ORIGINAL ARTICLE

Imputing missing data of function and disease activity in rheumatoid arthritis registers: what is the best technique?

Denis Mongin [iD],[1] Kim Lauper [iD],[1] Carl Turesson,[2,3] Merete Lund Hetland,[4,5] Eirik Klami Kristianslund,[6] Tore K Kvien,[6] Maria Jose Santos [iD],[7] Karel Pavelka,[8] Florenzo Iannone,[9] Axel Finckh,[1] Delphine Sophie Courvoisier[1]

Check for updates

ABSTRACT

Objective To compare several methods of missing data imputation for function (Health Assessment Questionnaire) and for disease activity (Disease Activity Score-28 and Clinical Disease Activity Index) in rheumatoid arthritis (RA) patients.

Methods One thousand RA patients from observational cohort studies with complete data for function and disease activity at baseline, 6, 12 and 24 months were selected to conduct a simulation study. Values were deleted at random or following a predicted attrition bias. Three types of imputation were performed: (1) methods imputing forward in time (last observation carried forward; linear forward extrapolation); (2) methods considering data both forward and backward in time (nearest available observation—NAO; linear extrapolation; polynomial extrapolation); and (3) methods using multi-individual models (linear mixed effects cubic regression—LME3; multiple imputation by chained equation—MICE). The performance of each estimation method was assessed using the difference between the mean outcome value, the remission and low disease activity rates after imputation of the missing values and the true value.

Results When imputing missing baseline values, all methods underestimated equally the true value, but LME3 and MICE correctly estimated remission and low disease activity rates. When imputing missing follow-up values at 6, 12, or 24 months, NAO provided the least biassed estimate of the mean disease activity and corresponding remission rate. These results were not affected by the presence of attrition bias.

Conclusion When imputing function and disease activity in large registers of active RA patients, researchers can consider the use of a simple method such as NAO for missing follow-up data, and the use of mixed-effects regression or multiple imputation for baseline data.

INTRODUCTION

Large observational studies become more common for the assessment of new treatments[1] and are encouraged by national or international health agencies.[2] Rheumatoid arthritis (RA) studies follows these trends and disease registers have been developed,[3–8]

Key messages

What is already known about this subject?

► Missing data is a major problem in large register of observational data. The use of complete case analysis is known to be a potential source of bias, and imputation of missing data is recommended.

What does this study add?

► The present work provides the first systematic study comparing the consequences of a wide range of strategies for dealing with missing values of function or disease activity in large registers of patients with rheumatoid arthritis (RA).

► Imputation methods considering data both forward and backward in time, such as using the nearest available observation, give a mean value and an associated remission or low disease activity rates almost unbiased for missing follow-up.

How might this impact on clinical practice or future developments?

► The results presented will help researchers to appropriately choose imputation methods when dealing with missing function or disease activity data in large register of observational data. This will ease studies exploring the effectiveness and safety of RA treatments in real-world populations, and thus indirectly impact clinical practices.

allowing to explore the effectiveness and safety of RA treatments in real-world populations. One of the key estimates of effectiveness is the calculation at a given follow-up time of the disease activity and the corresponding remission rate. However, observational register studies often face large amount of missing data, due to a variety of reasons, such as attrition of patients over time or due to missing visits at specific follow-up periods. Missing data of a variable Y can be of three type: (1) missing completely at random (MCAR) if the probability of missing is unrelated to any observed or unobserved variable, (2) missing

at random (MAR) if the probability of missing depends on observed variables but not on missing values of Y, and (3) missing not at random (MNAR) if the probability of missing depends on missing values of Y.[9 10] To avoid reducing the statistical power and to prevent the potential bias that complete case analyses (CCA) would cause in case of MAR or MNAR data,[11 12] the use of imputation techniques to handle missing data is advised. Imputation methods available for longitudinal data are numerous. Some impute forward in time, such as last observation carried forward (LOCF) or linear forward extrapolation (LFE). Others consider data both forward and backward in time, for instance nearest available observation (NAO),[13] linear extrapolation (LE), and polynomial extrapolation (PE), where the missing data is obtained by the adjustment of the variable time dependence with a third order polynomial. More advanced methods use all available data from all patients. Mixed effects regression, for example, can be used as an imputation method by estimating the mean and the individual time evolution of disease activity to extrapolate missing data.[14] Finally, the well-known multiple imputation using chained equation (MICE), uses data from other variables instead of using only the variable Y at different time points.[15 16] Despite the existence of this panel of techniques, an informal review of the articles using large registers published in two major rheumatology journals (*Annals of the Rheumatic Diseases* and *Arthritis & Rheumatology*) over the last 2 years demonstrated that almost 50% of them do not describe the strategy used to handle missing data, 25% use CCA and 5% LOCF. This observation, which is not specific to rheumatology,[17] is partly due to the absence of guidelines and the lack of studies tackling the problem of data imputation in large registers and its impact on outcomes estimates. Although imputation has been studied in clinical trials of rheumatic diseases[11 18–22] and in small registers,[23] the focus was generally limited to only a few imputation techniques and applied to a panel containing maximum four data points per patient. The aim of this work is thus to compare the effects of the main imputation techniques for MCAR or MAR data and the resulting impact on the estimation of mean function and disease activity and on the derived remission and low disease activity (LDA) rates in longitudinal observational data of RA patients. We will focus on the estimation of mean function and disease activity and remission rate at baseline and three specific but common follow-up: 6, 12 and 24 months.

METHODS
We used RA patient data from six national registries (Czech Republic, Denmark, Italy, Norway, Portugal and Sweden) of an existing register collaboration of RA registers (the Pan-European analysis of real-world effectiveness of abatacept (PanABA)).[8 24] Each register has obtained ethical approval for the use of anonymous data for research in their local ethics committee. Inclusion criteria were diagnosis of RA established by a rheumatologist, >18 years of age and initiation of a treatment with abatacept or tumor necrosis factor-inhibitors. Baseline (0 month) was defined as the initiation of a bDMARD treatment, and follow-up as the time elapsed since treatment initiation. The three outcomes of interest were the Health Assessment Questionnaire (HAQ) and the disease activity score based on 28 joints (DAS28) at four time points of interest: baseline, 6, 12 and 24 months. Because in real world observational data, patients' visits are rarely exactly at the specified time point, function and disease activity scores were taken within a time window from 0 to +1 month for baseline or treatment initiation, and within a 3-month time window centred at the time of interest for follow-up visits (6, 12, or 24 months).

Creation of the two datasets for the simulation study
First, a random sample of 1000 patients were selected from a pool of patients with complete data for the three outcomes at 0, 6, 12 and 24 months (DS1). The patients selected in for this simulation study had more visits on average (median 20, IQR 8–25) than patients from the registers not included in this study (median 6, IQR 2–7). In order to study the behaviour of imputation methods on a dataset representative of patients a commonly followed in registers, we created a second dataset DS2 with a number of visits per patient similar to the average patient in the registers. To this end, we randomly selected for each patient of DS1 a number of visits following the probability density function of the overall dataset of 49 319 patient stemming from 13 different countries (online supplementary figure S1). This procedure ensures that the DS2 dataset contained patients with similar visit numbers as the entire register collection. Main simulations were performed on DS2, while sensitivity analysis on availability of data near follow-up were performed on DS1 (see Sensitivity analysis section).

Creating missing data
For each follow-up time of interest, missing data in the two datasets were created by deleting all data in a given time window centred on the time of imputation for a number of patients corresponding to a predetermined percentage of missing data (eg, 60% missing data at 6 months, corresponds to 600 patients having no data available in a time windows centred on 6 months). The percentage of missing data at each time of imputation was varied from 10% to 90%. To simulate MCAR missing data, patients having missing data were chosen randomly, independently of any other variable. To simulate MAR missing data, patients with missing data were chosen by sampling them with a probability given by the odds of having a missing value for the considered outcome and follow-up time. For each outcome and follow-up, these odds were extracted from a generalised linear model applied on the entire register estimating the probability of having a missing outcome as a function of the treatment duration, the number of previous biological treatments and the baseline value of disease activity.

Sensitivity analysis

The length of time windows for the deletion of all observations used to create missing data was set to 3 months. As imputation performance may depend on the proximity of available data near follow-up, we varied the length of the time windows for deletions from 3 to 6 months, 12 months (only for 12 and 24 months follow-up to avoid deleting baseline data) and 18 months (for 24 months follow-up) when creating missing data in DS1.

Imputation methods

Schematic of the imputation methods is presented in online supplementary figure S2.

Last observation carried forward

LOCF sets the missing data to the last available value for each patient.

Linear forward extrapolation

LFE assumes a linear trend between the missing data and the last two available data. The predicted value $Y(t)$ of variable Y at follow-up t is:

$$Y(t) = \frac{Y(t_{-1}) - Y(t_{-2})}{t_{-1} - t_{-2}} (t - t_{-1}) + Y(t_{-1})$$

where t_{-1} is the last time with available data for Y and t_{-2} the second last one.

Nearest available observation

NAO sets the missing data as equal to the closest available data in time (backward or forward) for each patient.

Linear extrapolation

LE assumes a linear trend between the missing data and the two closest available data in time. Mathematical expression is the same as LFE with t_{-1} the closest time with available data for Y and t_{-2} the second closest.

Polynomial extrapolation

PE adjusts the individual variable Y with a K order time polynomial (for N measures, $K = N - 1$ and $K < 4$), and predict the Y missing value by calculating its value at follow-up t:

$$Y(t) = a + b1*t + b2*t^2 + b3*t^3$$

with a, b1, b2 and b3 the polynomial coefficients estimated using ordinary least square regression.

Linear mixed effects cubic regression

LME3 is a linear mixed effects model adjusting the variable to be imputed by a third order polynomial of time using all patients together. Considering Y_i the variable vector with the missing data and t_i the time vector associated for patient i, the regression reads:

$$Y_i \sim b_0 + u_{0i} + (b_1 + u_{1i}) t_i + (b_2 + u_{2i}) t_i^2 + (b_3 + u_{3i}) t_i^3$$

with b_n the fixed effect coefficients and u_{ni} their random counterpart. The missing data at follow-up t is predicted by directly estimating Y for each patient at t from the regression result.[25 26]

Multiple imputation by chained equation

Multiple imputation is a widely used technique of imputation developed by Rubin[27 28] based on a Monte Carlo technique in which the dataset with missing value is replaced by $m > 1$ complete dataset. We used here MICE[16] with the predictive mean matching technique implemented in the mice library in R with $n_{it} = 10$ iterations. Due to the high computational power required for our simulation, we chose m so that the SD of the imputation is only 1% wider than the one obtained with an infinite number of imputation[28]:

$$m = \frac{RE\lambda}{1 - RE} \approx 100\lambda$$

with λ being the rate of missing information, and RE the relative efficiency here set to 1%. It resulted in m varying between 22 and 35 depending on the parameters. In this simulation study, the variables used for the imputation were time elapsed since treatment initiation, age, sex, patient identifier, number of previous biological DMARD, type of concomitant synthetic DMARD, disease duration, seropositivity and when imputing at time different than baseline, the baseline values of DAS28 and HAQ.

As sensitivity analysis, other MICE methods were tried for missing values set at 60%: Bayesian linear regression, random sample from observed values, random forest imputations. Influence of iteration and imputation number has been tested by increasing both the imputation number ($m = 50$) and the iteration number ($n_{it} = 50$).

Complete case analysis

We also presented results based on CCA, for example, restricting results to follow-up with no missing values.

Aberrant values

When imputing with LE, LFE or PE and if the imputed value was negative or above maximal possible value, it was set to 0 or the maximum value respectively. Following recommendations,[29] imputed value of MICE were not changed.

Simulation

For each condition of missing data, follow-up, percentage of missing data at follow-up and size of the deletion time windows, the values of the outcomes were deleted and imputed with each imputation method 1000 times. Simulation was performed on the high performance computing facility of the Geneva University.

Statistical methods

For each parameter condition, the mean value and the relative bias for the DAS28 and HAQ were calculated and averaged over the 1000 runs of the simulation. The relative bias for a value α is

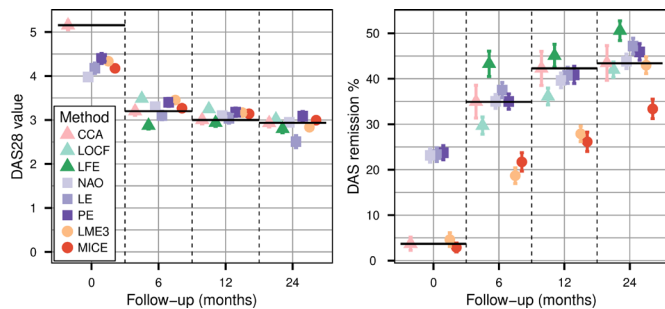$$\frac{\Delta\alpha}{\alpha} = \frac{\alpha_{imputed} - \alpha_{true}}{\alpha_{true}}$$

Figure 1 Mean disease activity score based on 28 joints (left panel) and corresponding remission rate (right panel) of the data panel where the 60% of data missing completely at random at follow-up were handled with complete case analysis (CCA), last observation carried forward (LOCF), linear forward extrapolation (LFE), nearest available observation (NAO), linear extrapolation (LE), polynomial extrapolation (PE), linear mixed effect cubic regression (LME3) and multiple imputation by chained equation (MICE). Error bars are the standard errors calculated over the 1000 simulation samples. LFE and LOCF estimation are missing for t=0 since these methods use previous values, which do not exist at baseline. Black horizontal lines represent the true values for each follow-up.

In addition, for DAS28, the percentage of remission and of LDA was also computed to better assess the potential clinical impact of the different imputation techniques.

R 3.6.1[30] software with the libraries *lme4, mice, parallel, ggplot2* and *data.table* were used to conduct the study. The code is provided as a supplementary file and can be used to implement these imputation techniques.

RESULTS
Data MCAR
Mean imputed values

For missing baseline values, all imputation methods underestimated the true mean value of DAS28 (figure 1, left panel and online supplementary table S1) and HAQ. PE and LME3 estimated the best results with around −20% bias. LOCF and LE cannot be used at baseline since no prior data are available. For missing 6, 12 and 24 months values, LOCF consistently overestimated the true values, while LFE underestimated them, the absolute bias decreasing for longer follow-ups. Among the methods using data points before and after the missing values, NAO was the most accurate. LME3 overestimated the true values, with biases below 10% in absolute value at 12 and 24 months. MICE also displayed small biases, consistently below 10%.

Compared with imputed missing DAS28, imputed HAQ values presented similar relative bias (online supplementary table S2). Varying the proportion of missingness between 10% and 90% at follow-up didn't affect the bias of the imputed values. It resulted in a linear relation between the bias of the mean outcome calculated over the whole dataset and the percentage of missing value (online supplementary figure S3), since an increase of

missing value increments the amount of biased imputed value entering the calculation of the mean value.

These results can be explained by the mean temporal shape of the main outcomes (online supplementary figure S4 for DAS28). With appropriate treatment, these outcomes begin by improving sharply in the first few months after treatment initiation, then they tend to level off until the outcomes reach a plateau.[31–35] At earlier time points, LOCF tends to overestimates the imputed values, as values before the follow-up tend to be higher than values at later follow-up times. The convex shape of this decrease also explains the underestimation of the LFE imputation, as the outcome deviates from a local linear approximation (see online supplementary figure S5). PE and LME3 probably perform best for missing baseline values because they fit more adequately the initial steeper change compared with the other models. Conversely, values imputed by LME3 and PE at 6 and 12 months are overestimated because of the influence of the higher baseline values. Imputed values at 24 months yielded low relative bias for almost all imputation methods because the outcome reaches a plateau.

Remission and LDA rate

When imputing rates of remission (figure 1, right panel and online supplementary table S3) or LDA derived from disease activity assessments, a large difference in results is observed between the various imputation methods. At baseline, LME3 and MICE yield remission rates close to the true rates, while the three other methods yield remission rates overestimated by three times. As LME3 and MICE imputed values are narrowly distributed around their estimated mean (figure 2, left panel), this does not lead to the overrating of small values produced by the other methods and so to the overestimation of remission rate.

For missing follow-up disease activity data, the narrow distribution of LME3 and MICE imputed value creates a strong underestimation of the true remission rate (figure 2, right panel). The NAO, LE and PE methods on the other hand provide an almost unbiased rate of remission (figure 1, right panel). NAO tends to perform best, because the distribution of NAO imputed data is comparable to the true one, whereas PE and LE tend to change the imputed value distribution because of their tendency to impute negative values (figure 2, right panel). LOCF and LFE yield respectively an underestimated and an overestimated remission rate with similar absolute difference (figure 1), in line with their observed tendency to respectively overproduce high and LDA values (figure 2).

The narrowing of the distribution of LME3 imputed values find its root in the mixed effect part of the regression, because it supposes that the polynomial coefficients are normally distributed across the patients, thus artificially reducing the range of allowed values. This is confirmed by the fact that PE, which is similar to LME3 but considers each patient individually, does not show this artefact (figure 2). The results for LDA yield results
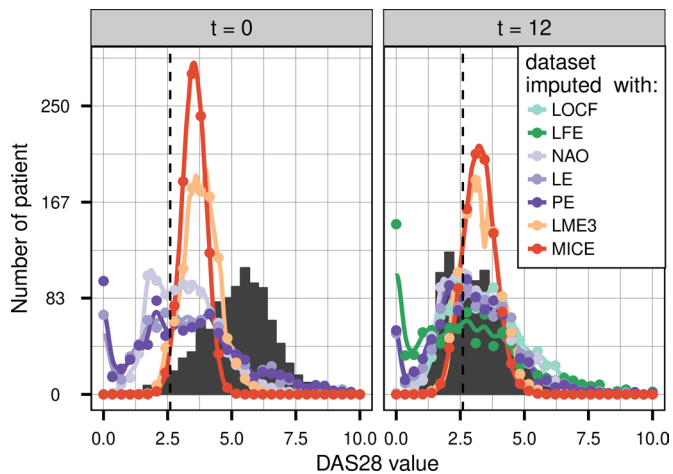
Figure 2 Distribution of disease activity score based on 28 joints values imputed with last observation carried forward (LOCF), linear forward extrapolation (LFE), nearest available observation (NAO), linear extrapolation (LE), polynomial extrapolation (PE), linear mixed effect cubic regression (LME3) and multiple imputation by chained equation (MICE), for two different follow-up (t=0 and t=12 months), 1000 imputations of 60% of data missing completely at random at follow-up. Black plain area is the true distribution, and colour lines are interpolations of the imputed dataset's distribution. The vertical line corresponds to the threshold of remission rate.

qualitatively similar to the remission simulations (see online supplementary figure S6 and online supplementary table S3), confirming our previous conclusions for remission rates.

Data MAR
Overall, results were generally similar when missing data patterns depended on covariates. There are two main differences. MICE, which requires MAR data and not MCAR, performs better, but still underestimates true remission rates by at least 10% at any follow-up visits (figure 3). The CCA applied on MAR data leads to biassed estimation of the remission rate, the bias reaching 12% at 2-year follow-up. Imputed mean value, remission rate or LDA by other imputation methods are not or only slightly affected by attrition bias, even though they do not explicitly take into account the covariates. This is certainly due to the fact that these methods use previous or subsequent measures of the outcome, which depend in part on the covariates, and can thus serve as proxy measures of the covariates influencing missing data.

Sensitivity analysis
In order to test how these results are impacted by the proximity of available data near follow-up, we repeated the analysis on DS1 with varying ranges of the deletion time windows when creating missing data. The results are presented for MAR in online supplementary figure S7 for the mean value, and online supplementary figure S8 for the remission rate. This sensibility study demonstrates that LOCF and LFE are most affected by the size of the deletion time window around the visits. When increasing
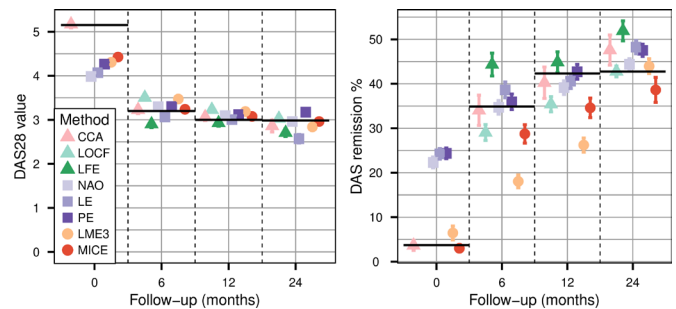

Figure 3 Mean disease activity score based on 28 joints (left panel) and corresponding remission rate (right panel) of the data panel where the 60% of data missing at random at follow-up were handled with complete case analysis (CCA), last observation carried forward (LOCF), linear forward extrapolation (LFE), nearest available observation (NAO), linear extrapolation (LE), polynomial extrapolation (PE), linear mixed effect cubic regression (LME3) and multiple imputation by chained equation (MICE). Error bars are the standard errors calculated over the 1000 simulation samples. LFE and LOCF estimation are missing for t=0 since these methods use previous values, which do not exist at baseline. Black horizontal lines represent the true values for each follow-up.

the deletion time windows, the bias of imputed values increases, since the available time points are further away in time (online supplementary figure S5).

In the sensitivity analysis of MICE parameters, the change in imputation method, number of imputed data sets or iteration number did not affect qualitatively the results presented.

A summary of the different results is presented in table 1.

DISCUSSION
This simulation study based on real world data allows to compare the consequences of a wide range of strategies when dealing with missing values for function or disease activity in large registers. However, many methods have been proposed and this simulation study does not test all of them.[25 36] The use of CCA can lead to bias greater than 10% in estimating the remission rate for follow-up superior to 1 year. Results for LFE indicates that it should be employed only for follow-up at 2 years or more, in agreement with previous results,[21] and that LOCF should be proscribed.[37] For other follow-up times, methods considering data both before and after follow-up should be preferred. Complex methods accounting for the whole population such as LME3 and MICE perform well in imputing the mean value of the outcome but can induce changes in the value distribution of the imputed variable, causing considerable bias when estimating derived response rates. The biases of the imputed values are mainly due to the imputation method itself and the evolution in time of the outcome, and thus are not affected by the attrition bias nor by the amount of missing data at follow-up.

Although MICE becomes more popular,[15 38] the misspecification of the model and the violation of underlying

Table 1 Summary of the main findings for each imputation method considered

Method	Imputed mean disease activity		Imputed remission rate		Effect of random missingness	Sensitivity to quantity of data around follow-up
	Baseline	Follow-up	Baseline	Follow-up		
LOCF	Unusable	Over-estimated	Unusable	Under-estimated	None	Yes
LFE	Unusable	Under-estimated	Unusable	Over-estimated	None	Yes
NAO	Strong under-estimation	Close to unbiased	Strong over-estimation	Close to unbiased	None	Small
LE	Strong under-estimation	Close to unbiased	Strong over-estimation	Slightly biassed	None	Small
PE	Strong under-estimation	Slightly biassed	Strong over-estimation	Close to unbiased	None	Yes
LME3	Strong under-estimation	Close to unbiased	Close to unbiased	Strong under-estimation	None	No
MICE	Strong under-estimation	Close to unbiased	Close to unbiased	Strong under-estimation	Reduction of bias of remission	No

LE, linear extrapolation; LFE, linear forward extrapolation; LME3, linear mixed effects cubic regression; LOCF, last observation carried forward; MICE, multiple imputation by chained equation; NAO, nearest available observation; PE, polynomial extrapolation.

assumptions such as conditional normality can lead to severe pitfalls[39 40] and the choice of covariate is critical.[41] Other studies reported changes in distribution of data imputed with MICE.[42] Though we included in the model several co-variates, they did not manage to reproduce the variability of the disease activity, even when they were the variables used to generate missing data. A possibility to overcome this problem could be the use of algorithms considering adjacent time.[43] Indeed, missing data imputation of longitudinal data with unequal time steps is a very active area of research.[43] But given the complexity of such method and the good results of simpler methods, MICE should be restricted to its use in regression models, where it has been widely studied and proven useful.[11 44 45]

There are strengths and limitations to the present study. The results of our study are supported by the use of data stemming from a large international collection of register of RA patient. Using purely simulated data may lead to excessively good results of the models due to having model assumptions being met too perfectly by the artificial data. Good care was taken for the simulation dataset to reproduce a disease duration and visit frequency structure, as well as pattern of missing data (MAR data), similar to the entire collection of register containing more than 49 000 patients. The large range of imputation techniques considered and the detailed effect of their use on both the mean value of disease activity and the remission rate is another asset of the present study. The results presented are reinforced by their robustness to the different sensitivity analysis performed.

However, using real data is also a limitation. Indeed, the main limitation of the present study is the use of a specific type of observational data coming from patients with active or established RA treated with biological DMARD. Although the results may be generalisable to cohorts with patients presenting a similar evolution of disease activity in time, they cannot be extrapolated without further studies to other types of cohort presenting a different

pattern of disease activity over time, such as early arthritis cohorts or remission cohorts. Furthermore, our study only considers missing data of patients still available in the cohort, and does not tackle the difficult problem of lost to follow-up data, for which imputation may not be the proper approach.[46]

The overall results of this simulation study suggest that imputation of function and disease activity assessments at baseline is best performed with linear mixed effects models with a cubic effect of time (LME3) or multiple imputation (MICE), and for imputation at follow-up times best with the nearest available observation (NAO).

Author affiliations
[1]Division of Rheumatology, Geneva University Hospitals, Geneva, Switzerland
[2]Department of Internal Medicine, Lund University, Lund, Sweden
[3]Department of Rheumatology, Skåne University Hospital, Malmö, Sweden
[4]Centre for Rheumatology and Spine Diseases, Rigshospitalet Glostrup, Glostrup, Denmark
[5]Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark
[6]Department of Rheumatology, Diakonhjemmet Hospital, Oslo, Norway
[7]Department of Rheumatology, Hospital Garcia de Orta, Almada, Portugal
[8]Institute of Rheumatology and Clinic of Rheumatology, Charles University, Prague, Czech Republic
[9]Department of Emergency and Transplantation, Rheumatology Unit, GISEA, University Hospital of Bari, Bari, Italy

ORCID iDs
Denis Mongin http://orcid.org/0000-0002-4801-8395
Kim Lauper http://orcid.org/0000-0002-4315-9009
Maria Jose Santos http://orcid.org/0000-0002-7946-1365

REFERENCES

1. Gill J, Prasad V. Improving observational studies in the era of big data. *The Lancet* 2018;392:716–7.
2. Sherman RE, Anderson SA, Dal Pan GJ, *et al*. Real-World evidence — what is it and what can it tell us? *N Engl J Med* 2016;375:2293–7.
3. Finckh A, Courvoisier D. Lessons learned from rheumatoid arthritis registries. *Joint Bone Spine* 2018;85:271–4.
4. Gabay C, Riek M, Hetland ML, *et al*. Effectiveness of tocilizumab with and without synthetic disease-modifying antirheumatic drugs in rheumatoid arthritis: results from a European collaborative study. *Ann Rheum Dis* 2016;75:1336–42.
5. Alemao E, Litman HJ, Connolly SE, *et al*. Do poor prognostic factors in rheumatoid arthritis affect treatment choices and outcomes? analysis of a US rheumatoid arthritis registry. *J Rheumatol* 2018;45:1353–60.
6. Soliman MM, Hyrich KL, Lunt M, *et al*. Effectiveness of rituximab in patients with rheumatoid arthritis: observational study from the British Society for rheumatology biologics register. *J Rheumatol* 2012;39:240–6.
7. Jinoos Y, Nick B, Megan C, *et al*. Rheumatology informatics system for effectiveness: a national Informatics-Enabled Registry for quality improvement. *Arthritis Care Res* 2016;68:1866–73.
8. Gottenberg JE, Courvoisier DS, Hernandez MV, *et al*. Brief report: association of rheumatoid factor and Anti-Citrullinated protein antibody positivity with better effectiveness of abatacept: results from the pan-European registry analysis. *Arthritis Rheumatol* 2016;68:1346–52.
9. Rubin DB. Inference and missing data. *Biometrika* 1976;63:581–92.
10. Schafer JL, Graham JW. Missing data: our view of the state of the art. *Psychol Methods* 2002;7:147–77.
11. Baron G, Ravaud P, Samson A, *et al*. Missing data in randomized controlled trials of rheumatoid arthritis with radiographic outcomes: a simulation study. *Arthritis Rheum* 2008;59:25–31.
12. Bell ML, Fairclough DL. Practical and statistical issues in missing data for longitudinal patient-reported outcomes. *Stat Methods Med Res* 2014;23:440–59.
13. England BR, Sokolove J, Robinson WH, *et al*. Associations of circulating cytokines and chemokines with cancer mortality in men with rheumatoid arthritis. *Arthritis Rheumatol* 2016;68:2394–402.
14. Bergstra SA, Branco JC, Vega-Morales D, *et al*. Inequity in access to bDMARD care and how it influences disease outcomes across countries worldwide: results from the METEOR-registry. *Ann Rheum Dis* 2018;77:1413–20.
15. Hayati Rezvan P, Lee KJ, Simpson JA. The rise of multiple imputation: a review of the reporting and implementation of the method in medical research. *BMC Med Res Methodol* 2015;15:30.
16. Azur MJ, Stuart EA, Frangakis C, *et al*. Multiple imputation by chained equations: what is it and how does it work? *Int J Methods Psychiatr Res* 2011;20:40–9.
17. Eekhout I, de Boer RM, Twisk JWR, *et al*. Missing data: a systematic review of how they are reported and handled. *Epidemiol Camb Mass* 2012;23:729–32.
18. DeSouza CM, Legedza ATR, Sankoh AJ. An overview of practical approaches for handling missing data in clinical trials. *J Biopharm Stat* 2009;19:1055–73.
19. Wong WK, Boscardin WJ, Postlethwaite AE, *et al*. Handling missing data issues in clinical trials for rheumatic diseases. *Contemp Clin Trials* 2011;32:1–9.
20. van der Heijde D, Fleischmann R, Wollenhaupt J, *et al*. Effect of different imputation approaches on the evaluation of radiographic progression in patients with psoriatic arthritis: results of the RAPID-PsA 24-week phase III double-blind randomised placebo-controlled study of certolizumab pegol. *Ann Rheum Dis* 2014;73:233–7.
21. Markusse IM, Landewé R, Wolterbeek R, *et al*. Linear extrapolation of missing radiographic change scores in clinical trials does not spuriously overestimate group radiographic changes in rheumatoid arthritis. *Rheumatology* 2016;55:1295–300.
22. Song J, Boscardin WJ, Furst DE, *et al*. A cautionary tale: dealing with missing data in clinical trials for rheumatic diseases. *Clin Exp Rheumatol* 2014;32:S-122–126.
23. MÁ D, Garcia VV, González-Alvaro I, *et al*. Tackling missing radiographic progression data: multiple imputation technique compared with inverse probability weights and complete case analysis. *Rheumatology* 2013;52:331–6.
24. Iannone F, Courvoisier DS, Gottenberg JE, *et al*. Body mass does not impact the clinical response to intravenous abatacept in patients with rheumatoid arthritis. Analysis from the "pan-European registry collaboration for abatacept (PANABA). *Clin Rheumatol* 2017;36:773–9.10.1007/s10067-016-3505-5
25. Engels J, Diehr P. Imputation of missing longitudinal data: a comparison of methods. *J Clin Epidemiol* 2003;56:968–76.
26. Zhang Z. Missing data imputation: focusing on single imputation. *Ann Transl Med* 2016;4.
27. Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. John Wiley & Sons, 2004.
28. Schafer JL. Multiple imputation: a primer. *Stat Methods Med Res* 1999;8:3-15.
29. Rodwell L, Lee KJ, Romaniuk H, *et al*. Comparison of methods for imputing limited-range variables: a simulation study. *BMC Med Res Methodol* 2014;14:57.
30. R Core Team. R: a language and environment for statistical computing. R foundation for statistical computing, 2017. Available: https://www.R-project.org/
31. Finckh A*et al*. Evidence for differential acquired drug resistance to anti-tumour necrosis factor agents in rheumatoid arthritis. *Ann Rheum Dis* 2006;65:746–52.
32. Cruyssen B, Van Looy S, Wyns B, *et al*. Four-Year follow-up of infliximab therapy in rheumatoid arthritis patients with long-standing refractory disease: attrition and long-term evolution of disease activity. *Arthritis Res Ther* 2006;8.
33. Gottenberg J-E, Brocq O, Perdriger A, *et al*. Non–TNF-Targeted biologic vs a second anti-TNF drug to treat rheumatoid arthritis in patients with insufficient response to a first anti-TNF drug. *JAMA* 2016;316:1172–80.
34. Rubbert-Roth A, Finckh A. Treatment options in patients with rheumatoid arthritis failing initial TNF inhibitor therapy: a critical review. *Arthritis Res Ther* 2009;11.
35. Finckh A, Möller B, Dudler J, *et al*. Evolution of radiographic joint damage in rituximab-treated versus TNF-treated rheumatoid arthritis cases with inadequate response to TNF antagonists. *Ann Rheum Dis* 2012;71:1680.2–5.
36. van BS. *Flexible Imputation of Missing Data, Second Edition*. 2 edition. Boca Raton: Chapman and Hall/CRC, 2018.
37. Lydersen S. Last observation carried forward. *Tidsskr Den Nor Legeforening*. [Epub ahead of print: 23 May 2019].
38. Cummings P. Missing data and multiple imputation. *JAMA Pediatr* 2013;167:656–61.
39. Sterne JAC, White IR, Carlin JB, *et al*. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009;338:b2393.
40. White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. *Stat Med* 2011;30:377–99.
41. Berchtold A, Surís J-C. Imputation of repeatedly observed multinomial variables in longitudinal surveys. *Commun Stat Simul Comput* 2017;46:3267–83.
42. Marston L, Carpenter JR, Walters KR, *et al*. Issues in multiple imputation of missing data for large general practice clinical databases. *Pharmacoepidemiol Drug Saf* 2010;19:618–26.
43. De Silva AP, Moreno-Betancur M, De Livera AM, *et al*. A comparison of multiple imputation methods for handling missing values in longitudinal data in the presence of a time-varying covariate with a non-linear association with time: a simulation study. *BMC Med Res Methodol* 2017;17.
44. Donders ART, van der Heijden GJMG, Stijnen T, *et al*. Review: a gentle introduction to imputation of missing values. *J Clin Epidemiol* 2006;59:1087–91.
45. Keogh RH, White IR. Using full-cohort data in nested case-control and case-cohort studies by multiple imputation. *Stat Med* 2013;32:4021–43.

46. Kristensen LE, Saxne T, Geborek P. The LUNDEX, a new index of drug efficacy in clinical practice: results of a five-year observational study of treatment with infliximab and etanercept among rheumatoid arthritis patients in southern Sweden. *Arthritis Rheum* 2006;54:600–6.