RESEARCH ARTICLE   **Open Access**

Missing not at random in end of life care studies: multiple imputation and sensitivity analysis on data from the ACTION study

Giulia Carreras[1*] (iD), Guido Miccinesi[1], Andrew Wilcock[2], Nancy Preston[3], Daan Nieboer[4], Luc Deliens[5], Mogensm Groenvold[6], Urska Lunder[7], Agnes van der Heide[4], Michela Baccini[8] and ACTION consortium

Abstract

Background: Missing data are common in end-of-life care studies, but there is still relatively little exploration of which is the best method to deal with them, and, in particular, if the missing at random (MAR) assumption is valid or missing not at random (MNAR) mechanisms should be assumed. In this paper we investigated this issue through a sensitivity analysis within the ACTION study, a multicenter cluster randomized controlled trial testing advance care planning in patients with advanced lung or colorectal cancer.

Methods: Multiple imputation procedures under MAR and MNAR assumptions were implemented. Possible violation of the MAR assumption was addressed with reference to variables measuring quality of life and symptoms. The MNAR model assumed that patients with worse health were more likely to have missing questionnaires, making a distinction between single missing items, which were assumed to satisfy the MAR assumption, and missing values due to completely missing questionnaire for which a MNAR mechanism was hypothesized. We explored the sensitivity to possible departures from MAR on gender differences between key indicators and on simple correlations.

Results: Up to 39% of follow-up data were missing. Results under MAR reflected that missingness was related to poorer health status. Correlations between variables, although very small, changed according to the imputation method, as well as the differences in scores by gender, indicating a certain sensitivity of the results to the violation of the MAR assumption.

Conclusions: The findings confirmed the importance of undertaking this kind of analysis in end-of-life care studies.

Keywords: Missing data, MAR, MNAR, Advance care planning, Oncology, Quality of life

Background

Missing data are common in palliative and end-of-life care studies, where 20–50% of participants withdraw early, mostly because of deterioration and/or death [1]. A systematic review and meta-analysis of randomized controlled palliative intervention trials found about one quarter of primary endpoint data missing [2]. In the literature, several methods have been proposed for dealing with missing data, but there has been relatively little exploration of which is best in the palliative care setting [3]. The appropriateness of the chosen method is strongly related to the nature of the mechanism generating missing data. Missing data can be categorized as: missing completely at random (MCAR), when the probability of an observation being missing does not depend on both observed and unobserved data; missing at

* Correspondence: g.carreras@ispro.toscana.it
[1]Oncological Network, Prevention and Research Institute (ISPRO), Florence, Italy
Full list of author information is available at the end of the article

Carreras *et al. BMC Medical Research Methodology*   (2021) 21:13

Page 2 of 12

random (MAR), when the probability of an observation being missing depends only on the observed; and missing not at random (MNAR), when the probability of an observation being missing also depends on unobserved data [4].

Complete-case analysis, that consists in performing the analysis on the subset of subjects with complete information, was the most common approach to treat missing data in randomized clinical trials as recently as 2013 [5, 6]. However, unless the missingness mechanism is MCAR, this may lead to biased results. When the MCAR assumption is not valid, alternative strategies can be adopted to deal with missing data: inverse probability weighting, doubly robust inverse probability weighting, maximum likelihood estimation, multiple imputation (MI). Among them, MI is widely recognised as the most appropriate one in many fields [7]. MI creates several complete versions of the data by replacing each missing value with more than one plausible value. Each of the resulting complete datasets is then analyzed with standard statistical methods and the results pooled for final inference using the Rubin's combination rule, to obtain a point estimate and a measure of precision which accounts for uncertainty due to missing information [4, 8]. There are several ways to implement MI that could be run under MAR and MNAR [9–12]. One of these is the Multivariate Imputation by Chained Equations (MICE) which relies on the MAR assumption [9, 10], but can be modified in order to account for MNAR mechanisms [13].

Missing data in end-of-life care studies can arise due to the fact that one or more questionnaire items are missing, or due to the fact that all the questionnaire items are missing (missing form). While missing items may be due to simple omissions in questionnaire compilation, reasons for missing a whole questionnaire may relate to sudden change in the patient's health status or to the patient's sensitivity to specific issues which could be not adequately measured by the collected variables. In these cases, the missingness mechanism may be MNAR [14–18]. Although it is known that performing MI under the assumption of MAR when the actual missingness mechanism is MNAR may produce biased estimates [4], this issue is not widely appreciated in dealing with missing data in palliative care studies [3]. Even if it is not possible to distinguish between MNAR and MAR patterns using observed data, the robustness of the MAR assumption can be investigated through sensitivity analyses [4, 14]: if the results obtained under MAR and specific MNAR assumptions are similar, one can conclude that the presence of unobserved factors does not affect the conclusions.

The aim of this study was to compare the performance of different MI methods within the ACTION (Advance care planning – an innovative palliative care intervention to improve quality of life in oncology) study by treating separately missing items and missing questionnaires [15]. Using a preliminary dataset, we handled the missing data by applying a MICE procedure under the standard assumption of MAR and also under MNAR by using a pattern mixture-model approach, [4, 16] which distinguished between missing values due to missing items and missing the whole questionnaires. We focused on target analyses to evaluate the sensitivity of the results to the use of the two procedures.

Methods
Data
The ACTION study is a phase III multicenter cluster randomised controlled trial, following the CONSORT guidelines, on the effects of advance care planning (ACP) in patients with advanced lung or colorectal cancer. The ACTION Respecting Choices ACP intervention involves trained healthcare professionals who assist patients and their relatives in reflecting on the patient's goals, values and beliefs and in discussing their healthcare wishes. The intervention has the potential to improve current and future healthcare decision making, provide patients with a sense of control, and improve quality of life. In total, 22 hospitals in 6 countries were randomised to be intervention or control sites, with up to 1360 patients participating. At the intervention sites, patients are offered interviews with a trained ACP facilitator, whereas in the control sites, patients receive usual care. All participating patients are asked to complete questionnaires at baseline, and again after 2.5 and 4.5 months; the questionnaires assess quality of life, and the extent to which care as received is aligned with their preferences, their evaluation of decision-making processes and quality of end-of-life care (see additional file 2) [15].

Ethical approval was obtained from the Research Ethics Committee (REC) of the coordinating centre ('Medische Ethische Toetsings Commissie Erasmus MC'), as well as RECs in all participating countries. The trial was registered in the International Clinical Trials Registry Platform (ISRCTN63110516) per 10/3/2014.

Within the ACTION study it was decided to perform a preliminary analysis on a first subsample of the enrolled subjects, with the aim to explore methods for dealing missing data. Our analysis was based on the records of 487 patients, representing the 36% of the final expected sample, containing information collected through questionnaires at baseline and at 2.5 months, for a total of 121 variables. Baseline data included personal information (gender, age, marital status, living with a partner or alone, living in a private household or in an institution/care facility, children, years of education, religiosity, ethnic group),

information on diagnosis (small cell or non-small cell lung cancer, colon cancer, rectal cancer), cancer stage (III or IV lung cancer; metachronous metastases or IV colorectal cancer), WHO performance status (a measure of how well a person with cancer is able to carry on ordinary daily activities), and current treatment (chemotherapy, radiation therapy, immunotherapy, targeted therapy). Using baseline and follow-up questionnaires, we calculated scores for quality of life and symptoms (primary endpoints) and for shared decision-making, satisfaction with care and coping with illness (secondary endpoints) (Table 1). All the scores, except that measuring patient involvement which was built ad hoc for the present study, are validated and largely used in the context of psychometrics [17–22]. They are continuous variables ranging between 0 and 100. Since an intermediate analysis on the treatment effect was not planned by protocol, this preliminary analysis was blinded in respect to treatment arm and country of the participants, as well as in respect to their survival. Due to the fact that few patients are expected to die during the first 2.5 months of follow up, for sake of simplicity, in our analysis we assumed that all patients were still alive at the completion of the second questionnaire, meaning that no form was missing because of patient's death (see discussion).

Statistical methods

In our analysis we first applied the MICE procedure under MAR assumption. The procedure was applied to all the variables with missing data, including variables on personal information, and on diagnosis and treatments. Then we performed a sensitivity analysis assuming alternative MNAR mechanisms, within a pattern mixture-model approach [4, 16]. MI under MNAR required the specification of additional assumptions on the missingness mechanism and the modification of the MICE algorithm.

According to the main purpose of the ACTION study, the sensitivity analysis focused on relevant outcomes of quality of life and symptoms scores. We estimated means and 90% confidence intervals of these outcomes on the overall sample and by gender and their correlations with other scores (i.e. shared decision making, satisfaction with care and coping with illness), under the different MI approaches. Differences among MI procedures were evaluated by comparing point estimates and their confidence intervals in a descriptive way. If, under all the MNAR models, the direction of the explored relationships was the same as that under MAR and the confidence intervals largely overlapped, we concluded that the results were robust to violation of the MAR assumption. Due to the fact that at this stage of the analysis we are not interested in performing statistical tests, but only in describing the results, we preferred to report the 90% confidence intervals instead of the 95% ones in order to discourage their use as a surrogate of p-value = 0.05 [23].

Multiple imputation by chained equations under MAR assumption

Let $\mathbf{Y} = (Y_1,...,Y_J)$ be the $n \times J$ matrix of the data, where $Y_j = (y_{1j}\ y_{2j}...y_{nj})^{\mathrm{T}}$ is the vector of length n representing the values assumed by the jth variable in the n subjects. Let $\mathbf{R} = (R_1,..., R_J)$ be a $n \times J$ matrix, with $R_j = (r_{1j}, r_{2j}, ...r_{nj})^{\mathrm{T}}$ vector of the missingness indicators for the jth variable ($r_{ij} = 0$ if y_{ij} is missing and $r_{ij} = 1$ otherwise). Let us denote the observed entries and the missing entries of $\mathbf{Y}$ as $\mathbf{Y}_{\mathrm{obs}}$ and $\mathbf{Y}_{\mathrm{mis}}$, respectively. Analogously, let $Y_{j\ \mathrm{obs}}$ and $Y_{j\ \mathrm{mis}}$ be the observed and the missing entries of the vector Y_j. Under the MAR

Table 1 Primary and secondary endpoints (scales) measured with questionnaires at baseline and follow-up

Scale	Questionnaire	Reference
Quality of life and symptoms (built on 17 17 items)		
Pain		
Dyspnoea		
Insomnia		
Appetite loss		
Constipation	EORTC QLQ-C15-PAL	21
Physical functioning		
Fatigue		
Nausea and vomiting		
Quality of life		
Emotional functioning	QLQ-C30	22
Shared decision-making (built on 5 items)		
Decision making participation self-efficacy	APECC	23
Patient involvement	Self-constructed questions	
Satisfaction with care (built on 7 items)		
Doctors information provision		
Nurses information provision	EORTC IN-PATSA32	24
Overall quality of care		
Coping with illness (built on 12 items)		
Active coping		
Planning		
Acceptance	COPE	25-26
Denial		

Carreras *et al. BMC Medical Research Methodology*   (2021) 21:13

Page 4 of 12

assumption of conditional independence between $\mathbf{Y}_{\text{mis}}$ and $\mathbf{R}$ given the observed data $\mathbf{Y}_{\text{obs}}$, $f(\mathbf{Y}_{\text{mis}}|\mathbf{Y}_{\text{obs}}, \mathbf{R}) = f(\mathbf{Y}_{\text{mis}}|\mathbf{Y}_{\text{obs}})$, the MICE algorithm required the specification of the univariate conditional distributions $f(Y_j|Y_{-j}, \boldsymbol{\theta}_j)$ in the form of regression models for the variables with missing values, where $Y_{-j} = (Y_1, ..., Y_{j-1}, Y_{j+1}, ..., Y_J)$ and $\boldsymbol{\theta}_j$ was a vector of unknown regression parameters [9]. In our analysis, in order to avoid collinearity and related computational problems, assuring at the same time good prediction, a preliminary selection of the predictors to be included in the regressions was performed (in practice some of the elements of $\boldsymbol{\theta}_j$ have been set to zero). In particular, we regressed each variable Y_j with missing entries on each of the other variables, completed by a preliminary imputation which draw from their empirical marginal distribution. Then, we calculated the Akaike's information criterion (AIC) for each regression and selected as predictors for Y_j the 15 variables which lead to the smallest AIC values [24]. Considering the number of observations and the large percentage of missing values in most of the follow up scores, 15 was considered an appropriate number of covariates [10]. The marginal regressions used in this preliminary procedure and in the following imputation algorithms varied according to the nature of the outcome variable (linear regression for continuous variables, logistic regression for factor variables with 2 levels, multinomial logistic regression for factor variables with more than 2 levels, and proportional odds models for ordered variables). Finally, for each incomplete variable, the set of the selected predictors was enriched by including the indicators of gender, age, type of cancer and WHO performance status. The selected predictors for each incomplete variable are reported in Figure A1 of the additional file 1.

Once defined the conditional distributions $f(Y_j|Y_{-j}, \boldsymbol{\theta}_j)$ and assumed non-informative priors on $\boldsymbol{\theta}_j$, the MICE algorithm consisted in the following steps. First, we randomly drew an initial imputation $\hat{Y}_1, ..., \hat{Y}_J$ for the missing values in $\mathbf{Y}_{\text{mis}}$, by sampling from the empirical marginal distributions of the variables with missing entries (*step 0*). Then, for the first variable Y_j with missing entries:

1 we sampled $\hat{\theta}_j$ from the posterior distribution $f(\theta_j|Y_{j\,\text{obs}}, \hat{Y}_{-j})$,
2 we drew $\hat{Y}_{j\,\text{mis}}$ from the posterior predictive distribution $f(Y_j|Y_{j\,\text{obs}}, \hat{Y}_{-j}, \hat{\theta}_j)$.

We repeated *steps 1* and *2* sequentially for each variable with missing entries in the dataset, and we repeated the entire procedure, excluding *step 0*, until algorithm convergence (100 iterations) [25]. At the end, we got a complete dataset. We created several complete versions of the data by repeating the procedure M times, choosing M according to the rule of thumb based on the average percentage rate of missingness [26].

The MICE procedure was implemented by using the `mice` library of the R software [27, 28].

Multiple imputation by chained equations under MNAR assumption

Let us suppose that the MAR assumption is not valid. This, in general, implies that the posterior predictive distribution of $Y_{j\,\text{mis}}$ at step 2 of the MICE algorithm depends on $\mathbf{R}$. In order to account for this dependency, according to the pattern mixture-model approach [4, 16], the MICE algorithm can be modified defining distinct posterior predictive distributions depending on the missing data patterns in $\mathbf{R}$. For example, let us assume that we are in the simple case in which the value of $Y_{j\,\text{mis}}$ depends on the observed and on the fact that Y_j is missing, but not on the fact that other variables are missing. In this case, we would have two distinct posterior predictive distributions, one for $R_j = 0$ and one for $R_j = 1$:

$$f(Y_j|Y_{j\,\text{obs}}, \hat{Y}_{-j}, \hat{\theta}_j^{R_j=0}) \text{ and } f(Y_j|Y_{j\,\text{obs}}, \hat{Y}_{-j}, \hat{\theta}_j^{R_j=1}),$$

and the step 2 of the MICE algorithm can be modified as follows. After having sampled $\hat{\theta}_j^{R_j=1}$ from the posterior distribution $f(\theta_j|Y_{j\,\text{obs}}, \hat{Y}_{-j})$, we generate $\hat{\theta}_j^{R_j=0}$ from the conditional distribution $f(\theta_j^{R_j=0}|\hat{Y}_{-j}, \hat{\theta}_j^{R_j=1})$, a priori defined according to an assumed hypothesis on the MNAR mechanism. Finally, we sample $\hat{Y}_{j\,\text{mis}}$ from the posterior predictive distribution $f(Y_j|Y_{j\,\text{obs}}, \hat{Y}_{-j}, \hat{\theta}_j^{R_j=0})$. More in general, when the value of $Y_{j\,\text{mis}}$ depends also on the fact that other variables are missing, there are in principle distinct predictive distributions for each pattern of missing data [29].

Partition of the missing values and modified MICE in the ACTION study

In our analysis, we explored the violation of the MAR assumption, making a distinction between missing items due to the fact that the patient did not reply to some items in a questionnaire and missing items due to the fact that the form was completely missing (this could happen for the second questionnaire, but not for the baseline one), for which a violation of the MAR assumption could be hypothesised. In particular, we introduced two matrices of missingness indicators, $\mathbf{R^I}$ which indicated the missing values of the first type and $\mathbf{R^F}$ which indicated the missing values of the second type. The two matrices described a partition of the missing values; for simplicity we called $\mathbf{M^I}$ the first element of the partition (collection of the missing data defined by $\mathbf{R^I}$) and $\mathbf{M^F}$ (collection of missing data defined by $\mathbf{R^F}$). Then, we assumed a possible MNAR mechanism for the missing values in $\mathbf{M^F}$:

$$f(\mathbf{Y}_{\mathrm{mis}}|\mathbf{Y}_{\mathrm{obs}}, \mathbf{R^I}, \mathbf{R^F}) = f(\mathbf{Y}_{\mathrm{mis}}|\mathbf{Y}_{\mathrm{obs}}, \mathbf{R^F}),$$

with f $(\mathbf{Y}_{\mathrm{mis}}|\mathbf{Y}_{\mathrm{obs}}, \mathbf{R^F})$ possibly different from f $(\mathbf{Y}_{\mathrm{mis}}|\mathbf{Y}_{\mathrm{obs}})$.

Two distinct posterior predictive distributions were defined for each Y_j with missing values in $\mathbf{M^F}$, one for $R^F_j = 0$ and one for $R^F_j = 1$:

$$f(Y_j|Y_{j\mathrm{obs}}, \hat{Y}_{-j}, \hat{\theta}_j^{R^F_j=0}) \text{ and } f(Y_j|Y_{j\mathrm{obs}}, \hat{Y}_{-j}, \hat{\theta}_j^{R^F_j=1}).$$

The described violation of the MAR assumption required a modification of the MICE algorithm that relied on a model for $f(\theta_j^{R^F_j=0}|\hat{Y}_{-j}, \hat{\theta}_j^{R^F_j=1})$. The modified MICE algorithm that we implemented for our analysis was the following. After having randomly drawn an initial imputation $\hat{Y}_1, \ldots, \hat{Y}_J$ for the missing values by sampling from the empirical marginal distributions of the variables with missing entries (*step 0*), we imputed the missing values for the first variable Y_j with missing entries, according to the following steps:

1. we sampled $\hat{\theta}_j$ from the posterior distribution $f(\theta_j|Y_{j\,\mathrm{obs}}, \hat{Y}_{-j})$,
2. we drew $\hat{Y}_{j\,\mathrm{mis}}$ from the posterior predictive distribution $f(Y_j|Y_{j\,\mathrm{obs}}, \hat{Y}_{-j}, \hat{\theta}_j)$,
3. if Y_j had missing entries in $\mathbf{M^F}$, we generated $\hat{\theta}_j^{R^F_j=0}$ from the conditional distribution $f(\theta_j^{R^F_j=0}|\hat{Y}_{-j}, \hat{\theta}_j^{R^F_j=1})$, a priori defined according to our hypothesis on the MNAR mechanism (see below), and we sampled $\hat{Y}_{j\,\mathrm{mis}}$ from the posterior predictive distribution $f(Y_j|Y_{j\,obs}, \hat{Y}_{-j}, \hat{\theta}_j^{R^F_j=0})$.

We repeated *step 1-step 3* for all variables with missing entries sequentially, until algorithm convergence (100 iterations). We repeated the whole procedure M times in order to create M complete versions of the data set. In running the algorithm, we ordered the variables according the number of missing values (from the variable with fewer missing entries to that one with more missing entries).

Assumptions on the MNAR mechanism
We assumed that the MAR assumption was violated with reference to the six primary endpoints expressing the patient's health status, as measured by the second questionnaire: follow-up scores pain (PA), dyspnoea (DY), emotional functioning (EF), physical functioning (PF), fatigue (FA) and quality of life (QOL). Under the general assumption that reasons for missing the second questionnaire as a whole may relate to sudden changes in the patient's health status or to sensitivity to specific issues not adequately measured by the other variables [4,

14, 30–33], patients with a poorer health status were those for whom we expected to have larger probability of missing form. Our definition of the MNAR mechanisms reflected these hypotheses.

Even if in principle the model for $f(\theta_j^{R^F_j=0}|\hat{Y}_{-j}, \hat{\theta}_j^{R^F_j=1})$ could be highly complex, we specified two simple alternative models [13, 34, 35]. The first model assumed a constant shift k in the expected value of each of the six variables between observed and missing observations belonging to $\mathbf{M^F}$. We considered four different shifts $k_1, \ldots, k_4$ defined as the equispaced values between 0 and ½ interquartile range of the variable. The second model allowed the shift k to vary among individuals according to their WHO performance status. Let Y_{WHO} be the variable expressing the WHO performance status, and $sd(Y_{j\,\mathrm{obs}})$ the standard error for one of the six primary endpoints, calculated on the observed values; the shift k was defined as $k = Y_{\mathrm{WHO}} \times sd(Y_{j\mathrm{obs}})/4$. The shift was always assumed to be in the direction of worsening the patient's health conditions (subtracted for QOL, PF, EF and added for the other scores).

Analysis of data
We implemented the algorithm for the MNAR model by modifying the existing `mice` function of the `mice` R library. We used the passive imputation built-in method in order to incorporate at each step of the imputation algorithm the transformations required by the MNAR assumptions [27, 28, 34].

At the end of the MI procedure (MICE or modified MICE), each of the M complete datasets was analysed applying standard statistical methods and the results were combined according to the Rubin's rules [4]. Indicating with Q the unknown parameter of interest, for example the average quality of life score, let $\hat{Q}_m$ be the point estimate of Q and $\hat{U}_m$ the estimate of its variance, arising from the analysis on the mth dataset ($m = 1, \ldots, M$). The combined estimate of Q was equal to $\overline{Q} = \sum_{m=1}^{M} \frac{\hat{Q}_m}{M}$ and its estimated variance was $T = U + (1 + 1/M)\ B$, where $U = \sum_{m=1}^{M} \frac{\hat{U}_m}{M}$ and $B = \frac{\sum_{m=1}^{M}(\hat{Q}_m - \overline{Q})^2}{(M-1)}$.

Results
In Tables 2 and 3, we report descriptive statistics and percentage of missing values for socio-demographic, diagnosis and treatment variables measured at baseline, and for baseline and follow-up scores of quality of life and symptoms, shared decision-making, satisfaction with care and coping with illness. Between 1 and 15% of socio-demographic data, clinical data and baseline scores

were missing. Missing values at follow-up were around 36–39%, mostly due to missing forms (170/487; 35%).

In the additional file 1, we reported the missing data pattern (figure A1), as well as the selected predictors for each univariate conditional regression used in the imputation procedures (figure A2).

Table 4 shows means and 90% confidence intervals of quality of life scores and symptoms at follow up, calculated from the observed data and after MI under MAR and MNAR assumptions setting $M = 40$, which approximately corresponded to the maximum percentage rate of missingness [26]. When compared with the values calculated on the observed data, the results under MAR suggested that missingness was related to poor health status and lower quality of life, with the means after MICE moving in the direction of a worse health, i.e. decreasing for QOL, PF, EF and increasing for PA, DY and FA. Assuming MNAR mechanisms the discrepancy between imputed values and observed values was markedly greater.

Table 5 shows the mean scores by gender, calculated on the observed, under the MAR model and under the MNAR model that assumed a degree of departure from MAR dependent on the WHO performance status. Under both imputation approaches, there was a clear difference between genders, with males reporting a better health state than females. The size of the difference between genders increased under the MNAR model.

All correlations of secondary endpoints (patient involvement, overall quality of care, active coping, denial) with primary endpoints related to symptoms (FA, PA, DY) were negative, whereas those with primary endpoints related to functioning (QOL, PF, EF) were positive, although very low. As an example, in Fig. 1 we show the correlation coefficients between QOL and the four selected secondary endpoints. Overall quality of care showed the higher correlations with primary endpoints. Although in general correlations after MICE were robust to violations of the MAR assumption, sometimes a certain discrepancy was observed between the different imputation methods (see additional file 1 Figures A3-A7). In particular, under the MNAR scenarios all correlations involving denial were weaker than under MAR. A similar behaviour was observed also for patient involvement.

Discussion

Reviews of randomized controlled trials indicate wide variation in how missing data are dealt with. Relatively few studies use MI (8%), and, when carried out, it is mainly under MAR assumption; few studies assume MNAR mechanisms [5, 6, 36]. As usual in end-of-life care studies, the proportion of missing values in the subset of the ACTION study data analysed in the present

paper was high, particularly in relation to quality of life and symptoms outcomes. The percentage of missing forms at follow up was around 37%, thus higher than the 23% reported in a recent meta-analysis of randomized controlled trials on palliative interventions [2].

Although the availability of many variables should make the MAR assumption reasonable, quality of life is a multidimensional and complex concept and it is difficult to exclude the presence of relevant unmeasured factors, particularly in end of life care studies where reasons for missing values may relate to the patient's health status or their sensitivity to specific issues [14, 30–33]. Because the violation of the MAR assumption can significantly affect final results, particularly when the proportion of missing values is high, we performed a sensitivity analysis to evaluate if the results obtained assuming MAR were robust to the presence of MNAR. In particular, we focused on means and correlations between primary and secondary outcomes or between primary outcomes and individual characteristics (gender).

Performing MI under both MAR and MNAR assumptions produced different estimates of the average score of the primary outcomes. The results under MAR suggested that missingness was related to poor health status and lower quality of life. This is consistent with the common sense idea that more critical patients could have greater problems in filling questionnaires. As expected, assuming MNAR mechanisms which explicitly assigned higher probabilities of missing form to those patients who experienced worse health status, the discrepancy between imputed values and observed values became markedly greater. Males showed a better health than females. After imputation, the difference became even more marked, particularly under the MNAR models. The marginal correlations between primary and secondary endpoints were consistently negative for symptoms and positive for functioning and quality of life. As an example, quality of life was positively associated with patient involvement, overall quality of care, active coping and denial. On the contrary, pain was negatively correlated with the same variables. In agreement with clinical expectations, overall quality of care had the strongest correlations with the primary outcomes. The correlations of the primary outcomes with denial appeared to be weaker under the MNAR scenarios than when assuming MAR. These findings suggest that the provisional indicators which we focused on in the present analysis were not always robust to violation of the MAR assumption.

The peculiar features of our approach are mainly two. First, we integrated the MNAR model within the MICE algorithm as implemented in the `mice` function through the passive imputation built-in method [34], so that imputation error was correctly propagated [37]. This is

Table 2 Descriptive statistics on personal information and diagnosis information at baseline

Personal information					
Age		**Live with a partner, n (%)**		**Religious, n(%)**	
missing, n (%)	5 (1.03)	missing	11 (2.26)	missing	13 (2.67)
mean (sd)	65.99 (9.90)	yes	340 (69.82)	yes	225 (46.20)
Gender, n (%)		no	136 (27.93)	no	185 (37.99)
missing	4 (0.82)	**Living place**, n (%)		prefer not to specify	64 (13.14)
male	292 (59.96)	missing	14 (2.87)	**Minority ethnic group**, n (%)	
female	191 (39.22)	private household	447 (91.79)	missing	20 (4.11)
Marital status, n (%)		institution/care facility	3 (0.62)	yes	7 (1.44)
missing	8 (1.64)	other	23 (4.72)	no	460 (94.46)
married	324 (66.53)	**Children**, n (%)		**Years of education**	
unmarried	41 (8.42)	missing	8 (1.64)	missing, n (%)	73 (14.99)
divorced/separed	61 (12.53)	yes	414 (85.01)	mean (sd)	13.39 (4.74)
widowed	53 (10.88)	no	65 (13.35)		
Diagnosis information					
Type of cancer, n (%)		**WHO performance status**, n (%)		**Stage of cancer**, n (%)	
missing	4 (0.82)	missing	5 (1.03)	missing	7 (1.44)
small cell lung	62 (12.73)	0 fully active	167 (34.29)	III lung	59 (12.11)
non-small cell lung	218 (44.76)	1 symptomatic but completely ambulatory	255 (52.36)	IV lung	224 (46.00)
colon	155 (31.83)	2 symptomatic, < 50% in bed during the day	58 (11.91)	IV colorectal	134 (27.52)
rectal	48 (9.86)	3 symptomatic, > 50% in bed, but not bedbound	2 (0.41)	metachronous metast colorectal cancer	63 (12.94)
Current treatment					
Chemotherapy, n (%)		**Immunotherapy**, n (%)			
missing	0	missing	0		
no	119 (24.44)	no	450 (92.40)		
yes	368 (75.56)	yes	37 (7.60)		
Radiation therapy, n (%)		**Targeted therapy**, n (%)			
missing	0	missing	0		
no	449 (92.20)	no	428 (87.89)		
yes	38 (7.80)	yes	59 (12.11)		

different from following a two-step procedure that first imputes each outcome under the MAR assumption, and then modifies the imputed values according to a specific model for MNAR [13, 31, 33, 38, 39]. Additionally, our algorithm accounts for the different nature of the missing values in the data set: missing entries satisfying the MAR assumption (15.3% of the total number of missing values) and missing entries, due to missing form, for which a MNAR mechanism could be hypothesised (84.7%).

Table 3 Means of the scores (calculated using observed data) at baseline and at 2.5 months of follow-up (sd: standard deviation)

Variable	Baseline		Follow-up	
	mean (sd)	missing (%)	mean (sd)	missing (%)
Pain	21.41 (24.03)	6 (1.23)	20.67 (25.16)	175 (35.93)
Dyspnoea	25.28 (28.08)	7 (1.44)	26.33 (27.83)	173 (35.52)
Insomnia	27.36 (28.03)	7 (1.44)	25.27 (26.35)	173 (35.52)
Appetite loss	21.14 (29.95)	6(1.23)	21.87 (29.24)	173 (35.52)
Constipation	19.72 (29.14)	7(1.44)	18.33 (26.84)	176 (36.14)
Emotional functioning	78.41 (20.50)	11(2.26)	79.84 (21.22)	179(36.76)
Physical functioning	69.58 (23.91)	9 (1.85)	68.69 (25.60)	177 (36.34)
Fatigue	41.66 (26.22)	9 (1.85)	41.68 (25.70)	174 (35.73)
Nausea and vomiting	11.54 (20.42)	6 (1.23)	11.52 (22.39)	173 (35.52)
Quality of life	65.05 (20.27)	13 (2.67)	64.16 (21.84)	174 (35.73)
Decision making participation self-efficacy	75.06 (23.77)	17 (3.49)	76.53 (22.59)	179 (36.76)
Patient involvement	74.71 (18.24)	19 (3.90)	75.80 (18.30)	181 (37.17)
Doctors information provision	72.37 (23.74)	8 (1.64)	70.67 (24.41)	181 (37.17)
Nurses information provision	71.82 (24.29)	21(4.31)	70.05 (23.99)	191 (39.22)
Overall quality of care	76.21 (21.48)	12 (2.46)	73.03 (22.16)	183 (37.58)
Active coping	71.34 (25.49)	20(4.11)	68.26(27.40)	184 (37.78)
Planning	66.35 (27.97)	17 (3.49)	66.17 (27.92)	185 (37.99)
Acceptance	74.34 (21.70)	21 (4.31)	73.58 (22.34)	182(37.37)
Denial	23.20 (27.04)	25 (5.13)	20.68(25.37)	186 (38.19)

In the literature, several algorithms have been proposed that modify MICE accounting for the violation of the MAR assumption [13, 29, 34, 35, 38]. Recently, Tompsett and colleagues [29] proposed an approach, called Not-At-Random Fully Conditional Specification (NARFCS), which generalizes the MICE algorithm to MNAR, by including in each univariate conditional regression model the missingness indicator of the variable to be imputed (the coefficient of which cannot be estimated from the data and is the object of the sensitivity analysis), as well as the missingness indicators of the other incomplete variables. In this way, it is possible to allow for the presence of correlations between each variable and the missingness indicators of the others. Unlike NARFCS, our approach assumed that these correlations were equal to zero, because it included a missing form indicator only in the univariate conditional models used for the imputation of quality of life and symptoms scores (i.e. the variable interested by MNAR). However, in a sensitivity analysis, we included the missing form

Table 4 Mean (90% confidence interval) of the scores at 2.5 months of follow-up up (Quality of life; Pain; Dyspnoea; Emotional functioning; Physical functioning; Fatigue) observed and imputed under the MAR assumption and using different models for departure from MAR (k_{WHO}; k_1; k_2; k_3; k_4)

Variable	observed	MAR	k_{WHO}	k_1	k_2	k_3	k_4
Quality of life	64.16 (64.09, 64.24)	63.02 (62.94, 63.09)	58.82 (58.74, 58.90)	57.12 (57.04, 57.19)	58.63 (58.56, 58.71)	59.97 (59.89, 60.04)	61.65 (61.57, 61.72)
Pain	20.67 (20.59, 20.76)	22.37 (22.28, 22.46)	28.03 (27.92, 28.14)	34.17 (34.04, 34.30)	30.41 (30.29, 30.52)	27.51 (27.40, 27.61)	24.93 (24.83, 25.02)
Dyspnoea	26.33 (26.23, 26.42)	27.03 (26.94, 27.12)	36.99 (36.86, 37.12)	44.22 (44.07, 44.37)	38.64 (38.51, 38.78)	34.00 (33.89, 34.12)	30.20 (30.09, 30.30)
Emotional functioning	79.84 (79.77, 79.91)	78.62 (78.54, 78.69)	73.55 (73.46, 73.64)	66.21 (66.09, 66.33)	70.11 (70.01, 70.21)	73.70 (73.61, 73.79)	76.31 (76.22, 76.39)
Physical functioning	68.69 (68.60, 68.77)	67.24 (67.15, 67.33)	62.35 (62.26, 62.44)	59.94 (59.85, 60.04)	61.60 (61.51, 61.69)	63.23 (63.14, 63.32)	65.43 (65.34, 65.51)
Fatigue	41.68 (41.59, 41.76)	42.90 (42.82, 42.99)	49.62 (49.52, 49.72)	51.14 (51.04, 51.25)	48.91 (48.81, 49.01)	46.83 (46.73, 46.92)	44.75 (44.66, 44.84)

Table 5 Mean (90% confidence interval) of the scores at 2.5 months of follow-up (Quality of life; Pain; Dyspnoea; Emotional functioning; Physical functioning; Fatigue) for males and females observed and imputed under the MAR assumption and under the MNAR model assuming a shift depending on the WHO score (k_{WHO})

Variable	observed		MAR		k_{WHO}	
	males	females	males	females	males	Females
Quality of life	66.67 (66.60,66.74)	60.06 (59.98,60.13)	65.69 (65.62,65.76)	58.94 (58.87,59.02)	61.32 (61.25,61.39)	55.01 (54.93,55.08)
Pain	19.22 (19.15,19.30)	22.92 (22.82,23.01)	20.71 (20.62,20.80)	24.90 (24.82,24.99)	26.01 (25.91,26.10)	31.12 (31.01,31.24)
Dyspnoea	25.04 (24.95,25.14)	28.69 (28.59,28.79)	25.75 (25.66,25.84)	29.00 (28.90,29.09)	35.34 (35.24,35.44)	39.50 (39.36,39.64)
Emotional functioning	83.96 (84.02,83.90)	73.02 (73.10,72.94)	82.88 (82.81,82.95)	72.11 (72.04,72.19)	78.64 (78.57,78.71)	65.80 (65.70,65.89)
Physical functioning	69.86 (69.77,69.95)	66.50 (66.41,66.58)	68.82 (68.74,68.90)	64.83 (64.74,64.93)	63.90 (63.82,63.98)	59.99 (59.90,60.09)
Fatigue	39.09 (39.01,39.18)	46.19 (46.10,46.28)	40.14 (40.05,40.22)	47.12 (47.04,47.21)	46.54 (46.45,46.63)	54.33 (54.22,54.43)

indicator in each univariate conditional distribution, thus allowing for the possible correlation between having a missing form and all the variables in the data set. The results of this sensitivity analysis were very similar to those reported in the paper (see additional file 1 Tables A1-A2).

Recently, imputation algorithms based on nonlinear models, such as classification and regression trees or random forests, have spread. These approaches allow to deal with complex interactions and nonlinearities in the prediction models. Additionally, they do not require pre-selection of predictors and can be used also when the number of covariates is larger than the number of observations [40, 41]. However, the performance of these methods is not well stated. It depends on the presence and relevance of possible interaction effects and on the correlation structure of the data, and it could be quite poor when data are highly skewed [42–44]. It would be interesting to investigate the robustness to departure from the MAR assumption for multiple imputations approaches based on recursive partitioning. However, at the best of our knowledge, there are no studies dealing with such algorithms under the MNAR assumption and extending the pattern-mixture model to this context deserves ad hoc investigation.

Limitations
We fixed the number of imputations to 40 following a rule proposed by White and colleagues [26]. Since in this preliminary analysis we focused on different quantities, all surrogate in respect to the primary endpoint of the ACTION study, we did not perform a detailed investigation in order to determine an optimal number of imputations, which in principle should be based on the Fraction of Missing Information evaluated for the parameter of interest [26] or on criteria aimed to assure results stability over repetitions of the MI procedure [45].

The present results could be sensitive to the prediction models specification. Models with a larger number of predictors or with predictors selected with different methods (e.g. LASSO) or models which includes interactions could result more or less robust to violation of the MAR assumption, leading to different conclusions. This point will be addressed when the mixed-pattern approach proposed in this paper will be applied on the complete ACTION data set. As one of the main limitations of the present study, we would like to remark that we did not account for the cluster randomized design of the study, because information on treatment assignment and cluster variables (country and hospitals) was blinded to us [46, 47]. Similarly, also patient's survival at follow up was not available, so that we had no information about possible truncation by death. At this stage of the study, truncation by death was likely a minor problem since the analyses have been carried out with a follow-up of 2.5 months from recruitment and most of the enrolled patients had a good WHO performance status at baseline (the median survival for both small and non-small cell lung cancer in both stage III and IV is estimated to be around 7–8 months [48, 49], and 2.6 and 1.7 years for patients with and without metastasectomy, respectively [50]). Truncation by death is however a very relevant point that should be address in the future in order to get reliable estimates of the treatment effect [51–54]. Moreover, the MI on the complete ACTION dataset will be performed separately by treatment arm [46].

Conclusions
In imputing missing data in end-of-life care studies, sensitivity analyses for the departure from MAR should be performed. We proposed a modification of the MICE algorithm which accounts for the presence in the data set of two kind of missing data: missing entries satisfying the MAR assumption and missing entries, due to missing form, for which a MNAR mechanism could be hypothesised. We found that the results obtained after having imputed the missing values through MICE were not always robust to possible violations of the MAR assumption.
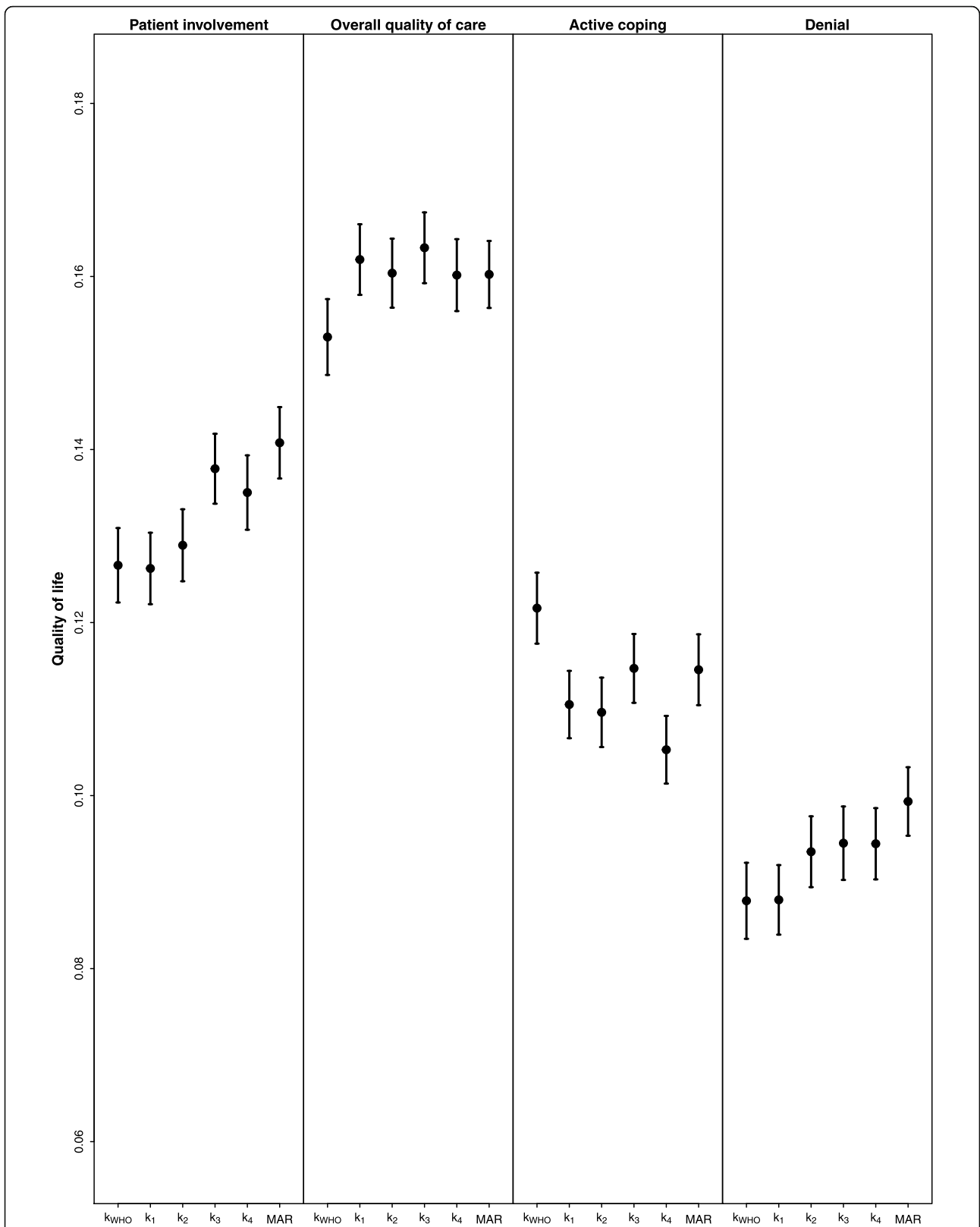
Fig. 1 Correlations (90% confidence intervals) between Quality of life and secondary endpoints (Patient involvement; Overall quality of care; Active coping; Denial) calculated after MI under the MAR assumption and under different MNAR models (k_{WHO}; k_1; k_2; k_3; k_4)

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12874-020-01180-y.

Additional file 1 Appendix. Appendix with supplementary material

Additional file 2. ACTION study patient questionnaire. ACTION study patient questionnaire

---

Abbreviations

ACP: Advance care planning; AIC: Akaike's information criterion; DY: Dyspnoea; EF: Emotional functioning; FA: Fatigue; MAR: The missing at random; MCAR: Missing completely at random; MI: Multiple imputation; MICE: Multivariate Imputation by Chained Equations; MNAR: Missing not at random; NARFCS : Not-At-Random Fully Conditional Specification; PA: Pain; PF: Physical functioning; QOL: Quality of life; REC: Research Ethics Committee

Authors' contributions

GC conceived the study, built the scores, developed the statistical models, analyzed the data and wrote the first version of the manuscript. GM built the scores, interpreted the data and contributed in writing the manuscript. MB conceived the study, developed the statistical models and wrote the first version of the manuscript. AW, NP, DN, LD, MG, UL, AH contributed in writing the manuscript. The authors read and approved the final manuscript.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate

Ethical approval for the ACTION project was obtained from the Research Ethics Committee (REC) of the coordinating centre ('Medische Ethische Toetsings Commissie Erasmus MC'), as well as RECs in all participating countries. Written informed consent to participate was obtained from all participants to the study.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Oncological Network, Prevention and Research Institute (ISPRO), Florence, Italy. [2]Department of Clinical Oncology, University of Nottingham, Nottingham, UK. [3]Lancaster University, International Observatory on end of life care, Lancaster, UK. [4]Department of Public Health, Erasmus University, Rotterdam, Netherlands. [5]Vrije Universiteit Brussel & Ghent University, Brussels, Belgium. [6]Department of Public Health, Copenhagen University, Copenhagen, Denmark. [7]University Clinic for Respiratory and Allergic Diseases, Golnik, Slovenia. [8]Department of Statistics, Computer Science, Applications 'G. Parenti' (DISIA), University of Florence, Florence, Italy.

References

1. Palmer JL. Analysis of missing data in palliative care studies. J Pain Symptom Manag. 2004;28:612–8.
2. Hussain JA, White IR, Langan D, Johnson MJ, Currow DC, Torgerson DJ, et al. Missing data in randomized controlled trials testing palliative interventions pose a significant risk of bias and loss of power: a systematic review and meta-analyses. J Clin Epidemiol. 2015;74:57–65.
3. Hussain JA, Bland M, Langan D, Johnson MJ, Currow DC, White IR. Quality of missing data reporting and handling in palliative care trials demonstrates that further development of the CONSORT missing data reporting guidance is required: a systematic review. J Clin Epidemiol. 2017;88:81–91.
4. Rubin DB. Multiple imputation for nonresponse in surveys. New York: Wiley; 1987.
5. Sullivan TR, Yelland LN, Lee KJ, Ryan P, Salter AB. Treatment of missing data in follow-up studies of randomised controlled trials: a systematic review of the literature. Clin Trials. 2017;14:387–95. https://doi.org/10.1177/1740774517703319.
6. Bell ML, Fiero M, Horton NJ, Hsu CH. Handling missing data in RCTs; a review of the top medical journals. BMC Med Res Methodol. 2014;14:118.
7. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, Wood AM, Carpenter JR. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ. 2009;338:b2393. https://doi.org/10.1136/bmj.b2393.
8. Little RJA, Rubin DB. Statistical analysis with missing data. New Jersey: Wiley; 2002.
9. Raghunathan TE, Lepkowski JM, van Hoewyk J, Solenberger P. A multivariate technique for multiply imputing missing values using a sequence of regression models. Surv Methodol. 2001;27:85–95.
10. Van Buuren S. Flexible imputation of missing data. Boca Raton: Chapman & Hall; 2018.
11. Jones M, Mishra GD, Dobson A. Analytical results in longitudinal studies depended on target of inference and assumed mechanism of attrition. J Clin Epidemiol. 2015;68:1165–75.
12. Hu B, Li L, Greene T. Joint multiple imputation for longitudinal outcomes and clinical events that truncate longitudinal follow-up. Stat Med. 2016;35:2991–3006.
13. Moreno-Betancur M, Chavance M. Sensitivity analysis of incomplete longitudinal data departing from the missing at random assumption: methodology and application in a clinical trial with drop-outs. Stat Methods Med Res. 2016;25:1471–89.
14. Preston NJ, Fayers P, Walters SJ, Pilling M, Grande GE, Short V, et al. Recommendations for managing missing data, attrition and response shift in palliative and end-of-life care research: part of the MORECare research method guidance on statistical issues. Palliat Med. 2013;27:899e907.
15. Post WJ, Buijs C, Stolk RP, de Vries EG, le Cessie S. The analysis of longitudinal quality of life measures with informative drop-out: a pattern mixture approach. Qual Life Res. 2010;19:137–48.
16. Fielding S, Fayers PM, McDonald A, McPherson G, Campbell MK, RECORD Study Group. Simple imputation methods were inadequate for missing not at random (MNAR) quality of life data. Health Qual Life Outcomes. 2008;6:57. https://doi.org/10.1186/1477-7525-6-57.
17. Pauler DK, McCoy S, Moinpour C. Pattern mixture models for longitudinal quality of life studies in advanced stage disease. Stat Med. 2003;22:795–809.
18. Diehr P, Johnson LL. Accounting for missing data in end-of-life research. J Palliat Med. 2005;8(Suppl 1):S50–7.
19. Rietjens JA, Korfage IJ, Dunleavy L, Preston NJ, Jabbarian LJ, Christensen CA, et al. Advance care planning- a multi-centre cluster randomised clinical trial: the research protocol of the ACTION study. BMC Cancer. 2016;8(16):264.
20. Ratitch B, O'Kelly M, Tosiello R. Missing data in clinical trials: from clinical assumptions to statistical analysis using pattern mixture models. Pharm Stat. 2013;12:337–47.
21. Groenvold M, Petersen MA, Aaronson NK, Arraras JI, Blazeby JM, Bottomley A, Fayers PM, de Graeff A, Hammerlid E, Kaasa S, Sprangers MA, Bjorner JB,

EORTC Quality of Life Group. The development of the EORTC QLQ-C15-PAL: a shortened questionnaire for cancer patients in palliative care. Eur J Cancer. 2006;42:55–64.

22. Fayers PM, Aaronson NK, Bjordal K, Groenvold M, Curran D, Bottomley A, on behalf of the EORTC Quality of Life Group. EORTC QLQ-C30 scoring manual. 3rd ed. Brussels: EORTC; 2001.

23. Arora NK, Weaver KE, Clayman ML, Oakley-Girvan I, Potosky AL. Physicians' decision-making style and psychosocial outcomes among cancer survivors. Patient Educ Couns. 2009;77:404–12. https://doi.org/10.1016/j.pec.2009.10.004.

24. Brédart A, Bottomley A, Blazeby JM, Conroy T, Coens C, D'Haese S, et al. An international prospective study of the EORTC cancer in-patient satisfaction with care measure (EORTC IN-PATSAT32). Eur J Cancer. 2005;41:2120–31.

25. Carver CS, Scheier MF, Weintraub JK. Assessing coping strategies: a theoretically based approach. J Pers Soc Psychol. 1989;56:267–83.

26. Stanton AL, Kirk SB, Cameron CL, Danoff-Burg S. Coping through emotional approach: scale construction and validation. J Pers Soc Psychol. 2000;78:1150–69.

27. Greenland S, Senn SJ, Rothman KJ, Carlin JB, Poole C, Goodman SN, Altman DG. Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. Eur J Epidemiol. 2016;31:337–50.

28. Li F, Baccini M, Mealli F, Zell ER, Frangakis CE, Rubin DB. Multiple imputation by ordered monotone blocks with application to the anthrax vaccine research program. J Comput Graphical Stat. 2014;23:877–92.

29. Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences. Stat Sci. 1992;7:457–511.

30. White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. Stat Med. 2011;30:377–99.

31. Van Buuren S, Groothuis-Oudshoorn K. MICE: Multivariate imputation by chained equations. R J Stat Softw. 2011;45:1–67.

32. R Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2013.

33. Tompsett DM, Leacy F, Moreno-Betancur M, Heron J, White IR. On the use of the not-at-random fully conditional specification (NARFCS) procedure in practice. Stat Med. 2018;37:2338–53.

34. Resseguier N, Giorgi R, Paoletti X. Sensitivity analysis when data are missing not-at-random. Epidemiology. 2011;22:282.

35. Leacy FP, Floyd S, Yates TA, White IR. Analyses of sensitivity to the missing-at-random assumption using multiple imputation with delta adjustment: application to a tuberculosis/HIV prevalence survey with incomplete HIV-status data. Am J Epidemiol. 2017;185:304–15.

36. Keene ON, Roger JH, Hartley BF, Kenward MG. Missing data sensitivity analysis for recurrent event data using controlled imputation. Pharmaceut Stat. 2014;13:258–64. https://doi.org/10.1002/pst.1624.

37. Little RJA. Selection and pattern-mixture models. In: Fitzmaurice G, Davidian M, Verbeke G, Molenberghs G, editors. Longitudinal data analysis, chapter 18. Boca Raton: CRC Press; 2009. p. 409–31.

38. Van Buuren S, Boshuizen HC, Knook DL. Multiple imputation of missing blood pressure covariates in survival analysis. Stat Med. 1999;18:681–94.

39. Carpenter JR, Kenward MG. Multiple imputation and its application. 1st ed. Chichester: Wiley; 2013.

40. Burgette LF, Reiter JP. Multiple imputation for missing data via sequential regression trees. Am J Epidemiol. 2010;172:1070–6.

41. Tang F, Ishwaran H. Random forest missing data algorithms. Stat Anal Data Min. 2017;10:363–77.

42. Doove LL, Van Buuren S, Dusseldorp E. Recursive partitioning for missing data imputation in the presence of interaction effects. Comput Stat Data Anal. 2014;72:92–104.

43. Hong S, Lynn HS. Accuracy of random-forest-based imputation of missing data in the presence of non-normality, non-linearity, and interaction. BMC Med Res Methodol. 2020;20:199.

44. Shah AD, Bartlett JW, Carpenter J, Nicholas O, Hemingway H. Comparison of random forest and parametric imputation models for imputing missing data using MICE: a CALIBER study. Am J Epidemiol. 2014;179:764–74.

45. Lu K. Number of imputations needed to stabilize estimated treatment difference in longitudinal data analysis. Stat Methods Med Res. 2017;26:674–90.

46. Sullivan TR, White IR, Salter AB, Ryan P, Lee KJ. Should multiple imputation be the method of choice for handling missing data in randomized trials? Stat Methods Med Res. 2018:962280216683570.

47. Speidel M, Drechsler J, Jolani S. R package hmi: a convenient tool for hierarchical multiple imputation and beyond. IAB-discussion paper. 2018. 16/2018.

48. Wang S, Tang J, Sun T, Zheng X, Li J, Sun H, Zhou X, Zhou C, Zhang H, Cheng Z, Ma H, Sun H. Survival changes in patients with small cell lung cancer and disparities between different sexes. Socioecon Statuses Ages Sci Rep. 2017;7:1339. https://doi.org/10.1038/s41598-017-01571-0.

49. Simmons CP, Koinis F, Fallon MT, Fearon KC, Bowden J, Solheim TS, Gronberg BH, McMillan DC, Gioulbasanis I, Laird BJ. Prognosis in advanced lung cancer--a prospective study examining key clinicopathological factors. Lung Cancer. 2015;88:304–9. https://doi.org/10.1016/j.lungcan.2015.03.020.

50. Siebenhüner AR, Güller U, Warschkow R. Population-based SEER analysis of survival in colorectal cancer patients with or without resection of lung and liver metastases. BMC Cancer. 2020;20(246). https://doi.org/10.1186/s12885-020-6710-1.

51. Yang F, Small DS. Using post-outcome measurement information in censoring-by-death problems. J Royal Stat Soc Series B. 2016;78:299–318.

52. Tchetgen Tchetgen EJ. Identification and estimation of survivor average causal effects. Stat Med. 2014;33:3601–780.

53. Wen L, Terrera GM, Seaman SR. Methods for handling longitudinal outcome processes truncated by dropout and death. Biostatistics. 2018;19:407–25.

54. Yang F, Ding P. Using survival information in truncation by death problems without the monotonicity assumption. Biometrics. 2018. https://doi.org/10.1111/biom.12883.

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