Missing forms and dropout in the TME quality of life substudy

Stiggelbout, A.M.; Putter, H.; Marijnen, C.A.M.; Klein-Kranenbarg, E.; Velde, C.J.H. van de

Quality of Life Research (2005) 14: 857–865

Missing forms and dropout in the TME quality of life substudy

H. Putter[1], C.A.M. Marijnen[2], E. Klein Kranenbarg[3], C.J.H. van de Velde[3] & A.M. Stiggelbout[4]
[1]*Department of Medical Statistics (E-mail: h.putter@lumc.nl);* [2]*Department of Clinical Oncology;* [3]*Department of Surgery;* [4]*Department of Medical Decision Making, Leiden University Medical Centre, Leiden, The Netherlands*

Accepted in revised form 14 June 2004

Abstract

Objective: Missing forms may pose problems in health related quality of life (QOL) studies, because the absence of a QOL measure may be related to the patient's health and hence to the patient's QOL itself. Studying patterns of missingness, dropout, and the possible impact of missing data on QOL measures is an important step in reporting outcomes of QOL studies. We study patterns of dropout and evaluate the impact of missing forms in the TME QOL substudy. *Methods*: Patients with rectal cancer, randomized to receive either radiotherapy plus total mesorectal excision (TME) or TME only were included in the TME trial. QOL was evaluated in 1302 Dutch patients, before treatment, and 3, 6, 12, 18 and 24 months after surgery. Here only the visual analogue score (VAS) was studied. *Results*: At baseline, differences between VAS scores were found with respect to whether the QOL forms were dated before or after radiotherapy and surgery. Differences were small between different statistical methods accounting for dropout; only a cross-sectional analysis gave biased results. *Conclusion*: The results of the sensitivity analysis indicated that a linear mixed model analysis is a reliable and attractive approach for this study.

Key words: Missing data, Non-random dropout, Pattern-mixture models, Rectal cancer

Introduction

The occurrence of missing values is a virtually unavoidable problem in clinical studies in general. In health related quality of life studies in particular, missing values pose problems since the absence of a value at a certain point in time may be related to the severity of disease of the patient. Since this severity of disease may determine the quality of life score, absence or presence of the value is then directly related to the unobserved outcome. For the subsequent statistical analysis, data missing not at random (MNAR) are a definite problem, since summarizing the available data at any time point will generally result in an overestimation of the quality of life, the reason being that patients with poor quality of life will have a higher probability of being missing. While missing data may pose serious problems, there are of course other reasons for data being missing that are perhaps not related to illness. Patients may have forgotten to return their questionnaires or may have returned them too late; logistic problems may have resulted in the questionnaires not having been processed adequately, either before sending or after receiving them [1]. This calls for the need to incorporate patterns of missingness or dropout into the statistical analysis. A first step towards formulating a model explicitly modelling dropout should be an assessment of how many missing values occur at each of the scheduled time points and why values are missing. Then, since many methods incorporating dropout often make untestable assumptions, a sensitivity analysis comparing a number of such methods allow for better evaluation of how conclusions depend on the

missing data mechanism. We illustrate these steps using data from a large randomized clinical trial in rectal cancer treatment.

In rectal cancer, the quality of life is related to the prognosis of the patient, but also to the impact of the disease and treatment on patients health status. Local recurrences are a major problem, since they cause severe disabling symptoms and are difficult to treat. Over the past decade, the introduction of a new surgical technique, total mesorectal excision (TME), has reduced the number of local recurrences considerably [2, 3]. In addition, short-term preoperative radiotherapy has demonstrated to improve local control in resectable rectal cancer [4]. In a large randomized study the value of short-term preoperative radiotherapy (5×5 Gy) was evaluated in patients with resectable rectal cancer and operated with the TME technique. Therapeutic outcomes have been reported previously and demonstrated a significant decrease in the local recurrence rate at 2 year: 8.2% in the TME alone group vs. 2.4% in the radiotherapy and surgery group [5]. Quality of life and cost outcomes were secondary endpoints in the trial. Questionnaires on health related quality of life were to be filled out and returned by the patients before randomization (pre), and at 3, 6, 12, 18 and 24 months after surgery.

The purpose of this paper is twofold: first, to give an assessment, as precise as possible, of how much data is missing at each scheduled quality of life measurement and for what reason. The second purpose is to compare a number of analysis strategies, some of them incorporating informative dropout in the methods, in order to establish a reliable and simple method of analysis.

Methods

Study design

From January 1996 until December 1999, 1861 patients were randomized to preoperative radiotherapy with 5×5 Gy followed by standardized TME surgery or to TME surgery only. In- and exclusion criteria have been described in detail elsewhere [5]. The majority of the included patients (1530) were from the Netherlands; the other 331 patients were included by Swedish, other European and Canadian co-investigators. Quality of Life was evaluated in Dutch patients only. Informed consent was obtained from all patients for the study and was separately obtained for the quality of life study. All patients underwent surgery according to the Total Mesorectal Excision principle, as advocated by Heald [2]. Patients assigned to preoperative radiotherapy received a total dose of 25 Gy in 5 fractions over 5–7 days, followed by TME within 10 days from start of radiotherapy.

Quality of life assessment

Patients were requested to fill out a quality of life questionnaire at the following time points: pretreatment, 3, 6, 12, 18 and 24 months after surgery. Patients who failed to return two consecutive questionnaires were considered as withdrawn from the study and did not receive further questionnaires. The following time-windows were allowed: 3 months (1.5–4.5), 6 months (4.5–9), 12 months (9–15), 18 months (15–21) and 24 months (21–27) months. Forms with no date on it or outside the time-windows were considered missing. The effect of pretreatment forms filled in before or after randomization, start of radiotherapy and surgery was evaluated.

Health-related quality of life was assessed by the patients using the Rotterdam Symptom Check List (RSCL) [6], a questionnaire with 38 items which contains a physical and a psychological distress sub-scale. In addition to this, patients were requested to score their general well-being over the last week on a Visual Analogue Scale (VAS). The patients were asked to rate their well-being by placing a mark on a 100 mm horizontal line, anchored by optimal health and death. The score is the number of millimeters from the 'death' anchor to the mark, divided by 100. For the purpose of this paper, only VAS scores are used.

Statistical analysis

Assessment of missingness

VAS scores were compared according to whether the quality of life forms were dated before, at, or after day of randomization, start of radiotherapy (RT + TME group only), and surgery, both for all patients and for the treatment groups separately.

One-way ANOVA with Bonferroni correction for multiple testing was used for testing differences. The independent samples *t*-test was used for testing differences between treatment groups.

Because many patients were excluded from the baseline measurement because the forms were filled in and/or returned too late, we divided patients with complete quality of life followup into three categories. In the first one all forms are present (note that a particular VAS score could still be missing due to missing the VAS item or undeterminable score), the second category consists of those patients of whom only the first quality of life form is missing, the third consists of all other patients and is a heterogenous mix of patients that have at least one form missing (excluding patients with only baseline form missing). Differences between VAS scores were investigated between these three groups using one-way ANOVA with Bonferroni correction for multiple testing.

All analyses in this subsection were done with SPSS 11.0 (Chicago).

Comparison of analysis strategies

Missing data are typically divided into three categories, according to the division made by Rubin [7]: missing completely at random (MCAR), missing at random (MAR) and missing not a random (MNAR).

Let y_i denote the vector of complete observations of patient i; denote the observed and missing values of the patient by y_i^{obs} and y_i^{miss} respectively and let r_i be a vector of 0's and 1's indicating whether or not the corresponding y_i component is missing or not.

The strictest assumption of MCAR specifies that the probability of a missing observation is independent both of observed and of missing observations, i.e.

$$P(r_i|y_i^{\text{miss}}, y_i^{\text{obs}}) = P(r_i).$$

The MAR assumption specifies that the probability of a missing observation may depend on the observed, but not on the missing data,

$$P(r_i|y_i^{\text{miss}}, y_i^{\text{obs}}) = P(r_i|y_i^{\text{obs}}).$$

Finally, the MNAR assumption specifies that the probability of a missing observation may depend both on the observed and on the missing observations,

$$P(r_i|y_i^{\text{miss}}, y_i^{\text{obs}}) = P(r_i|y_i^{\text{miss}}, y_i^{\text{obs}}).$$

When covariates are incorporated into the analysis, these assumptions need to be rephrased slightly, see Rubin [7].

In this paper we concentrate on dropout. Dropouts at a certain time point are defined as patients who did not submit a valid form at that specific time point or any later time point. For data that are missing before dropout we assume that they are MAR. Moreover, distinguishing between MAR and MNAR is intrinsically difficult and often requires untestable assumptions [8].

We report three different analyses, each of which will give unbiased results under different assumptions on missingness.

The available cases or cross-sectional analysis refers to a straight-forward analysis, where for each point in time only those cases are considered with non-missing data. This analysis gives unbiased results under MCAR. Differences between treatment groups are tested at each time-point with a two-sample *t*-test. *p*-values are presented without correction for testing at multiple time-points.

The mixed model [9] approach assumes random patient effects and fixed effects of time (as categorical variable) and treatment (and interaction). This implies that average VAS scores differ from patient to patient, but mean changes over time are identical for patients in the same treatment group. The random patient effects induce (positive) correlation between VAS scores of the same patient at different time points. This method is also known as repeated measures analysis of variance. Differences between treatment groups are tested by defining linear combinations of the estimated regression coefficients in the model (contrasts) and testing their significance using Wald's test. A mixed model will give unbiased results under MAR.

Pattern-mixture models [9–11] assume that data are missing at random within a number of prespecified patterns $p = 1, \ldots, P$; in our situation a pattern corresponds to a dropout time. Patients dropping out after baseline will be in pattern 1, patients dropping out after month 3 will be in pattern 2 and so on, and patients with a valid month 24 form will be in pattern $P = 6$. Note that within each pattern $p \neq 1$, patients may have missing forms before dropout. Also, patterns are

defined on the basis of valid forms; occasionally even if the form is valid, the VAS score may have not been filled in or may have been illegible. For a patient in pattern p with observations $y_{ij}, \ldots, y_{ip}$ (some of which may be missing), observations y_{ij} after dropout are imputed from the conditional distribution of y_{ij} given $y_{i1}, \ldots, y_{ip}$ from a higher pattern for which this distribution can be estimated. Pattern-mixture models can be further divided according to which pattern is used to impute missing observations from. Pattern-mixture models are underidentified and hence identifying restrictions need to be made in order to make the model identifiable. Since we are assuming MAR for intermediate missing observations, we only need to impute for patterns $p = 1, \ldots, P-1$. For a patient in pattern p, we only need to impute observations $p+1, \ldots, P$. We consider the following three methods (for more details refer to [9], Section 20.5):

Complete cases missing values (CCMV): All missing observations after dropout are imputed from the conditional distribution of the pattern corresponding to patients with complete data. In our case this is the last pattern, pattern P.

Neighboring cases missing values (NCMV): A missing observation y_{ij} for $j > p$ of patient i is imputed using the conditional distribution of pattern j.

Available cases missing values (ACMV): A missing observation y_{ij} for $j > p$ of patient i is imputed using the conditional distribution of a randomly chosen pattern $(j, \ldots, P)$. The probabilities $\pi_k, k = j, \ldots, P$ of imputing from pattern k are designed in such a way that dropout is missing at random.

Imputation of missing data is repeated a number of times; the completed data are analyzed for each of the patterns separately using mixed models and the estimates are combined afterwards. The variances of estimates within and between imputations are combined (see for instance [9], Section 20.3 for a more detailed description of this multiple imputation method).

Differences between treatment groups for each of the above methods were assessed by defining contrasts and testing their significance using Wald's test.

All analyses in this subsection were done with S-plus 2000 (Insightful Corporation).

Results

Assessment of missingness

A total of 1530 randomized Dutch patients were approached for the TME quality of life substudy. For the present analysis, we excluded 50 patients that were ineligible, 37 patients that never had any surgery, 52 that died in hospital, and 89 that did not consent. Excluding those patients left 1302 patients for the present study. Thirty of those patients never actually had a valid quality of life measurement, 906 completed quality of life followup. Of those, 609 (67%) had filled in each form on time, 143 (16%) only missed their pretreatment form.

For the baseline measurement, a total of 216 forms were received too late. Of those, 76 (6%) were received after start of radiotherapy, 93 (7%) after surgery, and 47 (4%) did not have a date on the form.

Table 1 summarizes baseline VAS scores, depending on whether the corresponding quality of life forms were dated before, at, or after day of randomization, radiotherapy and surgery, respectively. Baseline VAS scores did not differ according to whether forms were dated before, at or after day of randomization, neither for all patients, nor for treatment groups separately. Baseline VAS scores differed according to whether forms were dated before, at, or after day of radiotherapy (RT + TME group only, $p = 0.020$) and day of surgery ($p < 0.001$ for all patients; $p = 0.003$ for the RT + TME group, $p = 0.005$ for the TME group). A post-hoc comparison with Bonferroni correction indicated significant differences of baseline VAS scores between patients who filled in their forms before and after day of radiotherapy ($p = 0.016$) and between patients who filled in their forms before and after day of surgery ($p < 0.001$). We did not find any significant differences between treatment groups, for any of the time-frames (before, at, or after day of randomization, before, at, or after day of surgery). Based on these findings, pretreatment forms dated after start of radiotherapy and at or after surgery were considered missing; pretreatment forms dated at or after day of randomization but before start of radiotherapy and surgery were considered for analysis, resulting in 991 valid baseline forms.

Table 1. Comparison of VAS scores at baseline, according to whether the quality of life forms were filled in before, at or after day of randomization, radiotherapy (RT + TME group only), and surgery

		Total				RT + TME				TME			
		N	Mean	95% CI	*p*-Value	N	Mean	95% CI	*p*-Value	N	Mean	95% CI	*p*-Value
Randomization	Before	26	75.0	67.0–83.0		12	77.3	66.8–87.7		14	73.1	60.0–86.1	
	At	47	75.4	68.4–78.5		25	74.9	68.1–81.8		22	71.7	63.7–79.7	
	After	1068	75.2	74.2–76.2	0.79	526	76.0	74.6–77.3	0.91	542	74.4	72.9–75.9	0.77
Radiotherapy	Before					428	77.0	75.5–78.5					
	At					14	76.9	69.3–84.5					
	After					111	72.2	69.0–75.4	0.020				
Surgery	Before	1047	75.8	74.8–76.8		525	76.5	75.2–77.9		522	75.1	73.5–76.6	
	At	6	64.8	42.4–87.3		1	90.0			5	59.8	35.6–84.0	
	After	88	67.6	63.6–71.7	<0.001	37	67.5	61.1–73.8	0.003	51	67.8	62.4–73.1	0.005

p-Values refer to overall comparison of means with one-way ANOVA. Differences between treatment groups, for the different time-frames (before, at, or after day of randomization, before, at, or after day of surgery), were not significant.

Table 2 summarises how many forms were received too late, at baseline and at each following scheduled time point. Table 2 also shows the compliance at every measurement. Mean compliance was 84%; only at baseline compliance was below 80% (71.1% in the RT + TME group, 81.1% in the TME group; $p < 0.001$), the reason being that many forms were filled out after surgery or after start of radiotherapy. From month 3 on, compliance was consistently somewhat higher (1–2%) in the RT + TME group; this difference was nowhere significant, however.

Table 3 contains the results of the comparison for patients with complete quality of life followup, between the three categories described in section Assessment of missingness. The differences in n-values is explained by the fact that the subgroups are defined with respect to missing forms, not missing VAS scores, see also section Assessment of missingness. For month 3 and month 24, the VAS scores differ significantly between the three categories ($p = 0.034$ both for month 3 and month 24). Post-hoc tests indicate that the categories with complete data and baseline only missing do not differ, and that the 'other' category is lower than the baseline only missing group at month 3 (Bonferroni-corrected $p = 0.045$) and from the complete group at month 24 (Bonferroni-corrected $p = 0.032$). No statistically significant differences were found in either of the treatment groups, nor were there any statistically significant differences between treatment groups.

Comparison of analysis strategies

Table 4 contains the results of the available cases analysis and the mixed model approach. At baseline and at 3 months the estimated means are virtually identical for the two strategies; from month 6 on, the linear mixed model gives slightly lower means, the difference in means between the available cases and the mixed model analyses increases with time. At 24 months, the available cases analysis mean estimates are two points (2.5 standard errors) higher than those of the mixed model. As expected, the available cases analysis estimates are biased upwards, because it is based on a sample of healthier patients, the sick ones having dropped out. The differences between the treatment groups are nevertheless comparable for the two analyses.

Table 2. Assessment of missing forms

	Baseline	Month 3	Month 6	Month 12	Month 18	Month 24	Total
Total alive at previous time	1302	1302	1293	1280	1230	1180	
Dead in interval	0	9	13	50	50	50	172
No of forms possible	1302	1293	1280	1230	1180	1130	7415
Total missing	311	160	157	175	195	224	1222
Not received	95	88	119	138	164	190	
No disease	87	78	94	104	120	140	
LR			4	4	7	4	
DR		2	9	14	20	26	
LR + DR				1	4	8	
R2 or M1 at surgery	8	8	12	13	13	12	
Too late	169	34	5	12	6	14	
After surgery	93						
After RT	76						
Date missing	47	38	33	25	25	20	
Received on time	991	1133	1123	1055	985	906	6193
Compliance (%)	76.1	87.6	87.7	85.8	83.5	80.2	83.5

LR – local recurrence, DR – distant recurrence; R2 – residual tumor, M1 – metastasis, RT – radiotherapy. Note that 991 pre-forms were filled in on time. Of these, 26 were filled in before the day of randomization, as originally required, 47 on the day of randomization itself, and 918 after randomization.

Table 3. Comparison of VAS scores for patients with complete quality of life followup

			Mean	95% CI	*p*-Value
Baseline	Complete	(n = 605)	76.9	(75.6–78.2)	
	Baseline missing	(n = 0)	–	–	
	Other	(n = 105)	76.9	(73.8–79.9)	0.98
Months 3	Complete	(n = 596)	76.1	(74.8–77.4)	
	Baseline missing	(n = 141)	77.1	(74.5–79.7)	
	Other	(n = 94)	71.8	(68.1–75.5)	0.034
Month 6	Complete	(n = 602)	77.5	(76.3–78.7)	
	Baseline missing	(n = 142)	78.4	(75.6–81.2)	
	Other	(n = 115)	76.9	(74.4–79.3)	0.73
Month 12	Complete	(n = 598)	78.8	(77.5–80.1)	
	Baseline missing	(n = 141)	79.5	(76.6–82.3)	
	Other	(n = 110)	75.8	(73.0–78.6)	0.15
Month 18	Complete	(n = 599)	78.9	(77.6–80.1)	
	Baseline missing	(n = 137)	77.2	(74.3–80.1)	
	Other	(n = 98)	77.3	(73.3–81.2)	0.42
Month 24	Complete	(n = 604)	77.9	(76.6–79.2)	
	Baseline missing	(n = 140)	76.3	(73.1–79.5)	
	Other	(n = 147)	73.7	(70.4–77.1)	0.034

Patients are divided into three categories: complete forms, baseline forms missing (forms are complete with the exception of baseline forms), and other. *p*-Values are based on one-way analysis of variance.

For the pattern-mixture model analysis, the percentages of patients in patterns 1 through 6 were 2.5, 3.4, 5.7, 6.4, 10.2 and 71.8, respectively. We calculated the means and 95% confidence intervals of the VAS scores for each of the six patterns. The result is shown graphically in Figure 1.

Table 4. Estimated means (SE) of VAS scores for two approaches of dealing with missing values; available cases and a linear mixed model

	Available cases			Linear mixed model		
	RT + TME	TME	*p*-Value	RT + TME	TME	*p*-Value
Baseline	76.88 (0.74)	74.92 (0.78)	0.070	76.85 (0.79)	74.77 (0.74)	0.054
Month 3	73.29 (0.78)	74.41 (0.72)	0.29	73.28 (0.74)	74.48 (0.73)	0.25
Month 6	76.24 (0.69)	75.66 (0.75)	0.57	75.74 (0.74)	75.34 (0.73)	0.70
Month 12	76.16 (0.80)	77.27 (0.75)	0.32	75.60 (0.75)	76.68 (0.75)	0.31
Month 18	76.90 (0.80)	77.48 (0.80)	0.61	75.82 (0.77)	75.81 (0.77)	0.97
Month 24	77.11 (0.84)	76.96 (0.83)	0.90	75.82 (0.79)	75.03 (0.80)	0.49

p-Values refer to the hypothesis of equality of mean quality of life values for RT + TME and TME only groups.

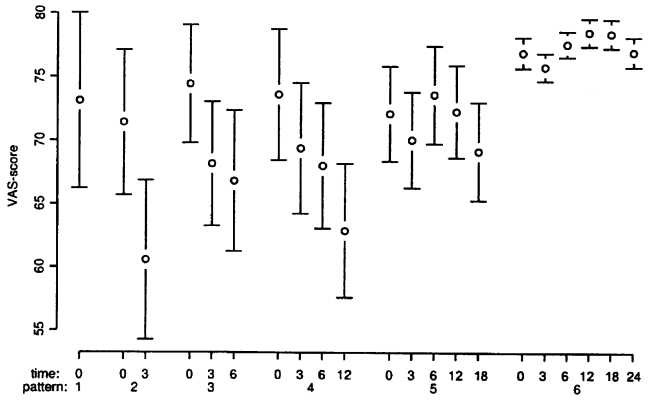

Figure 1. VAS means and error bars (95% confidence intervals) for different patterns.

It clearly shows that mean VAS scores are higher for the patients in pattern 6, strongly indicating a violation of the MCAR assumption. Each of the patterns shows a decrease at 3 months compared to baseline and a decline towards the time of dropout. A continuation of this decline until at or after time of dropout (not observed of course, but not unlikely either) would be a violation of the MAR assumption (i.e., implying MNAR).

Table 5 contains the results of the pattern-mixture model approach under three possible

Table 5. Estimated means (SE) of VAS scores for three strategies of imputing missing values in pattern-mixture models

	CCMV			NCMV			ACMV		
	RT + TME	TME	*p*-Value	RT + TME	TME	*p*-Value	RT + TME	TME	*p*-Value
Baseline	76.86 (0.75)	74.94 (0.70)	0.061	76.90 (0.75)	74.95 (0.70)	0.057	76.89 (0.76)	74.95 (0.71)	0.060
Month 3	73.09 (0.70)	74.70 (0.69)	0.099	73.18 (0.71)	74.49 (0.70)	0.18	73.12 (0.72)	74.65 (0.69)	0.12
Month 6	75.49 (0.70)	75.60 (0.68)	0.91	75.70 (0.82)	75.25 (0.72)	0.66	75.27 (0.71)	75.40 (0.69)	0.90
Month 12	75.17 (0.70)	76.94 (0.69)	0.067	74.73 (0.84)	75.85 (0.72)	0.27	74.74 (0.77)	76.49 (0.71)	0.077
Month 18	75.85 (0.69)	76.19 (0.68)	0.72	75.08 (0.75)	74.95 (0.72)	0.90	75.30 (0.75)	75.44 (0.71)	0.89
Month 24	76.10 (0.70)	75.53 (0.67)	0.55	75.93 (0.71)	75.02 (0.68)	0.34	75.73 (0.72)	75.28 (0.68)	0.64

p-Values refer to the hypothesis of equality of mean quality of life values for RT + TME and TME only groups.

strategies of imputing missing values: CCMV, NCMV, and ACMV.

The three imputation methods give very similar results. No clear pattern emerges in the sense of one method giving higher mean values than the other. Each of the three imputation methods is also very similar to the linear mixed model analysis of Table 4.

Discussion

We have assessed missing values and dropout in the TME quality of life substudy. With regard to baseline forms, we found that, although strictly speaking only forms dated before the day of randomization should be included, there is little harm in including also forms filled in at or even after day of randomization. Apparently, knowing what treatment a patient gets does not influence the VAS score. Of 991 eligible pre-forms, only 26 were filled in before the day of randomization, as originally required, 47 on the day of randomization itself, and 918 after randomization, so adhering to the protocol would leave only 26 or at most 73 patients with valid baseline forms. We have therefore decided to admit 918 baseline forms dated after randomization, but not after radiotherapy or surgery. One would expect radiotherapy and/or surgery to have a larger impact on quality of life scores. This was also confirmed by our findings, which clearly shows lower VAS scores at or after day of radiotherapy/surgery as compared to before day of radiotherapy/surgery.

We discussed a hierarchy of types of missingness: missing completely at random (MCAR), missing at random (MAR) and missing not at random (MNAR). An error-bar plot of VAS scores for each of six patterns separately (patterns defined by time of dropout) clearly indicated that the MCAR assumption was violated and even suggested the possibility of violation of MAR. Patients with complete quality of life followup had higher VAS scores; early dropout was associated with lower quality of life. The distinction between MAR and MNAR is intrinsically more difficult; in fact testing violation of MAR requires assumptions (to make the model identifiable) which are themselves untestable. Differences between estimates obtained from statistical models assuming MAR and MNAR could indicate violation of MAR, but could also be caused by an incorrect choice of identifying restrictions. To assess the influence of these identifying restrictions we fitted three models with different identifying restrictions: complete cases missing values (CCMV), neighboring cases missing values (NCMV) and available cases missing values (ACMV). This sensitivity analysis revealed no important differences between the estimates obtained from these models. Moreover, each of those methods gave very similar results to the linear mixed model approach, which is valid under the MAR assumption. Note that this does not necessarily mean that data are missing at random. In fact, Figure 1 is somewhat suggestive of violation of MAR, since one could well imagine the decline in, e.g., pattern 2 to continue past month 3. However, dropout rates were too low (pattern 2, e.g., constituting only 3.4% of all patients) for the pattern-mixture models (representing only moderate deviations from MAR) to pick up an effect of any violation of MAR on VAS scores. Since pattern-mixture models are computationally considerably involved, for this quality of life analysis the linear mixed model approach was found to be an attractive method of analysis yielding accurate estimates of mean VAS scores and differences between treatment groups.

Even if different statistical models yield different estimates of quality of life, when the main objective is to study differences in quality of life between two or more treatments, differences in quality of life estimates may not be a problem as long as estimates of differences in quality of life remain similar across different models. Differences in quality of life estimates but not in quality of life treatment differences may occur when dropout is considerable (and probably related to disease progression and quality of life) but comparable for the different treatments. For this quality of life substudy of the TME study there is a priori reason to suspect that dropout due to local recurrence is higher in the TME only arm. In fact, this expectation was one of the incentives to undertake this comparative analysis.

A number of other papers ([9], Section 20.6, [12]) have shown substantial differences between the quality of life estimates based on different choices of statistical models. Those were studies concerned with diseases in advanced stages like

metastatic breast cancer and metastatic lung cancer, based on patient populations with high mortality and high disease progression, resulting in high percentage of missing values. The fact that we did not find major differences in our data between the results from models assuming MAR and MNAR could be very well explained by the fact that in our study morbidity and mortality were low and the percentage of data missing was at most moderate. A number of recent studies on quality of life in patients with early-stage Hodgkin's disease [13] and non-small-cell lung cancer [14] with comparable or even higher degree of missingness also reported similar results under MAR and MNAR. Given this fact, it is all the more striking that the estimates obtained from the available cases analysis were so different. This is a clear warning to interpret results based on available cases analyses with considerable care.

Acknowledgement

Part of this research was funded by the Dutch Digestive Diseases Foundation (MLDS).

References

1. Bernhard J, Cella DF, Coates AS, et al. Missing quality of life data in cancer clinical trials: Serious problems and challenges. Stat Med 1998; 17: 517–532.
2. Heald RJ, Ryall RD. Recurrence and survival after total mesorectal excision for rectal cancer. Lancet 1986; 1: 1479–1492.
3. Enker WE, Havenga K, Polyak T, Thaler H, Cranor M. Abdominoperineal resection via total mesorectal excision and autonomic nerve preservation for low rectal cancer. World J Surg 1997; 21: 715–720.
4. Gray R, Hills R, Stowe R, et al. Adjuvant radiotherapy for rectal cancer: A systematic overview of 8507 patients from 22 randomised trials. Lancet 2001; 358: 1291–1304.
5. Kapiteijn E, Marijnen CAM, Nagtegaal ID, et al. Preoperative radiotherapy combined with total mesorectal excision for resectable rectal cancer. N Engl J Med 2001; 345: 638–646.
6. De Haes JCJM, Van Knippenberg FCE, Neijt JP. Measuring psychological and physical distress in cancer patients – Structure and application of the Rotterdam Symptom Checklist. Br J Cancer 1990; 62: 1034–1038.
7. Rubin DB. Inference and missing data. Biometrika 1976; 55: 581–592.
8. Curran D, Bacchi M, Schmitz SFH, Molenberghs G, Sylvester RJ. Identifying the types of missingness in quality of life data from clinical trials. Stat Med 1998; 17: 739–756.
9. Verbeke G, Molenberghs G. Linear Mixed Models for Longitudinal Data. Springer Series in Statistics. New York: Springer, New York, 2000.
10. Little RJA. Pattern-mixture models for multivariate incomplete data. J Amer Stat Assoc 1993; 88: 125–134.
11. Hogan JW, Laird NM. Mixture models for the joint distribution of repeated measures and event times. Stat Med 1997; 16: 239–257.
12. Fairclough DL, Peterson HF, Cella D, Bonomi P. Comparison of several model-based methods for analysing incomplete quality of life data in cancer clinical trials. Stat Med 1998; 17: 781–796.
13. Ganz PA, Moinpour CM, Pauler DK, et al. Health status and quality of life in patients with early-stage Hodgkin's disease treated on Southwest Oncology Group study 9133. J Clin Oncol 2003; 21: 3512–3519.
14. Moinpour CM, Lyons B, Grevstad PK, et al. Quality of life in advanced non-small-cell lung cancer: Results of a Southwest Oncology Group randomized trial. Qual Life Res 2002; 11: 115–126.

Address for correspondence: H. Putter, Department of Medical Statistics, Leiden University Medical Centre, P.O. Box 9604, 2300 RC Leiden, The Netherlands
Phone: +31-71-5276827; Fax: +31-71-5276799
E-mail: h.putter@lumc.nl