3(arm, x) = 1$ . The relationships among treatment  $arm$ , covariate  $x$  and the probabilities  $P_1(arm, x)$ ,  $P_2(arm, x)$ , and  $P_3(arm, x)$  are given by the follow equations:

$$P_1(arm, x) = \frac{\exp(\alpha_1 + \beta_1 arm + \gamma_1 x)}{\exp(\alpha_1 + \beta_1 arm + \gamma_1 x) + \exp(\alpha_2 + \beta_2 arm + \gamma_2 x) + 1}$$

$$P_2(arm, x) = \frac{\exp(\alpha_2 + \beta_2 arm + \gamma_2 x)}{\exp(\alpha_1 + \beta_1 arm + \gamma_1 x) + \exp(\alpha_2 + \beta_2 arm + \gamma_2 x) + 1}$$

$$P_3(arm, x) = \frac{1}{\exp(\alpha_1 + \beta_1 arm + \gamma_1 x) + \exp(\alpha_2 + \beta_2 arm + \gamma_2 x) + 1}$$

where  $\alpha_1, \beta_1, \gamma_1, \alpha_2, \beta_2,$  and  $\gamma_2$  are given parameters at each simulation.

We denoted  $Y(arm, x)$  truncated at 0 and  $Max$  by  $Y(arm, x)|_0^{Max}$ . We assumed that  $Y(arm, x)|_0^{max}$  follows a truncated NB distribution, denoted by  $NB(r(arm, x), p(arm, x))|_0^{Max}$ , with mean  $E_{r(arm, x), p(arm, x)}\{Y(arm, x)| 0 < Y(arm, x) < Max\}$  and standard deviation  $SD_{r(arm, x), p(arm, x)}\{Y(arm, x)| 0 < Y(arm, x) < Max\}$ . We further assumed that

$$E_{r(arm, x), p(arm, x)}\{Y(arm, x)| 0 < Y(arm, x) < Max\} = E_{r(arm, 0), p(arm, 0)}\{Y(arm, 0)| 0 < Y(arm, 0) < Max\} + a \cdot x$$

and  $p(arm, x) = p(arm, 0)$  for all  $x$ , where  $a$  is a parameter assumed. Thus, for given  $E_{r(arm, 0), p(arm, 0)}\{Y(arm, 0)| 0 < Y(arm, 0) < Max\}, SD_{r(arm, 0), p(arm, 0)}\{Y(arm, 0)| 0 < Y(arm, 0) < Max\}$ , and parameter  $a$  at each simulation, the NB parameters  $r(arm, x)$  and  $p(arm, x)$  can be derived for given  $arm$  and  $x$  (see the detailed derivation in Additional file 1).

For given  $arm$  and  $x$ , the mean of the productivity loss outcome  $Y(arm, x)$  can be calculated by

$$E\{Y(\text{arm}, x)\} = P_1(\text{arm}, x) \cdot 0 + P_2(\text{arm}, x) \cdot E_{r(\text{arm}, x), p(\text{arm}, x)}\{Y(\text{arm}, x) | 0 < Y(\text{arm}, x) < \text{Max}\} + P_3(\text{arm}, x) \cdot \text{Max}$$

**Distribution for covariate  $x$**

Since most RCTs are not randomized by patients’ productivity loss at baseline, which is highly correlated with productivity loss outcome, it is common to use regression models to adjust for baseline productivity loss. Therefore, in this study, we assumed  $x$  to be the productivity loss at baseline which follows a distribution with  $P$ , the probability of being zeros ( $P > 0$ ), and  $1 - P$ , the probability being non-zero values from a NB distribution truncated at 0 and  $\text{Max}$ ,  $NB(r_x, p_x)|_0^{\text{Max}}$ . That is,  $x$  is independent of treatments and all RCT participants should be working at baseline and thus their productivity loss does not equal to  $\text{Max}$ . The truncated NB distribution has mean =  $\mu_x$  and variance =  $sd_x$ . At each simulation,  $\mu_x$  and  $sd_x$  are given parameters and  $r_x$  and  $p_x$  are derived from  $\mu_x$  and  $sd_x$ .

**Simulation parameters**

We assumed the time period for estimating work productivity loss is 12 weeks and the maximum loss time is 60 days ( $\text{Max} = 60$ ). We considered three sets of parameters for the multinomial distributions for  $Y(\text{arm}, x)$ , three sets of parameters for  $NB(r(\text{arm}, x), p(\text{arm}, x))|_0^{60}$  and one set of parameters for baseline productivity loss  $x$  (see all parameters in Table 1). The parameters were chosen based on our review of recently published articles which measured absenteeism and presenteeism in an RCT [7–14, 27, 39].

**Number of observations each arm  $N_{obs}$**

Similarly, based on the common sample size in the previous RCT studies, we chose  $N_{obs} = 50, 100, \text{ and } 200$  participants who are working at baseline for each arm at each simulation. A sample size of 1000 and 2000 were also used to check whether bias (see definition below) varied by sample size.

**Simulation algorithm**

At each simulation, we generated  $N_{obs}$  samples for each  $\text{arm}$  in the following steps.

For  $\text{arm} = 0, 1$  and  $i = 1, 2, 3, \dots, N_{obs}$ ,

1. Randomly generate  $b_i$  from Bernoulli distribution Bernoulli ( $P$ ). If  $b_i = 1$ , then let  $x_i = 0$ . If  $b_i = 0$ , then randomly generate  $x_i$  from  $NB(r_x, p_x)|_0^{60}$ .
2. Randomly generate vector  $(K_1(\text{arm}, x_i), K_2(\text{arm}, x_i), K_3(\text{arm}, x_i))$  from the multinomial distribution  $(P_1(\text{arm}, x_i), P_2(\text{arm}, x_i), P_3(\text{arm}, x_i), 1)$ , where

$K_1(\text{arm}, x_i), K_2(\text{arm}, x_i)$ , and  $K_3(\text{arm}, x_i)$  are either 0 or 1 and  $\sum_{j=1}^3 K_j(\text{arm}, x_i) = 1$ .

3. Randomly generate  $Z(\text{arm}, x_i)$  from  $NB(r(\text{arm}, x_i), p(\text{arm}, x_i))|_0^{\text{Max}}$ .
4. The productivity loss outcome  $Y_i(\text{arm})$  is defined by

$$Y_i(\text{arm}) = \begin{cases} 0 & \text{if } K_1(\text{arm}, x_i) = 1 \\ Z(\text{arm}, x_i) & \text{if } K_2(\text{arm}, x_i) = 1 \\ 60 & \text{if } K_3(\text{arm}, x_i) = 1 \end{cases}$$

**Estimation methods**

Each simulated dataset was analyzed using the following five regression models:

1. OLS;
2. NB: generalized linear model for NB distribution;
3. ZTNB: two-part model – logistic regression for the probability of being zero, and generalized linear regression with zero-truncated NB distribution (Hurdle) for the non-zeros.
4. ZG: two-part model – logistic regression for the probability of being zero, and generalized linear regression with Gamma distribution for the non-zeros;
5. Three-part model: multinomial logistic regression for the probabilities of being zero and 60 and generalized linear regression with Beta distribution for the those with values in (0, 60) (transformed to (0, 1)).

**Estimand**

Our estimands in the simulation study were  $\theta(x) = E\{Y(1, x)\} - E\{Y(0, x)\}$  at the values of  $x$  of interest. The estimates of estimand  $\theta(x)$  in each regression, denoted by  $\hat{\theta}(x)$ , were derived from the given  $x$  and the estimated regression parameters. We used bootstrapping method, 1000 replications, to estimate the standard error (SE) of  $\hat{\theta}(x)$  in all regressions except OLS.

**Performance measures**

Our key performance measure of interest was bias  $\hat{\theta}(x) - \theta(x)$ . We assumed  $SD(\hat{\theta}(x)) \leq 4$  (the standard deviation of the estimators  $\hat{\theta}(x)$ ) and considered 0.08 as an acceptable Monte Carlo SE of bias. Accordingly, we needed at least 2500 repetitions based on the published simulation study guidance [38]. We chose our final number of repetitions,  $n_{sim} = 5000$  and thus Monte Carlo SE of coverage would be 0.3 and 0.7 for coverage of 95 and 50%, respectively, which are acceptable.

Our performance measures included bias, coverage, power, empirical and model-based SE, and the mean squared error (MSE) for  $\hat{\theta}(x)$  at  $x = \mu_x, 0$  and 30. Their

**Table 1** Parameters in the simulation study

Parameters	80:15:5/60:30:10	60:35:5/40:50:10	50:40:10/30:55:15
Number of observations each arm: $N_{obs}$	50; 100; 200	50; 100; 200	50; 100; 200
<b>Multinomial distribution of productivity loss outcomes</b>			
Probability of zero loss: $P_1(arm, 0)$			
Arm 1	0.80	0.60	0.50
Arm 0	0.60	0.40	0.30
Probability of loss between 0 and Max: $P_2(arm, 0)$			
Arm 1	0.15	0.35	0.40
Arm 0	0.30	0.50	0.55
Probability of max loss: $P_3(arm, 0)$			
Arm 1	0.05	0.05	0.10
Arm 0	0.10	0.10	0.15
Coefficient of $x$			
$\gamma_1$	-0.02	-0.02	-0.02
$\gamma_2$	-0.01	-0.01	-0.01
<b>Truncated negative binomial distribution of productivity loss outcomes between 0 and Max</b>			
Mean: $E$ for Arm 1	15.00	15.00	15.00
Mean: $E$ for Arm 0	20.00	20.00	20.00
Max loss	60.00	60.00	60.00
Equal scale			
SD for Arm 1	13.34	13.34	13.34
SD for Arm 0	14.63	14.63	14.63
Unequal scale			
SD for Arm 1	12.00	12.00	12.00
SD for Arm 0	16.00	16.00	16.00
Coefficient of $x$ : $a$	0.15	0.15	0.15
<b>Baseline productivity loss</b>			
Probability of zero loss: $P$	0.30	0.30	0.30
Mean for non-zero loss: $\mu_x$	20.00	20.00	20.00
SD for non-zero loss: $sd_x$	16.00	16.00	16.00

SD standard deviation

definitions can be found in Morris et al. [38]. All analyses were performed using SAS 9.4. SAS codes used for our simulation study are available in Additional file 1.

## Results

### Computational issues

We encountered convergence issues mainly in the scenarios with smaller number of observations and higher proportion of Group I and lower proportion of Group III. For example, when  $N_{obs} = 50$ , the simulated databases based on 80% zero loss could not generate enough samples for Group II and Group III and thus we did not compare the five models in this scenario. Similarly, the simulated databases based on 5% max loss could have no samples for Group III and thus the three-part model was considered for such scenarios. For some of the

remaining scenarios, the issue of quasi-complete separation (the maximum likelihood estimate may not exist) was detected while running multinomial logistic regression for the three-part model. Table 2 presents the number of databases with quasi-complete separation detected among the 5000 simulated databases by three different sets of distributions of productivity loss outcomes in two arms (i.e., the proportions of zero loss, some loss and max loss for productivity loss outcomes in the two arms);  $N_{obs} = 50, 100$  and 200; and whether the two arms have equal scale parameters for truncated NB distributions of their productivity loss outcomes. The number of simulations with the convergence issue was very small.

Table 3 presents the proportions of the total 10,000 simulations (5000 for equal scale and 5000 for unequal

**Table 2** Number of simulated databases with quasi-complete separation

Distribution of productivity loss outcome	Number of Observations in each arm	Truncated negative binomial distributions of productivity loss outcomes in the two arms	Number of databases with quasi-complete separation <sup>a</sup>
80:15:5/60:30:10	100	Equal Scale	3
80:15:5/60:30:10	100	Unequal Scale	7
80:15:5/60:30:10	200	Equal Scale	0
80:15:5/60:30:10	200	Unequal Scale	0
60:35:5/40:50:10	100	Equal Scale	0
60:35:5/40:50:10	100	Unequal Scale	0
60:35:5/40:50:10	200	Equal Scale	0
60:35:5/40:50:10	200	Unequal Scale	0
50:40:10/30:55:15	50	Equal Scale	0
50:40:10/30:55:15	50	Unequal Scale	2
50:40:10/30:55:15	100	Equal Scale	0
50:40:10/30:55:15	100	Unequal Scale	0
50:40:10/30:55:15	200	Equal Scale	0
50:40:10/30:55:15	200	Unequal Scale	0

<sup>a</sup>The maximum likelihood estimate may not exist while running multinomial logistic regression for the three-part models. For number of observations = 50 and 5% max loss, a large number of databases with quasi-separation and thus the three-part model was not considered for these scenarios

scale) that have the number of quasi-complete separation detected among their 1000 bootstraps equal to 0, 1–10, 11–100, 101–400, 401–700, 701–950, and 950–1000 by the distributions of productivity loss outcomes and  $N_{obs}$ . The issue was the most problematic in the scenario with the distribution of productivity loss outcome, 80%/15%/5% vs. 60%/30%/10%, and 100 observations. Only 38.03% of the 10,000 simulations did not have any bootstraps with quasi-complete separation, 41.97% had 1–10 bootstraps with quasi-complete separation and 0.09% had > 950 bootstraps with quasi-complete separation.

**Bias**

Figure 1 (for  $N_{obs} = 100$ ) and Supplementary Figure S1 ( $N_{obs} = 50$ ), S2 ( $N_{obs} = 200$ ), S3 ( $N_{obs} = 1000$ ) and S4

( $N_{obs} = 2000$ ) (see Additional file 2) present the bias for the five regression models by the value of baseline productivity loss,  $x$ ; three different sets of distributions of productivity loss outcomes in two arms; and whether the two arms have equal scale parameters for truncated NB distributions of their productivity loss outcomes.

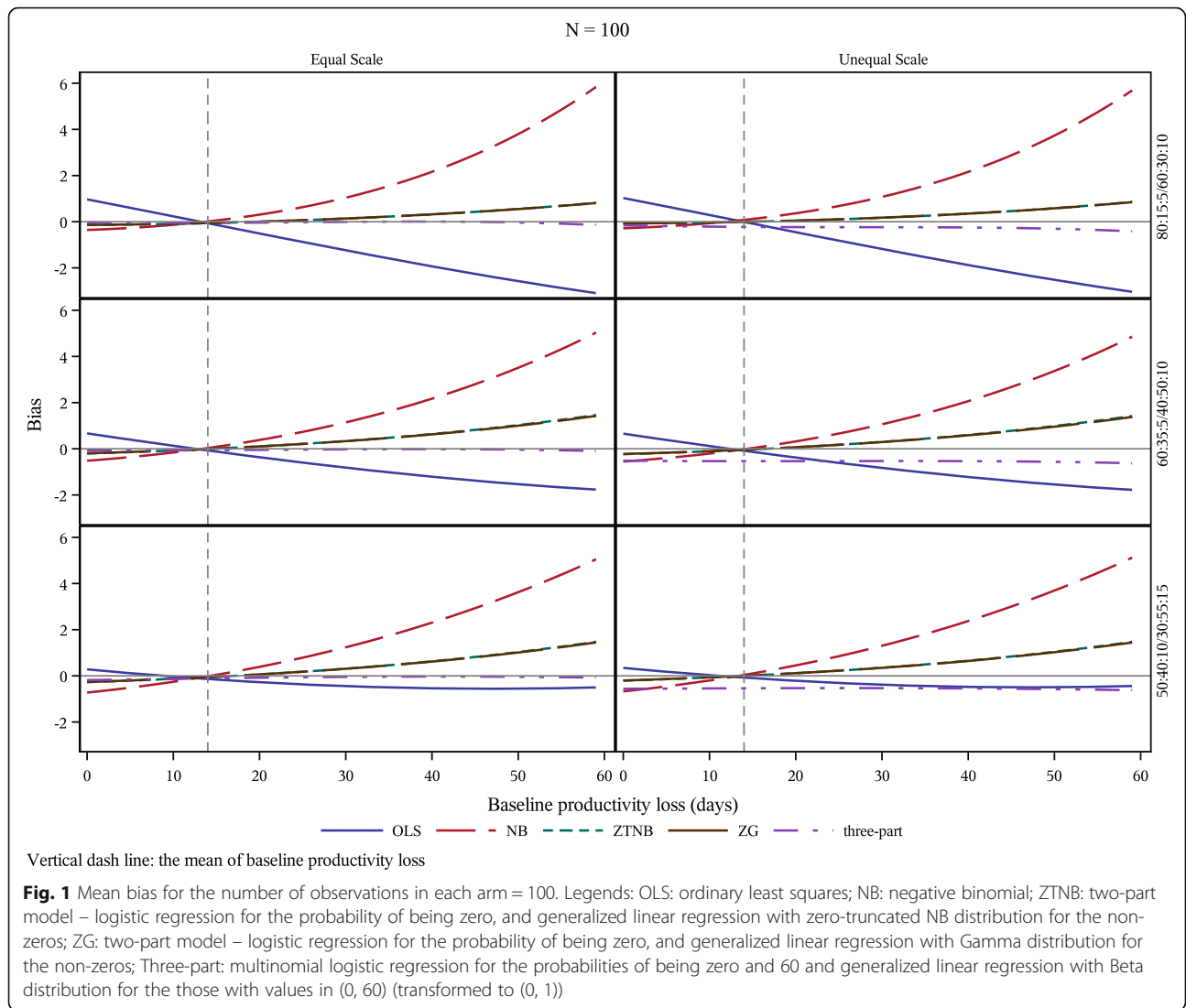
**Covariate  $x$**

All models performed similarly well at the mean of baseline productivity loss,  $x = 14$ . At other assessed  $x$  values, NB performed worst for all scenarios and three-part model performed best. Two-part models performed almost exactly the same for any  $x$  values regardless whether the second part was assumed to be a truncated NB or gamma distribution.

**Table 3** The proportions of simulations that having quasi-complete separation issues among their 1000 bootstraps

Distribution of productivity loss outcome	Number of observations in each arm	Number of quasi-complete separation detected in 1000 bootstraps						
		0	1–10	11–100	101–400	401–700	701–950	> 950
80:15:5/60:30:10	100	38.03	41.97	16.64	3.25	0.01	0.01	0.09
80:15:5/60:30:10	200	94.75	5.00	0.24	0.01	0	0	0
60:35:5/40:50:10	100	61.16	32.83	5.63	0.38	0	0	0
60:35:5/40:50:10	200	99.17	0.81	0.02	0	0	0	0
50:40:10/30:55:15	50	39.45	45.05	14.30	1.11	0.08	0.01	0
50:40:10/30:55:15	100	97.46	2.46	0.08	0	0	0	0
50:40:10/30:55:15	200	100	0	0	0	0	0	0

Among 10000 simulations for each row: 5000 for equal scale and 5000 for unequal scale for the truncated negative binomial distributions of productivity loss outcomes in the two arms



**Proportions of zero loss and max loss**

When productivity loss outcome had a high proportion of zero loss (> 50% in at least one arm) and a non-trivial proportion of max loss (> 5%), two-part models and three-part model performed better than OLS and NB models. The performance of OLS was getting better when the proportion of zero loss was lower ( $\leq 50\%$  in both arms) and could be better than two-part models when  $x$  values were higher.

**Different scale parameters between the two arms**

When the two arms had unequal scale parameters in the assumed truncated NB distributions for their productivity loss outcomes, three-part model provided bias estimators for all values of  $x$ . Two-part models could produce lower bias than three-part model at lower values of  $x$ . OLS performed best or as well as two-part models or three-part model in the databases with the

proportions of productivity loss outcomes at 50%/40%/10% vs. 30%/55%/15%.

**Sample size**

For  $N_{obs} = 50$ , we only applied the three-part model for databases with the proportions of productivity loss outcomes at 50%/40%/10% vs. 30%/55%/15%. The performance results based on bias did not change when  $N_{obs}$  increased from 50 to 100, 200, 1000 and 2000.

**Other performance measures**

Similar to bias, other performance measures, including coverage, power, empirical SE, model SE and MSE, were comparable among the five different models when  $x = 14$  (Tables 4, 5 and 6 and Supplementary Tables S1-S3 in Additional file 3). However, when  $x = 0$  or 30, the coverages of the two-part models and three-part model were similar; the coverage of the three-part model was slightly

**Table 4** Coverage for the number of observations each arm = 100

Distribution of productivity loss outcome	Scales in the two arms <sup>a</sup>	OLS	NB	ZTNB	ZG	Three-part
<b>x = mean</b>						
80:15.5/60:30:10	Equal Scale	95.2% (0.003)	94.7% (0.003)	94.5% (0.003)	94.5% (0.003)	94.6% (0.003)
80:15.5/60:30:10	Unequal Scale	94.7% (0.003)	94.4% (0.003)	94.3% (0.003)	94.3% (0.003)	94.4% (0.003)
60:35.5/40:50:10	Equal Scale	95.2% (0.003)	94.7% (0.003)	94.8% (0.003)	94.7% (0.003)	94.8% (0.003)
60:35.5/40:50:10	Unequal Scale	94.5% (0.003)	94.3% (0.003)	94.3% (0.003)	94.3% (0.003)	94.1% (0.003)
50:40:10/30:55:15	Equal Scale	95.2% (0.003)	95.0% (0.003)	94.9% (0.003)	94.9% (0.003)	95.2% (0.003)
50:40:10/30:55:15	Unequal Scale	95.4% (0.003)	94.9% (0.003)	94.8% (0.003)	94.8% (0.003)	94.8% (0.003)
<b>x = 0</b>						
80:15.5/60:30:10	Equal Scale	93.4% (0.004)	94.3% (0.003)	94.7% (0.003)	94.7% (0.003)	95.0% (0.003)
80:15.5/60:30:10	Unequal Scale	93.1% (0.004)	93.9% (0.003)	94.2% (0.003)	94.2% (0.003)	94.2% (0.003)
60:35.5/40:50:10	Equal Scale	94.2% (0.003)	94.3% (0.003)	94.4% (0.003)	94.4% (0.003)	95.1% (0.003)
60:35.5/40:50:10	Unequal Scale	94.3% (0.003)	93.6% (0.003)	94.3% (0.003)	94.3% (0.003)	94.4% (0.003)
50:40:10/30:55:15	Equal Scale	95.2% (0.003)	93.9% (0.003)	94.9% (0.003)	94.8% (0.003)	95.1% (0.003)
50:40:10/30:55:15	Unequal Scale	95.0% (0.003)	93.9% (0.003)	94.8% (0.003)	94.8% (0.003)	94.7% (0.003)
<b>x = 30</b>						
80:15.5/60:30:10	Equal Scale	92.6% (0.004)	94.1% (0.003)	94.5% (0.003)	94.5% (0.003)	94.6% (0.003)
80:15.5/60:30:10	Unequal Scale	92.0% (0.004)	93.7% (0.003)	94.3% (0.003)	94.3% (0.003)	94.4% (0.003)
60:35.5/40:50:10	Equal Scale	94.8% (0.003)	93.5% (0.003)	94.8% (0.003)	94.7% (0.003)	95.3% (0.003)
60:35.5/40:50:10	Unequal Scale	93.9% (0.003)	93.4% (0.004)	94.3% (0.003)	94.3% (0.003)	94.6% (0.003)
50:40:10/30:55:15	Equal Scale	95.1% (0.003)	94.0% (0.003)	95.0% (0.003)	94.9% (0.003)	95.4% (0.003)
50:40:10/30:55:15	Unequal Scale	95.2% (0.003)	93.5% (0.003)	94.8% (0.003)	94.8% (0.003)	95.0% (0.003)

OLS ordinary least squares, NB negative binomial, ZTNB two-part model – logistic regression for the probability of being zero, and generalized linear regression with zero-truncated NB distribution for the non-zeros, ZG two-part model – logistic regression for the probability of being zero, and generalized linear regression with Gamma distribution for the non-zeros; Three-part: multinomial logistic regression for the probabilities of being zero and 60 and generalized linear regression with Beta distribution for the those with values in (0, 60) (transformed to (0, 1))

<sup>a</sup>for truncated negative binomial distributions of productivity loss outcomes in the two arms

larger than that of OLS and NB except when the proportions of productivity loss outcomes were at 50%/40%/10% vs. 30%/55%/15% (Table 4 and Supplementary Tables S1-S3 in Additional file 3). When  $x = 0$ , the empirical SE and MSE of the NB, two-part models and three-part model were similar and they were all lower than those of OLS in all scenarios (Tables 5 and 6 and Supplementary Tables S1-S3 in Additional file 3). On the other hand, when  $x = 30$ , the empirical SE and MSE of OLS were the lowest, followed by the three-part model. Those of NB were the highest. The two-part models (ZTNB and ZG) performed the same using these performance measures.

## Discussion

In this paper, we compared different statistical models for analyzing productivity loss outcomes in RCTs by considering their data distribution characteristics with inflated zero and max values. We focused on five commonly used models, OLS, NB, ZTNB, ZG, and three-part model, adjusting for one covariate. From our simulation results, we found that NB performed worst

overall. Two-part models assuming the second part following zero-truncated NB (ZTNB) or Gamma distribution (ZG) performed the same in all scenarios. The performance of OLS, two-part models and three-part model varied in different scenarios. Based on our results, we provided the following practical recommendations if the treatment effect at any given values of one single covariate is of interest:

1. Check the sample size and the proportions of zero loss and max loss in each arm; If the sample size of each arm (i.e., the number of participants who are working at baseline) is  $\leq 50$  and there are  $\leq 5$  subjects with max loss, three-part model should not be considered. Two-part models (either ZTNB or ZG) should be used if the proportion of zero loss  $> 50\%$  in at least one arm and OLS should be used if the proportion of zero loss  $\leq 50\%$  in both arms.
2. If the sample size of each arm is  $> 50$  and the proportion of max loss  $> 5\%$  with more than 5 subjects, then check the scale parameter of the productivity loss outcome distribution between zero

**Table 5** Empirical standard error for the number of observations each arm = 100

Distribution of productivity loss outcome	Scales in the two arms <sup>a</sup>	OLS	NB	ZTNB	ZG	Three-part
<b>x = mean</b>						
80:15:5/60:30:10	Equal Scale	2.708 (0.027)	2.813 (0.028)	2.714 (0.027)	2.716 (0.027)	2.677 (0.027)
80:15:5/60:30:10	Unequal Scale	2.756 (0.028)	2.867 (0.029)	2.766 (0.028)	2.768 (0.028)	2.718 (0.027)
60:35:5/40:50:10	Equal Scale	2.734 (0.027)	2.793 (0.028)	2.750 (0.028)	2.752 (0.028)	2.671 (0.027)
60:35:5/40:50:10	Unequal Scale	2.744 (0.027)	2.806 (0.028)	2.761 (0.028)	2.763 (0.028)	2.689 (0.027)
50:40:10/30:55:15	Equal Scale	3.025 (0.030)	3.096 (0.031)	3.056 (0.031)	3.058 (0.031)	2.981 (0.030)
50:40:10/30:55:15	Unequal Scale	3.051 (0.031)	3.119 (0.031)	3.082 (0.031)	3.084 (0.031)	3.013 (0.030)
<b>x = 0</b>						
80:15:5/60:30:10	Equal Scale	2.708 (0.027)	2.447 (0.024)	2.414 (0.024)	2.416 (0.024)	2.415 (0.024)
80:15:5/60:30:10	Unequal Scale	2.756 (0.028)	2.499 (0.025)	2.465 (0.025)	2.467 (0.025)	2.456 (0.025)
60:35:5/40:50:10	Equal Scale	2.734 (0.027)	2.448 (0.024)	2.495 (0.025)	2.498 (0.025)	2.504 (0.025)
60:35:5/40:50:10	Unequal Scale	2.744 (0.027)	2.458 (0.025)	2.489 (0.025)	2.492 (0.025)	2.497 (0.025)
50:40:10/30:55:15	Equal Scale	3.025 (0.030)	2.711 (0.027)	2.779 (0.028)	2.782 (0.028)	2.781 (0.028)
50:40:10/30:55:15	Unequal Scale	3.051 (0.031)	2.746 (0.027)	2.818 (0.028)	2.822 (0.028)	2.827 (0.028)
<b>x = 30</b>						
80:15:5/60:30:10	Equal Scale	2.708 (0.027)	4.118 (0.041)	3.346 (0.033)	3.348 (0.033)	3.234 (0.032)
80:15:5/60:30:10	Unequal Scale	2.756 (0.028)	4.100 (0.041)	3.388 (0.034)	3.390 (0.034)	3.266 (0.033)
60:35:5/40:50:10	Equal Scale	2.734 (0.027)	3.613 (0.036)	3.228 (0.032)	3.227 (0.032)	3.019 (0.030)
60:35:5/40:50:10	Unequal Scale	2.744 (0.027)	3.623 (0.036)	3.252 (0.033)	3.250 (0.032)	3.060 (0.031)
50:40:10/30:55:15	Equal Scale	3.025 (0.030)	3.847 (0.038)	3.518 (0.035)	3.516 (0.035)	3.265 (0.033)
50:40:10/30:55:15	Unequal Scale	3.051 (0.031)	3.841 (0.038)	3.523 (0.035)	3.520 (0.035)	3.281 (0.033)

OLS ordinary least squares, NB negative binomial, ZTNB two-part model – logistic regression for the probability of being zero, and generalized linear regression with zero-truncated NB distribution for the non-zeros, ZG two-part model – logistic regression for the probability of being zero, and generalized linear regression with Gamma distribution for the non-zeros; Three-part: multinomial logistic regression for the probabilities of being zero and 60 and generalized linear regression with Beta distribution for the those with values in (0, 60) (transformed to (0, 1))

<sup>a</sup>for truncated negative binomial distributions of productivity loss outcomes in the two arms

and max loss (Group II) for each arm, which was an influencing factor on model performance:

- Our three-part model assumed a Beta distribution for Group II and thus the scale =  $\alpha + \beta$ , where  $\alpha$  and  $\beta$  are Beta distribution parameters derived from the mean and variance of Group II. If the two arms have similar scales, three-part model should be used.
- Otherwise, the two-part models should be used if the proportion of zero loss > 50% in at least one arm and OLS should be used if the proportion of zero loss  $\leq$  50% in both arms.

We chose the baseline productivity loss as a single covariate in our simulations, but the above practical recommendations would apply to any single covariate models for analyzing productivity loss outcomes in an RCT, in which the covariate is well balanced between treatment arms and associated with the productivity loss outcomes.

Our study followed the published guidance by Morris et al. [38] for design, execution, analysis, reporting, and

presentation of simulation studies. To the best of our knowledge, there have not been simulation studies considering data like productivity loss outcome with excess zero and max values. However, some previous simulation studies have compared different statistical models for data with excess zeros and found that zero inflated models or two-part models performed better than Poisson model, NB model or OLS [34, 35]. Our study showed consistent results that two-part models performed better than OLS if data had > 50% zeros in at least one arm of an RCT.

Our study has limitations. First, as mentioned above, we had convergence issues because of quasi-complete separation in simulated databases or their bootstraps. Thus, we did not apply the three-part model for  $N_{obs} = 50$  in scenarios with 5% max loss. However, for  $N_{obs} = 100$  and 200, the number of simulated databases with quasi-complete separation was very small (Table 2), which should have minimal impact on our mean bias estimates. The quasi-complete separation issues detected in bootstraps were also within an acceptable range (Table 3).



**Table 6** Mean squared error for the number of observations each arm = 100

Distribution of productivity loss outcome	Scales in the two arms <sup>a</sup>	OLS	NB	ZTNB	ZG	Three-part
<b>x = mean</b>						
80:15.5/60:30:10	Equal Scale	7.337 (0.146)	7.912 (0.163)	7.369 (0.148)	7.379 (0.148)	7.168 (0.145)
80:15.5/60:30:10	Unequal Scale	7.596 (0.152)	8.221 (0.168)	7.648 (0.153)	7.659 (0.153)	7.434 (0.148)
60:35.5/40:50:10	Equal Scale	7.477 (0.145)	7.800 (0.153)	7.562 (0.147)	7.573 (0.147)	7.139 (0.139)
60:35.5/40:50:10	Unequal Scale	7.538 (0.153)	7.872 (0.157)	7.625 (0.153)	7.636 (0.153)	7.519 (0.149)
50:40:10/30:55:15	Equal Scale	9.167 (0.181)	9.585 (0.190)	9.341 (0.185)	9.355 (0.185)	8.898 (0.175)
50:40:10/30:55:15	Unequal Scale	9.313 (0.182)	9.726 (0.188)	9.494 (0.183)	9.507 (0.184)	9.369 (0.183)
<b>x = 0</b>						
80:15.5/60:30:10	Equal Scale	8.273 (0.164)	6.117 (0.140)	5.846 (0.121)	5.853 (0.121)	5.832 (0.123)
80:15.5/60:30:10	Unequal Scale	8.649 (0.175)	6.323 (0.165)	6.081 (0.124)	6.090 (0.125)	6.053 (0.123)
60:35.5/40:50:10	Equal Scale	7.917 (0.153)	6.260 (0.123)	6.266 (0.123)	6.275 (0.123)	6.275 (0.127)
60:35.5/40:50:10	Unequal Scale	7.960 (0.163)	6.345 (0.129)	6.250 (0.127)	6.259 (0.127)	6.502 (0.131)
50:40:10/30:55:15	Equal Scale	9.228 (0.180)	7.873 (0.158)	7.795 (0.156)	7.811 (0.156)	7.765 (0.156)
50:40:10/30:55:15	Unequal Scale	9.426 (0.184)	7.988 (0.155)	7.985 (0.155)	8.001 (0.156)	8.305 (0.162)
<b>x = 30</b>						
80:15.5/60:30:10	Equal Scale	8.865 (0.173)	18.038 (0.564)	11.210 (0.231)	11.228 (0.232)	10.454 (0.215)
80:15.5/60:30:10	Unequal Scale	8.991 (0.174)	17.965 (0.493)	11.506 (0.238)	11.519 (0.238)	10.718 (0.215)
60:35.5/40:50:10	Equal Scale	8.139 (0.159)	14.363 (0.326)	10.529 (0.212)	10.519 (0.212)	9.113 (0.178)
60:35.5/40:50:10	Unequal Scale	8.219 (0.163)	14.250 (0.322)	10.661 (0.220)	10.643 (0.220)	9.646 (0.196)
50:40:10/30:55:15	Equal Scale	9.347 (0.187)	16.324 (0.365)	12.471 (0.253)	12.456 (0.253)	10.663 (0.211)
50:40:10/30:55:15	Unequal Scale	9.455 (0.185)	16.445 (0.347)	12.530 (0.247)	12.511 (0.247)	11.045 (0.218)

OLS ordinary least squares, NB negative binomial, ZTNB two-part model – logistic regression for the probability of being zero, and generalized linear regression with zero-truncated NB distribution for the non-zeros, ZG two-part model – logistic regression for the probability of being zero, and generalized linear regression with Gamma distribution for the non-zeros; Three-part: multinomial logistic regression for the probabilities of being zero and 60 and generalized linear regression with Beta distribution for the those with values in (0, 60) (transformed to (0, 1))

<sup>a</sup>for truncated negative binomial distributions of productivity loss outcomes in the two arms

Second, we compared five commonly used statistically methods to make more informative and practical recommendations to a broad clinical audience. However, there are other potential methods that could be used, for example, Poisson model, zero-inflated models and other mixture models. We chose NB model, two-part models and three-part model by assuming they would perform similarly to Poisson model, zero-inflated models and other mixture models, respectively.

Third, our simulation parameters were determined based on published RCTs, which were selected after a rapid literature review of RCT studies that measured and reported work productivity loss (absenteeism or presenteeism or both or all three subcomponents). However, the scenarios considered in our simulation study might not cover all possible scenarios of RCTs. Our simulation method can be used to compare statistical models in other scenarios we did not consider in the future.

## Conclusions

In summary, we conducted a simulation study to compare five statistical models for analyzing productivity loss outcomes in RCTs. Our findings suggest that NB model

performs worst. If treatment effect at any given values of a single covariate is of interest, the model selection among OLS, two-part models and three-part model depends on the sample size, the proportions of zero loss and max loss, and the scale of the productivity loss outcome distribution between zero and max loss in each arm of RCTs.

## Abbreviations

MSE: Mean squared error; NB: Negative binomial; OLS: Ordinary least squares; RCT: Randomized controlled trial; SD: Standard deviation; SE: Standard error; WPAI: Work Productivity and Activity Impairment Questionnaire; ZG: Two-part model – logistic regression for the probability of being zero, and generalized linear regression with Gamma distribution for the non-zeros; ZTNB: Two-part model – logistic regression for the probability of being zero, and generalized linear regression with zero-truncated NB distribution (Hurdle) for the non-zeros

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12874-021-01330-w>.

**Additional file 1.** Deviation for the Negative Binomial distribution parameters and SAS codes.

**Additional file 2: Supplementary Figure S1.** Mean bias for the number of observations in each arm = 50. **Supplementary Figure S2.**

Mean bias for the number of observations in each arm = 200. **Supplementary Figure S3.** Mean bias for the number of observations in each arm = 1000. **Supplementary Figure S4.** Mean bias for the number of observations in each arm = 2000.

**Additional file 3: Supplementary Table S1.** Performance measures for the number of observations each arm = 50. **Supplementary Table S2.** Performance measures for the number of observations each arm = 100. **Supplementary Table S3.** Performance measures for the number of observations each arm = 200.

#### Acknowledgements

Not applicable.

#### Authors' contributions

WZ conceived and designed the study, guided the simulations and analyses, and drafted and finalized the writing; HS conducted the simulations and analyses, and drafted and finalized the writing. All authors read and approved the final manuscript.

#### Authors' information

WZ is currently Assistant Professor of the School of Population and Public Health at the University of British Columbia and Program Head of Health Economics at the Centre for Health Evaluation and Outcome Sciences. One of the research areas of WZ is on measurement and valuation of work productivity loss due to health problems. HS is a Senior Statistician at the Centre for Health Evaluation and Outcome Sciences. She has a PhD in Mathematics and a Master's degree in Statistics. WZ and HS have co-worked on studies focusing on measuring and analyzing work productivity loss outcomes in different populations.

#### Funding

This work was supported by BC SUPPORT Unit Methods Cluster Project Award [RWCT-301]. Wei Zhang would like to acknowledge the support of the Michael Smith Foundation for Health Research Scholar Award. The funding body had no role in the design, methods, data collection, analysis, data interpretation, or preparation of the article.

#### Availability of data and materials

All data generated or analyzed during this study and the corresponding SAS codes are included in this published article and its supplementary information files.

#### Declarations

##### Ethics approval and consent to participate

This study was approved by the University of British Columbia-Providence Health Care Research Ethics Board (Ethics Certificate No. (H19-04005). In this study, all our data were based on simulations and we did not directly recruit study participants or had direct contact with them.

##### Consent for publication

Not applicable.

##### Competing interests

The authors declare that they have no competing interests.

Received: 8 March 2021 Accepted: 4 June 2021

Published online: 24 June 2021

#### References

- Kuenzig ME, Lee L, El-Matary W, Weizman AV, Benchimol EI, Kaplan GG, et al. The impact of inflammatory bowel disease in Canada 2018: indirect costs of IBD care. *J Can Assoc Gastroenterol*. 2019;2(Suppl 1):S34–41. <https://doi.org/10.1093/jcag/gwy050>.
- Deb A, Thornton JD, Sambamoorthi U, Innes K. Direct and indirect cost of managing alzheimer's disease and related dementias in the United States. *Expert Rev Pharmacoecon Outcomes Res*. 2017;17(2):189–202. <https://doi.org/10.1080/14737167.2017.1313118>.
- Rice DP. Cost of illness studies: what is good about them? *Injury Prev*. 2000;6(3):177–9. <https://doi.org/10.1136/ip.6.3.177>.
- Centers for Disease Control and Prevention. Cost of Illness. 2019. <https://www.cdc.gov/policy/polaris/economics/cost-of-illness.html>. Accessed 7 Feb 2021.
- Neumann PJ, Sanders GD, Russell LB, Siegel JE, Ganiats TG, editors. Cost-effectiveness in health and medicine. 2nd ed. Oxford; New York: Oxford University Press; 2016. <https://doi.org/10.1093/acprof:oso/9780190492939.001.0001>.
- Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes. 4th ed. Oxford; New York: Oxford University Press; 2015.
- Aggarwal J, Vera-Llonch M, Donepudi M, Suthoff E, Younossi Z, Goss TF. Work productivity among treatment-naïve patients with genotype 1 chronic hepatitis C infection receiving telaprevir combination treatment. *J Viral Hepat*. 2014;22:8–17.
- Viestar L, Verhagen EALM, Bongers PM, van der Beek AJ. The effect of a health promotion intervention for construction workers on work-related outcomes: results from a randomized controlled trial. *Int Arch Occup Environ Health*. 2015;88(6):789–98. <https://doi.org/10.1007/s00420-014-1007-9>.
- Arends I, Van Der Klink JLL, Van Rhenen W, De Boer MR, Bültmann U. Prevention of recurrent sickness absence in workers with common mental disorders: results of a cluster-randomised controlled trial. *Occup Environ Med*. 2014;71(1):21–9. <https://doi.org/10.1136/oemed-2013-101412>.
- Soler-Font M, Ramada JM, Van Zon SKR, Almansa J, Bültmann U, Serra C, et al. Multifaceted intervention for the prevention and management of musculoskeletal pain in nursing staff: results of a cluster randomized controlled trial. *PLoS One*. 2019;14:1–16.
- Keyser JJ, LaValley MP, Brown C, Felson DT, AlHeresh RA, Vaughan MW, et al. Efficacy of a work disability prevention program for people with rheumatic and musculoskeletal conditions: a single-blind parallel-arm randomized controlled trial. *Arthritis Care Res*. 2018;70(7):1022–9. <https://doi.org/10.1002/acr.23423>.
- Brämberg EB, Bergström G, Jensen I, Hagberg J, Kwak L. Effects of yoga, strength training and advice on back pain: a randomized controlled trial. *BMC Musculoskelet Disord*. 2017;18:1–11.
- van Vollenhoven RF, Cifaldi MA, Ray S, Chen N, Weisman MH. Improvement in work place and household productivity for patients with early rheumatoid arthritis treated with adalimumab plus methotrexate: work outcomes and their correlations with clinical and radiographic measures from a randomized controlled trial companion study. *Arthritis Care Res*. 2010;62(2):226–34. <https://doi.org/10.1002/acr.20072>.
- Zhang W, Bansback N, Sun H, Pedersen R, Kotak S, Anis AH. Impact of etanercept tapering on work productivity in patients with early rheumatoid arthritis: results from the PRIZE study. *RMD Open*. 2016;2(2):e000222. <https://doi.org/10.1136/rmdopen-2015-000222>.
- Duijzer G, Bukman AJ, Meints-Groenveld A, Haveman-Nies A, Jansen SC, Heinrich J, et al. Cost-effectiveness of the SLIMMER diabetes prevention intervention in Dutch primary health care: economic evaluation from a randomised controlled trial. *BMC Health Serv Res*. 2019;19:1–10.
- Chatterton ML, Mihalopoulos C, O'Neil A, Itsiopoulos C, Opie R, Castle D, et al. Economic evaluation of a dietary intervention for adults with major depression (the "SMILES" trial). *BMC Public Health*. 2018;18:1–12.
- Goorden M, Huijbregts KML, van Marwijk HWJ, Beekman ATF, van der Feltz-Cornelis CM, Hakkaart-van RL. Cost-utility of collaborative care for major depressive disorder in primary care in the Netherlands. *J Psychosom Res*. 2015;79(4):316–23. <https://doi.org/10.1016/j.jpsychores.2015.06.006>.
- Bouwisma EVA, Bosmans JE, Van Dongen JM, Brölmann HAM, Anema JR, Huijre JAF. Cost-effectiveness of an internet-based perioperative care programme to enhance postoperative recovery in gynaecological patients: economic evaluation alongside a stepped-wedge cluster-randomised trial. *BMJ Open*. 2018;8:1–11.
- Zhang W, Bansback N, Anis AH. Measuring and valuing productivity loss due to poor health: a critical review. *Soc Sci Med*. 2011;72(2):185–92. <https://doi.org/10.1016/j.socscimed.2010.10.026>.
- Reilly MC, Zbrozek AS, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument. *Pharmacoeconomics*. 1993;4(5):353–65. <https://doi.org/10.2165/00019053-199304050-00006>.
- Reilly Associates. WPAI Studies. [http://www.reillyassociates.net/WPAI\\_References.html](http://www.reillyassociates.net/WPAI_References.html). Accessed 8 Aug 2017.
- Zhang W, Bansback N, Boonen A, Severens JL, Anis AH. Development of a composite questionnaire, the valuation of lost productivity, to value

- productivity losses: application in rheumatoid arthritis. *Value Health*. 2012; 15(1):46–54. <https://doi.org/10.1016/j.jval.2011.07.009>.
23. Zhang W, Bansback N, Kopec J, Anis AH. Measuring time input loss among patients with rheumatoid arthritis: validity and reliability of the valuation of lost productivity questionnaire. *J Occup Environ Med*. 2011;53(5):530–6. <https://doi.org/10.1097/JOM.0b013e318218abf1>.
  24. Li X, Gignac MAM, Anis AH. The indirect costs of arthritis resulting from unemployment, reduced performance, and occupational changes while at work. *Med Care*. 2006;44(4):304–10. <https://doi.org/10.1097/01.mlr.0000204257.25875.04>.
  25. Zhang W, Gignac MAM, Beaton D, Tang K, Anis AH. Productivity loss due to presenteeism among patients with arthritis: estimates from 4 instruments. *J Rheumatol*. 2010;37(9):1805–14. <https://doi.org/10.3899/jrheum.100123>.
  26. Zhang W, Bansback N, Guh D, Li X, Nosyk B, Marra CA, et al. Short-term influence of adalimumab on work productivity outcomes in patients with rheumatoid arthritis. *J Rheumatol*. 2008;35(9):1729–36.
  27. Zhang W, Bansback N, Sun H, Pedersen R, Kotak S, Anis AH. Estimating the monetary value of the annual productivity gained in patients with early rheumatoid arthritis receiving etanercept plus methotrexate: interim results from the PRIZE study. *RMD Open*. 2015;1(1):e000042. <https://doi.org/10.1136/rmdopen-2014-000042>.
  28. Lerner D, Reed JI, Massarotti E, Wester LM, Burke TA. The work limitations Questionnaire's validity and reliability among patients with osteoarthritis. *J Clin Epidemiol*. 2002;55(2):197–208. [https://doi.org/10.1016/S0895-4356\(01\)00424-3](https://doi.org/10.1016/S0895-4356(01)00424-3).
  29. Kessler RC, Greenberg PE, Mickelson KD, Meneades LM, Wang PS. The effects of chronic medical conditions on work loss and work cutback. *J Occup Environ Med*. 2001;43(3):218–25. <https://doi.org/10.1097/00043764-200103000-00009>.
  30. Puolakka K, Kautiainen H, Möttönen T, Hannonen P, Korpela M, Hakala M, et al. Early suppression of disease activity is essential for maintenance of work capacity in patients with recent-onset rheumatoid arthritis: five-year experience from the FIN-RACo trial. *Arthritis Rheum*. 2005;52(1):36–41. <https://doi.org/10.1002/art.20716>.
  31. Geuskens GA, Hazes JMW, Barendregt PJ, Burdorf A. Predictors of sick leave and reduced productivity at work among persons with early inflammatory joint conditions. *Scand J Work Environ Health*. 2008;34(6):420–9. <https://doi.org/10.5271/sjweh.1298>.
  32. Gilthorpe MS, Frydenberg M, Cheng Y, Baelum V. Modelling count data with excessive zeros: the need for class prediction in zero-inflated models and the issue of data generation in choosing between zero-inflated and generic mixture models for dental caries data. *Stat Med*. 2009;28(28):3539–53. <https://doi.org/10.1002/sim.3699>.
  33. Sarma S, Simpson W. A microeconomic analysis of Canadian health care utilization. *Health Econ*. 2006;15(3):219–39. <https://doi.org/10.1002/hec.1057>.
  34. Erdman D, Jackson L, Sinko A. Zero-inflated Poisson and zero-inflated negative binomial models using the COUNTREG procedure. Paper 322-2008, SAS Conference Proceedings: SAS Global Forum 2008, San Antonio. 2008. <https://support.sas.com/resources/papers/proceedings/pdfs/sgf2008/322-2008.pdf>. Accessed 30 Aug 2020.
  35. Yang S, Harlow L, Puggioni G, Redding C. A comparison of different methods of zero-inflated data analysis and an application in health surveys. *J Mod Appl Stat Methods*. 2017;16(1):518–43. <https://doi.org/10.22237/jmasm/1493598600>.
  36. Basu A, Manca A. Regression estimators for generic health-related quality of life and quality-adjusted life years. *Med Decis Mak*. 2012;32(1):56–69. <https://doi.org/10.1177/0272989X11416988>.
  37. Kleinman NL, Brook RA, Patel PA, Melkonian AK, Brizee TJ, Smeeding JE, et al. The impact of gout on work absence and productivity. *Value Health*. 2007;10(4):231–7. <https://doi.org/10.1111/j.1524-4733.2007.00173.x>.
  38. Morris TP, White IR, Crowther MJ. Using simulation studies to evaluate statistical methods. *Stat Med*. 2019;38(11):2074–102. <https://doi.org/10.1002/sim.8086>.
  39. Anis A, Zhang W, Emery P, Sun H, Singh A, Freundlich B, et al. The effect of etanercept on work productivity in patients with early active rheumatoid arthritis: results from the COMET study. *Rheumatology (Oxford)*. 2009;48(10):1283–9. <https://doi.org/10.1093/rheumatology/kep239>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

