The Stata Journal (2016)
16, Number 3, pp. 650–661

Using mi impute chained to fit ANCOVA models in randomized trials with censored dependent and independent variables

Andreas Andersen
Research Center for Vitamins and Vaccines
Statens Serum Institut
Copenhagen, Denmark
a.andersen@bandim.org

Andreas Rieckmann
Research Center for Vitamins and Vaccines
Statens Serum Institut
Copenhagen, Denmark
and
Odense Patient data Explorative Network
Institute of Clinical Research
University of Southern Denmark
Odense, Denmark
anri@ssi.dk

Abstract. In this article, we illustrate how to use `mi impute chained` with `intreg` to fit an analysis of covariance analysis of censored and nondetectable immunological concentrations measured in a randomized pretest–posttest design.

Keywords: st0447, mi impute chained, multiple imputation, intreg, censoring, detection limit, ANCOVA

1 Introduction

Antibody and cytokine concentrations are often measured in studies of the immune response to, for example, Bacillus Calmette–Guérin or measles vaccination (MV). The effect of such an exposure or a treatment on a continuous outcome variable can be evaluated by a randomized control-group design with a pretest or baseline measurement and a posttest or follow-up measurement (Jensen et al. 2014). Three common analyses of such designs are 1) analysis of the posttest measurement, 2) analysis of the posttest–pretest difference (change score), and 3) analysis of the posttest measurement adjusted for the pretest. The posttest analysis (POST) and the change score analysis (CHANGE) are conducted by analysis of variance. The third analysis is conducted by analysis of covariance (ANCOVA) with the posttest measurement as outcome and the pretest measurement as independent variable. A treatment indicator is included to estimate the treatment effect. In randomized trials, the difference at baseline between the treatment

and control groups is expected to be zero; therefore, POST, CHANGE, and ANCOVA are expected (in the statistical sense) to measure the same thing (Senn 2006). That is, the expected value of the treatment-effect estimates are the same. However, ANCOVA is the most efficient method with the highest statistical power to detect the treatment effect (Vickers 2001; Vickers and Altman 2001; Van Breukelen 2006).

Immunological measurement methods are often bounded by lower and upper detection limits (DL) creating nondetectable (ND) concentrations. In statistical terms, these observations can be regarded as left- and right-censored; the true concentration is only known to be smaller or greater than the lower or upper DL. Because the pretest and posttest measurements are measures of the same quantity, it is common that both variables have censored observations. Hence, in the ANCOVA analysis, we both have a censored outcome and a censored covariate. The issues of a left-censored outcome variable (Uh et al. 2008; Lubin et al. 2004) or a left-censored covariate (Nie et al. 2010; Austin and Hoch 2004) have previously been studied separately. Andersen et al. (2013) explored methods to conduct the CHANGE analysis with censored pretest and posttest measurements, showing that tobit regression and multiple imputation give very similar results. As of Stata 12, new extensive multiple-imputation commands are available, including the possibility to apply tobit regression in multivariate imputation methods using chained equations. This facilitates estimation of ANCOVA with censored pretest and posttest measurements. However, it is not obvious that this useful tool exists and how it should be applied. Our aim is to make users aware of this tool and to give them a practical guide and examples of how to apply it.

2 Methods

For independent subjects $i = 1, \ldots, n$, we denote the pretest measurement by X_i and the posttest measurement by Y_i. Let T_i be the treatment indicator. The POST, CHANGE, and ANCOVA analyses can be written as

$$\begin{aligned} Y_i &= \alpha_1 + \gamma_1 T_i + \epsilon_{1i}, & \epsilon_{1i} &\sim \text{i.i.d. } N(0, \sigma_1^2) \quad \text{(POST)} \\ Y_i - X_i &= \alpha_2 + \gamma_2 T_i + \epsilon_{2i}, & \epsilon_{2i} &\sim \text{i.i.d. } N(0, \sigma_2^2) \quad \text{(CHANGE)} \\ Y_i &= \alpha_3 + \gamma_3 T_i + \beta_3 X_i + \epsilon_{3i}, & \epsilon_{3i} &\sim \text{i.i.d. } N(0, \sigma_3^2) \quad \text{(ANCOVA)} \end{aligned} \tag{1}$$

ANCOVA with censored pretest and posttest measurements can be estimated by a multivariate imputation method using chained equations, also known as fully conditional specification (van Buuren, Boshuizen, and Knook 1999) or sequential regression multivariate imputation (Raghunathan et al. 2001). The Stata command is `mi impute chained` (see [MI] **mi impute chained**). The pretest and posttest variables are imputed iteratively through a sequence of univariate imputation models with fully conditional specification of the prediction equation. In iteration $(t + 1)$, $X^{(t)}$ is included in the prediction equation for $Y^{(t+1)}$, and $Y^{(t+1)}$ is included in the prediction equation for $X^{(t+1)}$. Each univariate imputation model is fit by tobit regression, and the draws are made from the estimated distribution truncated at the DL ([MI] **mi impute intreg**). Formally, the imputations are drawn from the prediction equations

$$Y^{(t+1)} \sim g_y(Y|X^{(t)}, T, \phi_y) \quad (2)$$
$$X^{(t+1)} \sim g_x(X|Y^{(t+1)}, T, \phi_x) \quad (3)$$

where ϕ_x and ϕ_y are the model parameters with uniform prior distributions. The univariate imputation models $g_x(\cdot)$ and $g_y(\cdot)$ are normal-based tobit models. After the censored observations have been imputed, the final treatment-effect estimate is obtained by `mi estimate` (see [MI] **mi estimate**).

3 ANCOVA: mi impute chained using intreg

To fit ANCOVA, we can use `mi impute chained` (see [MI] **mi impute chained**) with `intreg` (see [R] **intreg**) as the univariate imputation method.

3.1 Syntax

`mi impute chained` *lhs* [`=` *indepvars*] [*if*] [*weight*] [`,` *impute_options options*]

3.2 Description

Before we can use the `mi` commands, we must `mi set` the dataset. All the `mi set` styles can be used.

In `mi impute chained`, the left-hand side, *lhs*, specifies the prediction (2) and (3) used to impute values of the censored variables. Each censored variable requires specification of the univariate imputation method (*uvmethod*) and the name of the variable containing the imputations (*newivar*). With two censored variables, the syntax for *lhs* is `(`*uvmethod1*`)` *newivar1* `(`*uvmethod2*`)` *newivar2*. The *uvmethod* for the censored pretest and posttest variables is the tobit model implemented in the `intreg` command for interval-censored regression. The censoring interval is specified by the two dependent variables, *dep1var* and *dep2var*. For data right-censored at an upper DL, the interval for the censored observation is [DL,∞), and the dependent variables should be given the values *dep1var* = DL and *dep2var* = . as described in [R] **intreg**. For data left-censored at a lower DL, the interval for the censored observation is (-∞,DL], and the dependent variables should be given the values *dep1var* = . and *dep2var* = DL. For an uncensored observation with point value *a*, the dependent variables should be set to *dep1var* = *a* and *dep2var* = *a*. Because `intreg` is used with `mi impute chained`, the censoring variables must be specified as options: `intreg`, `ll(`*dep1var*`)`, and `ul(`*dep2var*`)`. The optional right-hand side [`=` *indepvars*] specifies complete, uncensored variables (that is, `Z1` and `Z2`) that may be included in the prediction equation for the censored variables. Overall, the syntax is `mi impute chained (intreg, ll(`*dep1var1*`) ul(`*dep2var1*`))` *newivar1* `(intreg, ll(`*dep1var2*`) ul(`*dep2var2*`))` *newivar2* `= Z1 Z2`.

3.3 Options

The options `add()`, `burnin()`, and `chainonly()` for `mi impute chained` are used to control the number of imputations to be drawn and the number of burn-in runs between each imputation.

4 Example: Immunological study of MV

In Jensen et al. (2014), we conducted an immunological substudy within a randomized trial of early MV at 4.5 and 9 months of age compared with the existing schedule giving MV at 9 months of age. We obtained blood samples before randomization (pretest) and six weeks after randomization (posttest) and measured concentrations of cytokines and soluble receptors. In the present example, we consider the concentration of IL-10 in response to in vitro challenge with purified protein derivative. The manufacturer of the cytokine assay defined a lower DL of 5pg/ml for IL-10. In the control and randomization groups, 23% and 19%, respectively, of the pretest IL-10 concentrations were ND below the DL, while 25% and 26%, respectively, of the posttest concentrations were below the DL. Distributions of immunological markers often resemble lognormal distributions. Figure 1 shows histograms and normal densities of the log concentrations of IL-10 in the pretest and posttest samples and in the two randomization groups. In the histogram, ND concentrations are placed in the DL. The mean and variance parameters of the normal density functions are estimated by tobit regression. Arguably, the lognormal distribution provides a reasonable fit to the observed distributions.

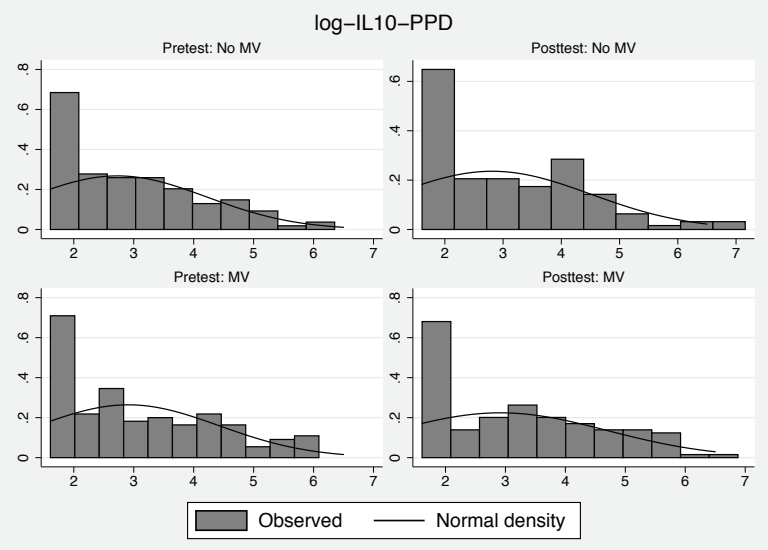

Figure 1. Histograms of pretest and posttest IL-10 concentrations

In the following, we show how we estimated the intervention effect of early MV in a pretest–posttest ANCOVA analysis using multiple imputation to account for the censoring below the DL. Because there can be substantial variation in cytokine concentrations across assay plates, we adjusted the analyses for assay plates in the variable called `sele`. In Jensen et al. (2014), we presented geometric mean ratios antilogging the regression coefficients. For convenience, we here present the direct log-scale coefficients.

Before we get to the ANCOVA analysis, we first conduct the POST analysis by tobit regression using `intreg`. The posttest variable is called `il10_ppd2`. We first generate the dependent variables to be used for `intreg`. The left-censored observations (`il10_ppd2` $\leq$ 5) are specified by a missing value of the first dependent variable `dp1v2`. The POST analysis gives an MV effect of 0.16 (95% confidence interval (CI): $[-0.29; 0.60]$).

```
. use mv_il10ppd
. generate dp1v2=log(il10_ppd2) if il10_ppd2>5
(63 missing values generated)
. generate dp2v2=log(il10_ppd2)
. quietly intreg dp1v2 dp2v2 i.mv i.sele
. intreg, cformat(%5.2f)
Interval regression                             Number of obs     =        249
                                                LR chi2(2)        =       3.78
Log likelihood = -426.64584                     Prob > chi2       =     0.1514
```

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
mv						
early mv	0.16	0.23	0.70	0.485	-0.29	0.60
2.sele	0.43	0.23	1.87	0.061	-0.02	0.88
_cons	2.60	0.20	12.95	0.000	2.21	3.00
/lnsigma	0.55	0.06	9.90	0.000	0.44	0.65
sigma	1.73	0.10			1.55	1.92

```
     63  left-censored observations
    186     uncensored observations
      0 right-censored observations
      0       interval observations
```

However, the pretest measurements (in the variable `il10_ppd1`) are not balanced between the two randomization groups. In fact, the mean difference of 0.20 (95% CI: $[-0.18; 0.58]$) is larger in the pretest samples than in the posttest samples.

```
. generate dp1v1=log(il10_ppd1) if il10_ppd1>5
(51 missing values generated)
. generate dp2v1=log(il10_ppd1)
. quietly intreg dp1v1 dp2v1 i.mv i.sele
. intreg, cformat(%5.2f)
Interval regression                             Number of obs     =        249
                                                LR chi2(2)        =       1.65
Log likelihood = -412.32184                     Prob > chi2       =     0.4384
```

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
mv						
early mv	0.20	0.20	1.02	0.308	-0.18	0.58
2.sele	0.17	0.20	0.88	0.381	-0.21	0.56
_cons	2.65	0.17	15.34	0.000	2.31	2.98
/lnsigma	0.40	0.05	7.66	0.000	0.30	0.51
sigma	1.50	0.08			1.35	1.66

```
     51  left-censored observations
    198     uncensored observations
      0 right-censored observations
      0       interval observations
```

Thus we observe a (nonsignificant) difference in posttest concentrations between the two groups, but this could be caused by the baseline imbalance in the pretest concentrations before randomization. ANCOVA (1) adjusts for the pretest imbalance by subtracting from the posttest difference the pretest difference scaled by the pretest–posttest correlation.

$$\widehat{\gamma_3} = \left(\overline{Y}_1 - \overline{Y}_0\right) - \widehat{\beta_3}\left(\overline{X}_1 - \overline{X}_0\right)$$

We have seen that there are ND concentrations among both the pretest and the posttest samples. The ANCOVA model can be fit by the sequential multiple imputation of the ND pretest and posttest concentrations using `mi impute chained` with `intreg` as the univariate imputation method. The ANCOVA analysis gives an MV effect of 0.10 (95% CI: $[-0.33; 0.54]$).

```
. preserve
. set seed 1000
. mi set mlong
. mi impute chained (intreg, ll(dp1v1) ul(dp2v1)) x1imp (intreg, ll(dp1v2)
> ul(dp2v2)) x2imp=i.mv i.sele, add(20) burnin(5)
Conditional models:
             x1imp: intreg x1imp x2imp i.mv i.sele , ll(dp1v1) ul(dp2v1)
             x2imp: intreg x2imp x1imp i.mv i.sele , ll(dp1v2) ul(dp2v2)
Performing chained iterations ...
Multivariate imputation                     Imputations =       20
Chained equations                                 added =       20
Imputed: m=1 through m=20                       updated =        0
Initialization: monotone                     Iterations =      100
                                                burn-in =        5
             x1imp: interval regression
             x2imp: interval regression
```

	Observations per *m*			
Variable	Complete	Incomplete	Imputed	Total
x1imp	198	51	51	249
x2imp	186	63	63	249

```
(complete + incomplete = total; imputed is the minimum across m
 of the number of filled-in observations.)
. mi estimate, nimputations(20) cformat(%5.2f) noheader: regress x2imp i.mv
> c.x1imp i.sele
```

x2imp	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
mv						
early mv	0.10	0.22	0.47	0.640	-0.33	0.54
x1imp	0.25	0.08	3.31	0.001	0.10	0.40
2.sele	0.39	0.23	1.71	0.089	-0.06	0.84
_cons	1.94	0.29	6.76	0.000	1.37	2.51

In figure 2, we show Q–Q plots of the completed pretest and posttest distributions after the first iteration of the imputation algorithm. The plots illustrate a good fit to the normal distribution, except for a few very low imputed values of the posttest measurements.

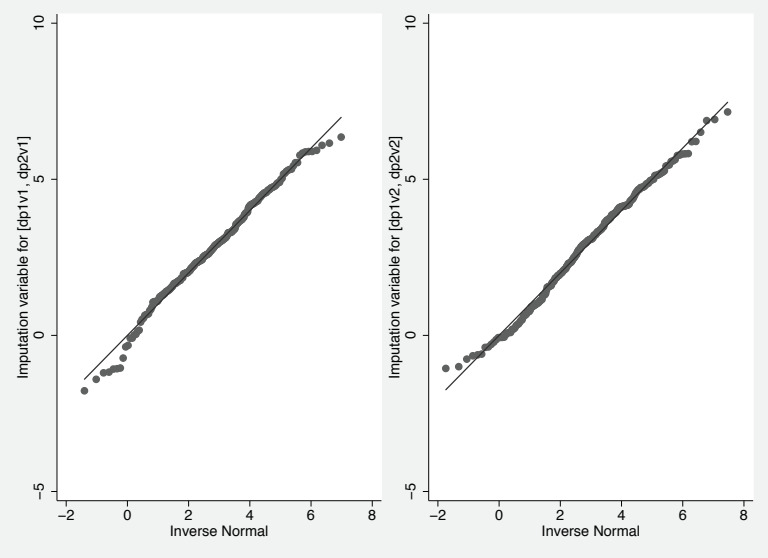

Figure 2. Q–Q plots of the completed values of the pretest and posttest measurements from one iteration

To assess convergence of the algorithm, we use `chainonly` and `savetrace()` to perform chained iterations without generating imputations and save the means of the imputed values in `impstats.dta`. The means are plotted along the iterations in figure 3. The two chains reach a steady level basically from the first iteration.

```
. restore
. set seed 1000
. mi set mlong
. mi impute chained (intreg, ll(dp1v1) ul(dp2v1)) x1imp (intreg, ll(dp1v2)
> ul(dp2v2)) x2imp=i.mv i.sele, burnin(50) chainonly chaindots
> savetrace(impstats, replace)
Conditional models:
             x1imp: intreg x1imp x2imp i.mv i.sele , ll(dp1v1) ul(dp2v1)
             x2imp: intreg x2imp x1imp i.mv i.sele , ll(dp1v2) ul(dp2v2)
Performing chained iterations:
  burn-in 50 .........10.........20.........30.........40.........50 done
Note: No imputation performed.
. use impstats, clear
(Summaries of imputed values from -mi impute chained-)
. scatter x1imp_mean x2imp_mean iter, c(l l) msymbol(i i) lcolor(black gs7)
```

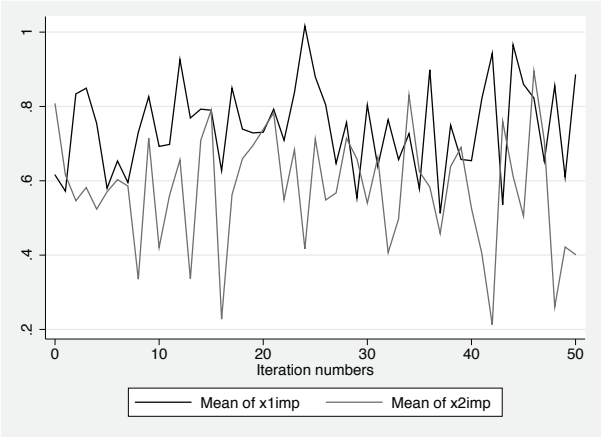

Figure 3. The means of the imputed pretest concentrations (`x1imp`) and the imputed posttest concentrations (`x2imp`) are presented to evaluate the convergence of the chained iterations

5 Simulation study of bias and power

We conducted a simulation study of the methods to fit POST, CHANGE, and ANCOVA. We compare the methods on bias (relative to the true effect) and power. Observations X, Y are drawn from a bivariate normal distribution with mean $(2.5, 3)$ in the control group and mean $(2.5, 3.15)$ in the treatment group. Hence, we simulate a pretest difference of 0 and a posttest difference of 0.15. The variances are $\sigma_x = \sigma_y = 1$ for both groups. We generated 5,000 datasets with $N = 400$ observations drawn from distributions with pretest–posttest correlations $\rho = 0.25$, 0.50, 0.75. The pretest and posttest measurements are artificially censored at 20%, 40%, and 60%.

Figure 4 shows that all three methods estimate the true effect of 0.15 unbiased for all pretest and posttest censoring schemes. Figure 5 shows power to detect the treatment effect of the three methods. As expected, in general, ANCOVA has the highest power. With low pretest–posttest correlation ($\rho = 0.25$), POST has similar power to ANCOVA, while CHANGE has lower power. With high pretest–posttest correlation ($\rho = 0.75$), POST has considerably lower power than ANCOVA, while CHANGE is closer to ANCOVA. Increasing censoring of the posttest outcome does not seem to decrease power. In contrast, power decreases with increasing pretest censoring. However, it is still more powerful to conduct ANCOVA with a censored pretest than to disregard the pretest in the POST analysis.

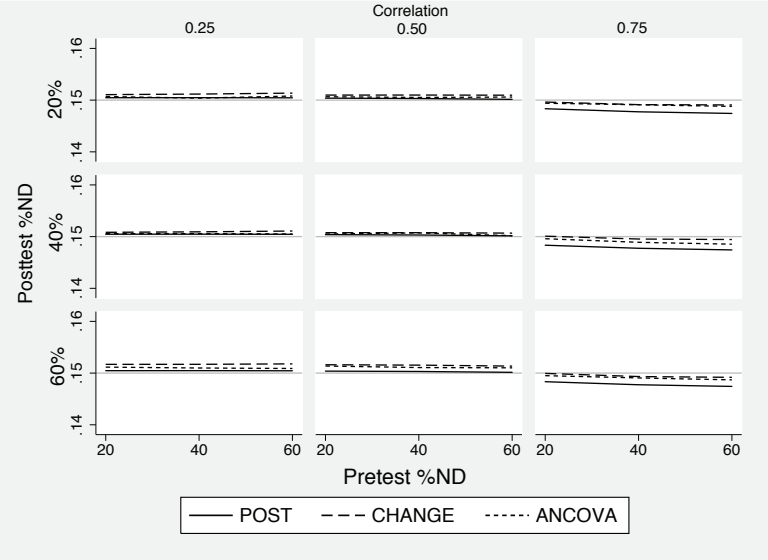

Figure 4. The average treatment-effect estimate in the simulation studies with various pretest–posttest correlations and various censoring schemes. The true underlying effect is 0.15

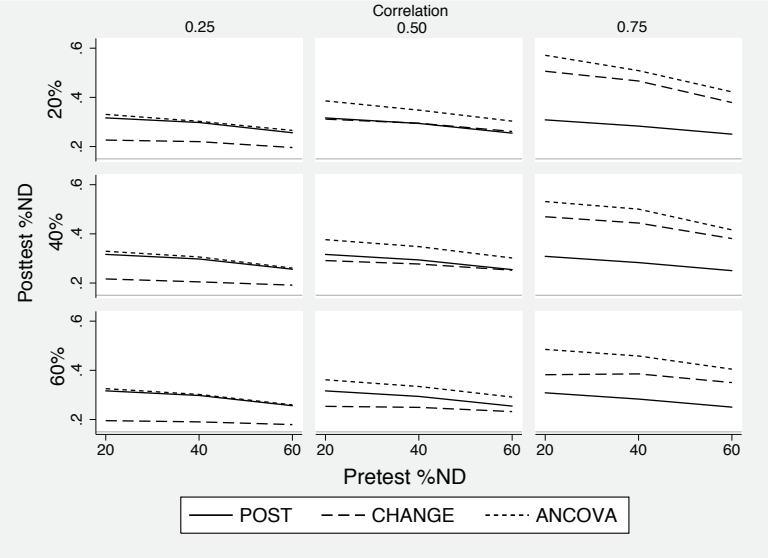

Figure 5. The average power to detect a significant treatment effect in the simulation studies with various pretest–posttest correlations and various censoring schemes

6 Conclusion

The command `mi impute chained` can be used to fit ANCOVA analyses of censored normally distributed data from randomized pretest–posttest trials and performs with higher power than POST and CHANGE analyses.

7 References

Andersen, A., C. S. Benn, M. J. Jørgensen, and H. Ravn. 2013. Censored correlated cytokine concentrations: Multivariate Tobit regression using clustered variance estimation. *Statistics in Medicine* 32: 2859–2874.

Austin, P. C., and J. S. Hoch. 2004. Estimating linear regression models in the presence of a censored independent variable. *Statistics in Medicine* 23: 411–429.

Jensen, K. J., M. Søndergaard, A. Andersen, E. Sartono, C. Martins, M.-L. Garly, J. Eugen-Olsen, H. Ullum, M. Yazdanbakhsh, P. Aaby, C. S. Benn, and C. Erikstrup. 2014. A randomized trial of an early measles vaccine at 4.5 months of age in Guinea-Bissau: Sex-differential immunological effects. *PLOS ONE* 9: e97536.

Lubin, J. H., J. S. Colt, D. Camann, S. Davis, J. R. Cerhan, R. K. Severson, L. Bernstein, and P. Hartge. 2004. Epidemiologic evaluation of measurement data in the presence of detection limits. *Environmental Health Perspectives* 112: 1691–1696.

Nie, L., H. Chu, C. Liu, S. R. Cole, A. Vexler, and E. F. Schisterman. 2010. Linear regression with an independent variable subject to a detection limit. *Epidemiology* 21: S17–S24.

Raghunathan, T. E., J. M. Lepkowski, J. Van Hoewyk, and P. Solenberger. 2001. A multivariate technique for multiply imputing missing values using a sequence of regression models. *Survey Methodology* 27: 85–95.

Senn, S. 2006. Change from baseline and analysis of covariance revisited. *Statistics in Medicine* 25: 4334–4344.

Uh, H.-W., F. C. Hartgers, M. Yazdanbakhsh, and J. J. Houwing-Duistermaat. 2008. Evaluation of regression methods when immunological measurements are constrained by detection limits. *BMC Immunology* 9: 59.

Van Breukelen, G. J. P. 2006. ANCOVA versus change from baseline had more power in randomized studies and more bias in nonrandomized studies. *Journal of Clinical Epidemiology* 59: 920–925.

van Buuren, S., H. C. Boshuizen, and D. L. Knook. 1999. Multiple imputation of missing blood pressure covariates in survival analysis. *Statistics in Medicine* 18: 681–694.

Vickers, A. J. 2001. The use of percentage change from baseline as an outcome in a controlled trial is statistically inefficient: A simulation study. *BMC Medical Research Methodology* 1: 6.

Vickers, A. J., and D. G. Altman. 2001. Analysing controlled trials with baseline and follow up measurements. *British Medical Journal* 323: 1123–1124.

About the authors

Andreas Andersen is a researcher and statistician interested in censored observations due to DL in immunological studies and missing or misclassified data in vaccine studies.

Andreas Rieckman is a PhD student interested in methodological challenges in epidemiology.