STATISTICS IN MEDICINE
Statist. Med. 2007; **26**:5586–5595
Published online 18 September 2007 in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sim.3061

Estimating adjusted NNT measures in logistic regression analysis

Ralf Bender[1,2,*,†,‡], Oliver Kuss[3], Mandy Hildebrandt[1,4] and Ulrich Gehrmann[1]

[1]*Department of Medical Biometry, Institute for Quality and Efficiency in Health Care (IQWiG), Cologne, Germany*
[2]*Faculty of Medicine, University of Cologne, Germany*
[3]*Institute for Medical Epidemiology, Biostatistics, and Informatics (IMEBI), Martin-Luther-University Halle-Wittenberg, Halle (Saale), Germany*
[4]*Institute for Medical Biometry, Epidemiology and Informatics (IMBEI), Johannes-Gutenberg-University Mainz, Mainz, Germany*

SUMMARY

The number needed to treat (NNT) is a popular measure to describe the absolute effect of a new treatment compared with a standard treatment or placebo in clinical trials with binary outcome. For use of NNT measures in epidemiology to compare exposed and unexposed subjects, the terms 'number needed to be exposed' (NNE) and 'exposure impact number' (EIN) have been proposed. Additionally, in the framework of logistic regression a method was derived to perform point and interval estimation of NNT measures with adjustment for confounding by using the adjusted odds ratio (OR approach). In this paper, a new method is proposed which is based upon the average risk difference over the observed confounder values (ARD approach). A decision has to be made, whether the effect of allocating an exposure to unexposed persons or the effect of removing an exposure from exposed persons should be described. We use the term NNE for the first and the term EIN for the second situation. NNE is the average number of unexposed persons needed to be exposed to observe one extra case; EIN is the average number of exposed persons among one case can be attributed to the exposure. By means of simulations it is shown that the ARD approach is better than the OR approach in terms of bias and coverage probability, especially if the confounder distribution is wide. The proposed method is illustrated by application to data of a cohort study investigating the effect of smoking on coronary heart disease. Copyright © 2007 John Wiley & Sons, Ltd.

KEY WORDS: confounding; epidemiology; exposure impact number (EIN); logistic regression; number needed to be exposed (NNE); number needed to treat (NNT)

*Correspondence to: Ralf Bender, Department of Medical Biometry, Institute for Quality and Efficiency in Health Care (IQWiG), D-51105 Cologne, Germany.
†E-mail: Ralf.Bender@iqwig.de, Ralf@rbsd.de
‡Statistician.

Contract/grant sponsor: German Research Foundation (DFG); contract/grant number: BL 443/5-1

Received 8 August 2007
Accepted 8 August 2007

1. INTRODUCTION

The number needed to treat (NNT) is a popular measure to describe the absolute effect of a new treatment compared with a standard treatment or placebo in clinical trials with binary outcome [1]. Recently, NNT measures have also been developed for use in case–control [2, 3] and cohort studies [3, 4]. As the term 'number needed to treat' makes no sense if the explanatory factor is an exposure rather than a treatment, the terms number needed to be exposed (NNE) [4] and exposure impact number (EIN) [3] have been proposed to apply the NNT concept in epidemiological studies. Regardless of terminology, in the simplest case NNT measures (NNT, NNE, EIN) are calculated by taking the reciprocal of the difference of two risks given by a 2×2 table. The use of simple 2×2 tables may be appropriate in randomized controlled trials. However, in observational studies usually confounding factors have to be taken into account to minimize bias. Bender and Blettner [4] have derived a method based upon multiple logistic regression analysis to perform point and interval estimation of NNT measures with adjustment for confounding factors. In this approach, adjusted NNT measures are calculated from the adjusted odds ratio (OR) and the mean risk of unexposed persons estimated by means of logistic regression (OR approach). In this paper, a new method is proposed, which takes the distribution of the considered confounders into account by using the average risk difference (ARD) estimated from multiple logistic regression (ARD approach). A simulation study is performed to compare the two approaches for calculating adjusted NNT measures. The new ARD approach for computing point and interval estimates of adjusted NNT measures is illustrated by using data of a cohort study investigating the effect of smoking on coronary heart disease (CHD) in white males (Evans County Cohort Study).

2. ADJUSTED NNT MEASURES

2.1. OR approach

Let π_0 and π_1 be the proportions of a control and experimental (treatment, exposure) group, respectively, which experience an outcome event. We consider the typical situation in epidemiological studies, where the outcome is an adverse event such as death or disease and the risk of the exposed persons is higher than that of the unexposed persons, i.e. $\pi_1>\pi_0$. In this case, $\mathrm{NNE}=1/(\pi_1-\pi_0)$ represents the number needed to be exposed for one person to be harmed. More precisely, NNE is the expected number of persons who must be exposed in order to have one additional event in the group of exposed persons compared with the group of unexposed persons. If we denote the odds ratio by $\mathrm{OR}=[\pi_1\times(1-\pi_0)]/[\pi_0\times(1-\pi_1)]$, then the relationship

$$\mathrm{NNE}=\frac{1}{(\mathrm{OR}-1)\times\pi_0}+\frac{\mathrm{OR}}{(\mathrm{OR}-1)\times(1-\pi_0)} \tag{1}$$

between NNE, OR and π_0 holds [4].

On the basis of formula (1), Bender and Blettner [4] derived a method to estimate NNE with adjustment for confounding factors by using the observed unexposed event rate (UER) for π_0 and the adjusted OR estimated by multiple logistic regression for OR (OR approach). Within the framework of logistic regression, the adjusted NNE can be presented as a function of the logistic regression coefficients [4]. Thus, the multivariate delta method [5] can be used to calculate approximate confidence intervals (CIs) for the adjusted NNE [4, 6].

DOI: 10.1002/sim

2.2. ARD approach

The drawback of the OR approach is that the adjusted NNE is calculated only at one point of the risk relation, namely for the observed UER. In the logistic regression model the risk of the unexposed persons is dependent on the confounders. The observed UER represents the mean of the risks of the unexposed persons over the distribution of the confounder values. As such, the OR approach is appropriate only if there is a small variation of the risks around the mean. However, if the distribution of the confounders is wide, this variation has to be taken into account in estimating adjusted NNT measures. The main idea is to use the expected difference of the risks with and without exposure concerning the distribution of the confounders. The reciprocal of the expected risk difference (ERD) represents the adjusted NNT measure. However, as the distribution of confounders may be different between unexposed and exposed persons, two different measures are obtained.

Consider two populations of exposed and unexposed persons, where Z is the binary exposure status with value z ($1=$exposed, $0=$unexposed). Let $(X_1,\ldots,X_k)$ be a vector of k binary or continuous confounder variables with values $(x_1,\ldots,x_k)$ and $F^{\circ}(x_1,\ldots,x_k)$ and $F^{\bullet}(x_1,\ldots,x_k)$ the distribution functions for the unexposed and the exposed persons, respectively. Furthermore, let $\pi(x_1,\ldots,x_k,z)$ describe the risk depending on the confounders and the exposure status. Then the ERD concerning the exposure status can be defined for each population by

$$\mathrm{ERD}^{\circ}=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}(\pi(x_1,\ldots,x_k,1)-\pi(x_1,\ldots,x_k,0))\,\mathrm{d}F^{\circ}(x_1,\ldots,x_k) \tag{2}$$

$$\mathrm{ERD}^{\bullet}=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}(\pi(x_1,\ldots,x_k,1)-\pi(x_1,\ldots,x_k,0))\,\mathrm{d}F^{\bullet}(x_1,\ldots,x_k) \tag{3}$$

It should be noted that in formula (2), $\pi(x_1,\ldots,x_k,1)$ represents a hypothetical risk, because in the population of unexposed persons nobody is exposed. Accordingly, in formula (3), $\pi(x_1,\ldots,x_k,0)$ is a hypothetical risk, because in the population of exposed persons nobody is unexposed.

The reciprocals of ERD° and $\mathrm{ERD}^{\bullet}$ represent appropriate NNT measures to describe the impact of the exposure for the populations of unexposed and exposed persons, respectively. For application, a decision has to be made whether the effect of the exposure should be described for the confounder distribution of the unexposed or the exposed persons. We propose to use the term 'number needed to be exposed' (NNE) for the first and the term 'exposure impact number' (EIN) for the second situation, i.e.

$$\mathrm{NNE}=\frac{1}{\mathrm{ERD}^{\circ}} \tag{4}$$

$$\mathrm{EIN}=\frac{1}{\mathrm{ERD}^{\bullet}} \tag{5}$$

NNE refers to the situation where the effect of an exposure in a population of unexposed persons is described. On the other hand, the term EIN was proposed for situations where the effect of removing an exposure from the population is considered [3]. Thus, the appropriate reference group is the population of exposed persons. In the case of no confounders, NNE and EIN are identical. However, in the case of confounding with different distributions of the confounders in unexposed and exposed persons, NNE and EIN may be different.

DOI: 10.1002/sim

For practical applications, formulas to calculate the risks in dependence on the confounders and the exposure status are required. Let n_0 and n_1 be the numbers of the unexposed and exposed persons, respectively, i.e. $n_0+n_1=n$. Let z_i and $x_{1i},\ldots,x_{ki}$ be the observed values of the exposure Z and the confounders $X_1,\ldots,X_k$ for $i=1,\ldots,n_0,n_0+1,\ldots,n_0+n_1=n$. By using logistic regression in a prospective study design, the risk π_i of the outcome event for person i is given by

$$\pi_i=\frac{\exp(\alpha+\gamma z_i+\beta_1 x_{1i}+\cdots+\beta_k x_{ki})}{1+\exp(\alpha+\gamma z_i+\beta_1 x_{1i}+\cdots+\beta_k x_{ki})} \quad \text{for } i=1,\ldots,n_0,n_0+1,\ldots,n_0+n_1=n \tag{6}$$

where $\alpha,\gamma,\beta_1,\ldots,\beta_k$ are the logistic regression coefficients. If these coefficients were known, we could calculate the hypothetical risk for each unexposed person if this person would be exposed by

$$\pi_i^{\circ}=\frac{\exp(\alpha+\gamma+\beta_1 x_{1i}+\cdots+\beta_k x_{ki})}{1+\exp(\alpha+\gamma+\beta_1 x_{1i}+\cdots+\beta_k x_{ki})} \quad \text{for } i=1,\ldots,n_0 \tag{7}$$

On the other hand, we could also calculate the hypothetical risk for each exposed person if this person would not be exposed by

$$\pi_i^{\bullet}=\frac{\exp(\alpha+\beta_1 x_{1i}+\cdots+\beta_k x_{ki})}{1+\exp(\alpha+\beta_1 x_{1i}+\cdots+\beta_k x_{ki})} \quad \text{for } i=n_0+1,\ldots,n_0+n_1 \tag{8}$$

For large samples, the ERD in (2) and (3) can be approximated by the corresponding ARDs

$$\mathrm{ARD}^{\circ}=\frac{1}{n_0}\sum_{i=1}^{n_0}(\pi_i^{\circ}-\pi_i) \tag{9}$$

$$\mathrm{ARD}^{\bullet}=\frac{1}{n_1}\sum_{i=n_0+1}^{n_0+n_1}(\pi_i-\pi_i^{\bullet}) \tag{10}$$

The estimation of ARD° and $\mathrm{ARD}^{\bullet}$ can be performed by substituting the usual estimates of the logistic regression coefficients into formulas (6)–(8). The estimates of NNE and EIN are obtained by taking the reciprocals of the estimates of ARD° and $\mathrm{ARD}^{\bullet}$ (ARD approach).

2.3. Confidence intervals

In general, the distribution of estimated NNT measures cannot be approximated by the normal distribution [7, 8]. Thus, CIs for NNT measures are frequently calculated indirectly via CIs for the corresponding risk difference [9]. Obviously, ARD° and $\mathrm{ARD}^{\bullet}$ represent functions of the logistic regression coefficients. Thus, the standard errors of these risk differences can be calculated by using the multivariate delta method [5] in the same way as described for the OR approach [4]. CIs for the adjusted NNE and EIN can then be calculated by inverting and exchanging the corresponding confidence limits for ARD° and $\mathrm{ARD}^{\bullet}$. The unusual scale of the NNT measure, especially the fact that the point of the zero effect is given by infinity, has to be taken into account. If the CI for ARD encloses 0, the corresponding confidence region for the adjusted NNE or EIN is given by the union of two half intervals [7, 9, 10]. For example, a 95 per cent CI for ARD of $[-0.2, 0.2]$ corresponds to the confidence region of $[5,\infty[\cup]-\infty,-5]$ for NNE or EIN. Bender and Blettner [4] with reference to Altman [10] proposed to write such confidence regions in the form of 'NNEH 5 to ∞ to NNEB 5', where NNEH (NNEB) means 'number needed to be exposed for one person

DOI: 10.1002/sim

to be harmed' (benefit). To avoid an additional abbreviation we use the abbreviation CI also for confidence regions. An SAS/IML® program to calculate adjusted NNEs and EINs with 95 per cent CIs on the basis of multiple logistic regression for use in cohort studies is available on the Internet (http://www.rbsd.de/SOFTWARE/nne_ein.sas).

3. SIMULATION STUDY

In order to investigate the features of the new ARD approach in comparison with the OR approach, a simulation study was performed. Without loss of generality, only the measure NNE (4) was considered. We generated data similar to a typical cohort study in which the association between a binary outcome Y and a binary exposure Z for a fixed follow-up time, say 5 years, is investigated and a continuous confounder X, say age, is taken into account. The sample size was set to $n=2000$ ($n_1=1000$ exposed and $n_0=1000$ unexposed persons). Values for the considered variables were generated as follows. The confounder X was normally distributed with varying standard deviations $\sigma=1,2,3,5,8$ and mean 40 for the exposed persons and mean 45 for the unexposed persons. As the exposed persons are on average 5 years younger than the unexposed persons, age is a strong confounder in this situation. The probability of an outcome event was modelled by means of the logistic regression equation

$$\pi=\frac{\exp(\alpha+\gamma z+\beta x)}{1+\exp(\alpha+\gamma z+\beta x)} \tag{11}$$

with $\alpha=-10$, $\gamma=1.0986$ and $\beta=0.1823$, which means that the true OR for the exposure is given by OR$=3$ and the true OR for the confounder age is given by OR$=1.2$ per year. The response Y was generated by assigning the value 1 (event) with probability π and the value 0 (no event) with probability $1-\pi$. For each of the five situations with different values for σ 1000 replications were performed.

For each generated data set we calculated the true adjusted NNE on the basis of the true logistic regression coefficients. For estimation of the model parameters the SAS® procedure PROC LOGISTIC was used (SAS® version 9.1.3). On the basis of the estimates of the adjusted OR and the logistic regression coefficients, we estimated the adjusted NNE with 95 per cent CIs by means of the OR and the ARD approach, respectively. We expressed the relative bias of the adjusted estimates by calculating the mean percent error (MPE)

$$\text{MPE}=100\times\frac{1}{r}\sum_{j=1}^{r}\frac{\hat{\theta}_j-\theta}{\theta} \tag{12}$$

where θ is the parameter of interest, $\hat{\theta}_j$ its estimate in replication j and $r=1000$ (number of replications). It should be noted that the calculation of the MPE is not meaningful if negative NNE estimates occur. Thus, we used only positive parameter estimates in formula (12), which is a suitable approach if negative estimates occur only rarely. We calculated the empirical coverage probability (CP) of the approximate 95 per cent CIs for the adjusted NNE (relative frequency of CIs containing the true parameter) and the average width (AW) from the differences between the upper and lower confidence limits. As the AW cannot be calculated if the upper confidence limit is negative, this calculation was restricted to simulations with positive upper limits. Additionally, the median width (MW) of all CIs (including those with a negative upper limit) was calculated,

DOI: 10.1002/sim

Table I. Bias of crude and adjusted NNE estimates for the OR and the ARD approach.

					Adjusted estimates based upon logistic regression				
	True parameters		Crude estimates			OR approach		ARD approach	
σ	OR	NNE	OR	NNE	OR	NNE	MPE (per cent)	NNE	MPE (per cent)
1	3	5.27	1.21	43.9	3.20	6.16*	16.9*	6.20*	17.7*
2	3	5.27	1.18	42.2	3.06	5.38	2.1	5.55	5.3
3	3	5.31	1.20	52.6	3.05	5.05	−4.8	5.42	2.1
5	3	5.48	1.20	46.2	3.02	4.76	−13.1	5.61	2.2
8	3	6.13	1.28	30.6	3.04	4.39	−28.5	6.21	1.2

The table shows the mean of the estimates based upon 1000 simulation runs of generated cohort data with sample size $n=2000$ and the mean percent error (MPE) as a measure of the relative bias of the adjusted NNE estimates; σ represents the standard deviation of the normally distributed continuous confounder.

*In 1 simulation run a negative NNE estimate was obtained. This calculation is based upon 999 simulation runs with positive NNE estimates.

Table II. Coverage probability and width of approximate 95 per cent confidence intervals or regions for the adjusted NNE based upon the OR and the ARD approach.

		OR approach			ARD approach		
σ	True NNE	CP (per cent)	AW	MW	CP (per cent)	AW	MW
1	5.27	94.2	50.88*	17.17	94.2	31.16†	16.42
2	5.27	93.5	6.57	5.39	93.6	6.32	5.18
3	5.31	94.9	3.80	3.49	95.5	3.71	3.41
5	5.48	82.0	2.88	2.70	95.2	3.00	2.83
8	6.13	34.1	2.41	2.27	95.6	3.00	2.84

The table shows the empirical coverage probability (CP), the average width (AW) and the median width (MW) of approximate 95 per cent confidence intervals or regions (CIs) for the adjusted NNE based upon 1000 simulation runs of generated cohort data with sample size n=2000; σ represents the standard deviation of the normally distributed continuous confounder.

*In 172 simulation runs a negative upper confidence limit for the adjusted NNE was obtained. This calculation is based upon 828 simulation runs with positive upper confidence limits.

†In 171 simulation runs a negative upper confidence limit for the adjusted NNE was obtained. This calculation is based upon 829 simulation runs with positive upper confidence limits.

which was possible because the proportion of negative upper confidence limits was lower than 50 per cent in all cases. The main results of the simulations are summarized in Tables I and II.

Owing to the strong confounding effect of age, the crude estimates of OR and NNE show a large bias so that the crude estimates are unacceptable. The adjusted ORs estimated by means of multiple logistic regression still show a small bias which is, however, negligible in practice (Table I). The two methods for estimating the adjusted NNE show a large upward bias in simulation setting 1 with a very low variability of the continuous confounder ($\sigma=1$). However, this situation is unrealistic, because this would mean that 95 per cent of the exposed persons are between 38 and 42 years of age and 95 per cent of the unexposed persons are between 43 and 47 years of age (or a similar confounder). A regression model with a continuous covariate having such a low overlap

DOI: 10.1002/sim

is unstable and its estimation should be avoided. Nevertheless, the empirical CP is quite close to the nominal level of 95 per cent in this situation for both approaches (Table II). With increasing σ the upward bias of the adjusted NNE estimates disappears for both methods. For the OR approach only situations 2 and 3 ($\sigma=2,3$) have acceptable bias values (|MPE|<5 per cent), but for larger values of σ the upward bias switches to a downward bias reaching the bias of MPE$=-28.5$ per cent for the situation with very large confounder variability ($\sigma=8$). On the other hand, the ARD approach shows an upward bias in all cases which decreases with increasing σ and is below 5 per cent if the confounder variability is not very small (Table I).

With increasing σ the empirical CP for the OR approach decreases and reaches only 34.1 per cent in situation 5 ($\sigma=8$) because of the strong downward bias. On the other hand, the ARD approach shows a CP quite close to the nominal level of 95 per cent in all cases. With the exception of situation 1 ($\sigma=1$) the AW and MW of the CIs are small in all cases for both methods. The average CI width of the ARD approach is lower than that of the OR approach with the exception of the cases where the OR approach shows a large downward bias and an unacceptably low CP (Table II).

In summary, the OR approach leads to reasonable results only in the case of continuous confounders with small variability, whereas the new ARD approach is valid in all cases where the underlying multiple logistic regression model has adequate features.

4. EXAMPLE: THE EVANS COUNTY COHORT STUDY

For illustration, we reanalyse the data of the Evans County Cohort Study [11]. In this study, 609 white males were followed for 7 years to investigate the effect of several covariates on the outcome CHD. We applied the logistic regression model as described by Kleinbaum and Klein [11], but focus on smoking as exposure variable and use catecholamine level, age, cholesterol, electrocardiogram abnormality and high blood pressure as confounders. Additionally, the two interaction terms between catecholamine level and cholesterol and between catecholamine level and high blood pressure are included (see [11]). The estimated adjusted OR for smoking estimated by the corresponding multiple logistic regression model is given by $\widehat{\mathrm{OR}}=2.17$ (95 per cent CI: 1.14–4.12, $p=0.018$).

To describe the absolute effect of smoking on the outcome CHD adjusted NNT measures can be used. By applying the ARD approach the estimated ARD based upon the distribution of confounders of the non-smoking males is given by $\widehat{\mathrm{ARD}}^{\circ}=0.0597$ (SE$=0.0235$, 95 per cent CI: 0.0138–0.1057) leading to $\widehat{\mathrm{NNE}}=16.74$ (95 per cent CI: 9.46–72.67). The estimated ARD based upon the distribution of confounders of the smoking males is given by $\widehat{\mathrm{ARD}}^{\bullet}=0.0573$ (SE$=0.0223$, 95 per cent CI: 0.0135–0.1011) leading to $\widehat{\mathrm{EIN}}=17.44$ (95 per cent CI: 9.89–73.81).

These results mean that on average 16–17 males from the population of non-smoking males are needed to be exposed to smoking to observe one extra case of CHD in 7 years. On the other hand, on average 17–18 males from a population with a confounder distribution equal to that of the smoking males are needed for one extra case of CHD in 7 years compared with non-smoking males with the same confounder distribution. As there is not much difference between NNE and EIN in this example, we can say that on average, among 16–18 males followed for 7 years one case of CHD can be attributed to smoking adjusted for all considered confounders.

DOI: 10.1002/sim

5. DISCUSSION

The use of NNT measures to describe the absolute effect of treatments or exposures has increased in the past years [1]. Especially, the demand for NNT measures also in non-randomized studies requested the development of methods to estimate adjusted NNT measures. Unfortunately, although multiple logistic regression is routinely used to estimate adjusted ORs, in current practice adjusted ORs are frequently complemented by crude NNT estimates in medical journals. For example, in the percutaneous coronary intervention (PCI)-CLARITY trial [12] the effect of clopidogrel pretreatment before PCI compared with clopidogrel treatment initiated at the time of PCI on the composite outcome cardiovascular death, recurrent myocardial infarction or stroke was investigated. Because this trial does not represent a randomized study, multiple logistic regression was used to adjust for possible selection bias [12]. Although adjusted ORs based upon the multiple logistic regression model were presented, calculation of NNTs was performed on the basis of the crude event rates without adjustment for possible selection bias. The reason for such invalid presentations of study results is probably that established statistical methods to calculate adjusted effect measures are available for relative effect measures such as the OR or the hazard ratio but not for NNT measures. This underlines the need to develop and apply statistical methods for the calculation of adjusted NNT measures.

In this paper, a new approach to estimate adjusted NNT measures based upon the reciprocal of ARDs within the framework of multiple logistic regression is developed. The method is applicable in all prospective studies where multiple logistic regression is an adequate method for data analysis, such as cohort studies, non-randomized controlled trials and randomized controlled studies with fixed follow-up time, in which the consideration of multiple predictor variables seems to be appropriate. We have shown by means of simulations that the ARD approach is preferable to the formerly proposed OR approach. The OR approach does not take the variability of continuous confounders into account and leads to a downward bias of NNTs (overestimation of the effect) if the confounder distribution is not quite narrow. The ARD approach outperforms the OR approach in terms of bias of the point estimates as well as CP of 95 per cent CIs. In cases with adequate CP, the AW and MW of the 95 per cent CIs of the ARD approach are slightly smaller than those of the OR approach showing a better estimation precision of the new method.

One disadvantage of using NNT measures is given by the unusual scale leading to the possible problem of negative estimates and negative upper confidence limits. In the case of negative parameter estimates the usual formulas to calculate bias are not applicable and in the case of negative upper confidence limits the width of the CIs cannot be calculated. In this paper we restricted these calculations to simulations with positive parameter estimates and positive upper confidence limits, which was suitable because negative estimates and negative upper confidence limits occurred only in very few simulation runs. In simulation studies considering other data situations, e.g. with small effects or low sample size, the problem may be more pronounced. However, in this paper the main aim was to propose a method to calculate adjusted NNT measures in situations where the use of NNT measures is suitable (large exposure effect, sufficient sample size).

Absolute risk differences and NNT measures are frequently sensitive to changes of covariate values and changes of the distribution of covariates in the considered population. This important issue should be taken into account when NNT measures are used to present study results. An illustrative example is given by Bender and Blettner [4], who calculated varying NNEs between 70 and 6 for age values between 30 and 50. For application of the ARD approach in epidemiology and public health research to calculate adjusted NNT measures, a decision has to be made whether

DOI: 10.1002/sim

the exposure effect should be described for the confounder distribution of the unexposed persons (NNE) or that of the exposed persons (EIN). The use of NNE is appropriate if the effect of allocating an exposure to unexposed persons should be described. The second alternative is appropriate in the situation where the effect of removing the exposure from the exposed persons in the population is considered. This is the same rationale as in situations where the population attributable risk is applied [13]. Alternative impact numbers to describe the absolute effect of exposures from a population point of view have been proposed [3, 14]. Methods for interval estimation of impact numbers on the basis of simple 2×2 tables have been summarized [15]. If these impact numbers are to be applied in the case of confounding, the ARD approach developed here should be adapted to perform point and interval estimation of adjusted impact numbers.

An alternative to the use of logistic regression is given by the linear probability model [16]. An advantage of this model is that risk differences can be estimated directly. Thus, adjusted NNT measures as inverse values of risk differences are independent of the confounder values so that the linear probability model is an interesting model for the estimation of adjusted NNT measures, especially if a logistic regression model seems to be inappropriate for the considered data. The choice of a specific model should be based upon a careful investigation of the goodness of fit [11, 17, 18]. In this paper, a method is developed to perform point and interval estimation of adjusted NNT measures within the framework of multiple logistic regression, which represents the most frequent regression model in cohort studies with fixed follow-up times. We will consider the linear probability model as basis for the estimation of adjusted NNT measures in future studies.

The current practice of presenting adjusted ORs together with crude NNTs in medical research papers should be avoided. If—supplementary to relative effect measures—NNT measures are used to present study results and if it is required to take confounders into account, the use of adjusted NNT measures is required. On the basis of our simulation study, the proposed ARD approach is an appropriate method to estimate adjusted NNT measures in all situations where multiple logistic regression is an adequate method for data analysis.

ACKNOWLEDGEMENTS

This work is supported in part by German Research Foundation (DFG) grant BL 443/5-1. We thank the anonymous referees and the special issue editor for the helpful comments.

REFERENCES

1. Bender R. Number needed to treat (NNT). In *Encyclopedia of Biostatistics* (2nd edn), vol. 6. Armitage P, Colton T (eds). Wiley: Chichester, 2005; 3752–3761.
2. Bjerre LM, LeLorier L. Expressing the magnitude of adverse effects in case–control studies: 'The number of patients needed to be treated for one additional patient to be harmed'. *British Medical Journal* 2000; **320**:503–506.
3. Heller RF, Dobson AJ, Attia J, Page J. Impact numbers: measures of risk factor impact on the whole population from case–control and cohort studies. *Journal of Epidemiology and Community Health* 2002; **56**:606–610.
4. Bender R, Blettner M. Calculating the 'number needed to be exposed' with adjustment for confounding variables in epidemiological studies. *Journal of Clinical Epidemiology* 2002; **55**:525–530.
5. Bishop YMM, Fienberg SE, Holland PW. *Discrete Multivariate Analysis*: *Theory and Methods*. MIT Press: Cambridge, MA, 1975.
6. Bender R, Kuss O. Confidence intervals for adjusted NNEs: a simulation study (Letter). *Journal of Clinical Epidemiology* 2003; **56**:205–206.
7. Lesaffre E, Pledger G. A note on the number needed to treat. *Controlled Clinical Trials* 1999; **20**:439–447.
8. Thabane L. A closer look at the distribution of number needed to treat (NNT): a Bayesian approach. *Biostatistics* 2003; **4**:365–370.

DOI: 10.1002/sim

9. Bender R. Calculating confidence intervals for the number needed to treat. *Controlled Clinical Trials* 2001; **22**:102–110.
10. Altman DG. Confidence intervals for the number needed to treat. *British Medical Journal* 1998; **317**:1309–1312.
11. Kleinbaum DG, Klein M. *Logistic Regression*: *A Self-Learning Text* (2nd edn). Springer: New York, NY, 2002.
12. Sabatine MS, Cannon CP, Gibson CM, Lopez-Sendon JL, Montalescot G, Theroux P, Lewis BS, Murphy SA, McCabe CH, Braunwald E. Effect of clopidogrel pretreatment before percutaneous coronary intervention in patients with ST-elevation myocardial infarction treated with fibrinolytics: the PCI-CLARITY study. *Journal of the American Medical Association* 2005; **294**:1224–1232.
13. Walter SD. The estimation and interpretation of attributable risk in health research. *Biometrics* 1976; **32**:829–849.
14. Heller RF, Dobson A. Disease impact number and population impact number: population perspectives to measures of risk and benefit. *British Medical Journal* 2000; **321**:950–952.
15. Hildebrandt M, Bender R, Gehrmann U, Blettner M. Calculating confidence intervals for impact numbers. *BMC Medical Research Methodology* 2006; **6**:32.
16. Agresti A. *Categorical Data Analysis* (2nd edn). Wiley: New York, NY, 2002.
17. Hosmer Jr DW, Lemeshow S. *Applied Logistic Regression* (2nd edn). Wiley: New York, NY, 2000.
18. Blizzard L, Hosmer DW. Parameter estimation and goodness-of-fit in log binomial regression. *Biometrical Journal* 2006; **48**:5–22.

DOI: 10.1002/sim