

University of Groningen

The analysis of longitudinal quality of life measures with informative drop-out

Post, W.J.; Buijs, C.; Stolk, R.P.; de Vries, E.G.E.; le Cessie, S.

Published in:
Quality of Life Research

DOI:
[10.1007/s11136-009-9564-1](https://doi.org/10.1007/s11136-009-9564-1)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2010

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Post, W. J., Buijs, C., Stolk, R. P., de Vries, E. G. E., & le Cessie, S. (2010). The analysis of longitudinal quality of life measures with informative drop-out: A pattern mixture approach. *Quality of Life Research*, 19(1), 137-148. <https://doi.org/10.1007/s11136-009-9564-1>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

The analysis of longitudinal quality of life measures with informative drop-out: a pattern mixture approach

Wendy J. Post · Ciska Buijs · Ronald P. Stolk ·
Elisabeth G. E. de Vries · Saskia le Cessie

Accepted: 30 November 2009 / Published online: 30 December 2009
© The Author(s) 2009. This article is published with open access at Springerlink.com

Abstract

Purpose The analysis of longitudinal health-related quality of life measures (HRQOL) can be seriously hampered due to informative drop-out. Random effects models assume Missing At Random and do not take into account informative drop-out. We therefore aim to correct the bias due to informative drop-out.

Methods Analyses of data from a trial comparing standard-dose and high-dose chemotherapy for patients with breast cancer with respect to long-term impact on HRQOL will serve as illustration. The subscale Physical Function (PF) of the SF36 will be used. A pattern mixture approach is proposed to account for informative drop-out. Patterns are defined based on events related to HRQOL, such as death and relapse. The results of this pattern mixture approach are compared to the results of the commonly used random effects model.

Results The findings of the pattern mixture approach are well interpretable, and different courses over time in

different patterns are distinguished. In terms of estimated differences between standard dose and high dose, the results of both approaches are slightly different, but have no consequences for the clinical evaluation of both doses. **Conclusion** Under the assumption that drop-out is at random within the patterns, the pattern mixture approach adjusts the estimates to a certain degree. This approach accounts in a relatively simple way for informative drop-out.

Keywords Health-related quality of life · Informative drop-out · Longitudinal studies · Missing values · Pattern mixture model

Introduction

Treatment for cancer can affect patient quality of life [1, 2]. Therefore, health-related quality of life (HRQOL) measures are frequently included in randomized clinical trials in oncology. The interpretation of results for HRQOL within these randomized trials can be seriously hampered by drop-out due to relapse and death, especially because these events do affect HRQOL.

In longitudinal studies, observations of patients can be missed at a certain time point because they miss visits or do not fill in certain questionnaires, although they still participate in the study. Consequently, these patients have responses on a subset of outcome measures at different time points. Alternatively patients can drop-out, i.e., not participating in the study after a certain moment. These patients are lost to follow-up, and no information is available. In our study, we define drop-out as not filling in the HRQOL questionnaire after a certain time point, while there could be information about the survival status after that time.

W. J. Post (✉) · R. P. Stolk
Department of Epidemiology, University Medical Center
Groningen, Hanzeplein 1, P.O. 30001, 9700 RB Groningen,
The Netherlands
e-mail: w.j.post@rug.nl

C. Buijs · E. G. E. de Vries
Department of Medical Oncology, University Medical Center
Groningen, Groningen, The Netherlands

S. le Cessie
Department of Medical Statistics and Bioinformatics, Leiden
University Medical Center, Leiden, The Netherlands

S. le Cessie
Department of Clinical Epidemiology, Leiden University
Medical Center, Leiden, The Netherlands

There exist different reasons for drop-out. Except for drop-out caused by relapse or death, drop-out occurs due to censoring, administrative, or person-related reasons. The reasons for drop-out are often referred to as the drop-out mechanism [3]. It is important to distinguish different types of drop-out mechanisms, because each drop-out mechanism requires different analyses to get valid results. A common classification of drop-out mechanisms has been introduced by Rubin and can be summarized as follows: Missing Completely At Random (MCAR), Missing At Random (MAR), and Missing Not At Random (MNAR) [4].

Drop-out can be considered MCAR if the probability that a patient drops out is independent of the observed (past) response measurements and of current and future response measurements. Or, to put in other words: patients drop-out due to completely unrelated reasons. For example, when HRQOL data are lost due to administrative failure, the drop-out can be considered MCAR: the drop-out is unrelated to HRQOL scores, and therefore does not depend on the observed nor on unobserved HRQOL scores. Under the assumption of MCAR, complete case analysis (including only those patients with complete data) will give valid results.

Drop-out can be considered MAR, if the probability of drop-out depends on previously observed response measures, but not on the current and future measures. However, given the observed response measures, drop-out is assumed to be independent of the unobserved (current and future) response measures. For example, when a patient drops out because the HRQOL score is lower at earlier measurements, the drop-out can be considered MAR. In this situation, one is able to predict the measurement after drop-out based on past measurements, and random effects models (mixed models) would provide unbiased results if the model is correctly specified. In many situations, this drop-out mechanism will be present in the data.

Drop-out is missing not at random (MNAR) if the probability of drop-out is also dependent on the unobserved (current and future) response measures. In this situation, the missing measurement itself is informative. For example, when a patient has a sudden decrease in HRQOL and decides not to fill in the questionnaire, the drop-out is MNAR. In this situation, we can not predict the missing HRQOL of this patient without making further (untestable) assumptions. The reason of a sudden decrease could be the occurrence of tumor relapse, or the fact that a patient is dying. So, drop-out due to tumor relapse and death might be informative.

In trials, it is quite common to analyze longitudinal data with random effects models. This type of models can handle missing data and give valid results if MAR can be assumed. In these models, the patients' courses over time are specified using regression equations per patient, taking the correlation between the different measurements within

patients into account. Therefore, a random effects model helps understanding about how individual patients change across time [5]. However, without modeling, the drop-out mechanism informative drop-out is not taken into account. In the analysis of a trial, this might lead to serious bias, and therefore to wrong conclusions about HRQOL. To understand this, suppose that in a clinical trial, the experimental arm reduces or postpones an event as tumor relapse. Suppose, additionally, that the experimental arm has a larger (negative) impact on HRQOL than the other arm. If patients are less inclined to fill in HRQOL questionnaires after a relapse, then patients in the experimental arm are more likely to fill in the HRQOL questionnaires than patients in the control arm. So, more patients with lower HRQOL will remain in the experimental arm compared to the control arm. In this case, complete case analysis will yield biased results in favor of the standard arm, because patients with missing observations after relapse are not taken into account. A similar argument holds when using last observation carried forward, since the high HRQOL values before relapse are imputed for the missing HRQOL values after death and relapse. It is therefore important to model implicitly or explicitly the drop-out mechanism.

There are several approaches for the analysis of informative drop-out. They are rather complicated. Therefore, their complexity may be the reason why these approaches are not routinely employed in the analysis of randomized clinical trials.

Historically, there are two different likelihood-based approaches for analyses of informative drop-out that differ with respect to the way the joint distribution of responses and drop-out process is factorized (Curran et al. [6]; Michiels et al. [7]). These approaches are the pattern mixture model approach (Hedeker and Gibbons [8], Pauler et al. [9]) and the selection model approach (Diggle and Kenward [10], Little [11], and Curran et al. [12]). In the present paper, we will focus on the pattern mixture approach, because this is a relatively simple extension of the commonly used random effects model assuming MAR.

Another class of models is based on the joint modeling of response measures (e.g., HRQOL) and times to events (e.g., relapse and death). These models estimate the HRQOL measures given that a person is in a certain state at a certain time point. Examples are the models of Kurland and Heagerty [13], and of Diggle et al. [14]. These models take the ordering of event times explicitly into account and regard drop-out as a consequence of events in the past. In this sense, these models can be used for prediction. This is in contrast to the pattern mixture approach where the drop-out process is reflected in patterns and where these patterns are determined retrospectively. This approach is therefore primary appropriate for adjusting the estimates for the missing data mechanism.

In this paper, we will use data of a Dutch multi-center randomized clinical trial in patients with breast cancer as illustration for the use of the pattern mixture approach. In this trial, the long-term impact of two different chemotherapy schedules (high dose vs. standard dose) on HRQOL was compared. The results of the study with respect to the primary outcomes, relapse-free survival, and overall survival have been published by Rodenhuis et al. [15, 16]. Buijs et al. [17] analyzed the HRQOL data comparing both treatment arms assuming that patients remained disease free.

We aimed in this paper to correct the HRQOL estimation for informative drop-out in order to get a valid comparison between schedules, taking events as death and relapse into account. Since we expect that drop-out due to relapse and death are the two main sources of informative drop-out, we propose a pattern mixture approach in which patterns are defined based on the different disease states a person can have at the end of the study period. The different states we use for the patterns are quality of life related such as ‘death’, ‘alive with tumor relapse’, and ‘disease free’. We start considering the ‘usual’ random effects model analysis assuming MAR. Then, we will analyse the data with the pattern mixture approach with HRQOL-related patterns. Thereafter, outcomes of both approaches will be compared.

Methods

Dutch multi-center randomized clinical breast cancer trial

From August 1993 to July 1999, 804 patients from 10 Dutch centers were randomized in the HRQL part of the trial comparing two different chemotherapy schedules. The standard chemotherapy arm consisted of five cycles of 5-fluorouracil (500 mg/m²), epirubicin (90 mg/m²), and cyclophosphamide (500 mg/m²) (FEC). The high-dose chemotherapy arm consisted of 4 identical cycles of FEC followed by one cycle of high-dose chemotherapy comprising cyclophosphamide (6 g/m²), thiotepa (480 mg/m²), and carboplatin (1600 mg/m²) administered over 4 days followed by peripheral blood progenitor cell reinfusion on day 7. There were 885 patients randomized; 804 of them participated in the HRQOL part of the trial of which 405 patients were randomly assigned to the standard therapy and 399 patients to the experimental (high dose) therapy. The main study was a survival study. The primary outcomes were the overall survival and the disease-free survival. Secondary outcome measures were health-related quality of life measures. The Medical Ethical Committees of all participating centers approved the study. All patients gave informed consent before study entry.

Subgroup of patients

At the time of the analysis in 2003, there appeared to be no significant differences following intention to treat analyses with respect to overall and disease-free survival (Rodenhuis et al. [15]). However, a subgroup of patients with normal HER2 expression did benefit of high-dose chemotherapy as far as disease free ($P = 0.002$) and overall survival ($P = 0.02$) were concerned (Rodenhuis et al. [16]). In this paper, we focus therefore on this subgroup. Only patients participating in the HRQOL part of the trial are considered, of which 273 patients were randomized to standard dose and 288 to high-dose chemotherapy.

Quality of life assessment

The pattern mixture model is illustrated for one of the subscales of the SF-36 namely the physical functioning (PF) subscale, consisting of items which indicate the ability to perform physical activities as walking, carrying shopping, and climbing stairs. The scores range from 0 to 100, with higher scores representing higher level of functioning. The choice of this subscale was mainly based on illustrative reasons: preliminary analyses on this subscale demonstrated significant and relative large differences in favor of the standard arm. Other subscales demonstrated smaller (and/or not significant) differences.

The questionnaires were sent by mail before randomization (baseline), after chemotherapy (about 3 months), after radiotherapy (about 6 months), and thereafter every 6 months.

Due to the variation in entry date, not all patients could be studied for the maximum period of 5 years and 12 measurements points. Each patient had a study period follow-up of at least 3 years and eight measurement points. We defined patients who were alive but did not have the maximum study period follow-up as censored.

Statistical analysis

The analyses were performed with SPSS (version 16), MLWin (version 2.02), and R (version 2.6.2).

Descriptive

Analyses were based on intention to treat. A cross-sectional comparison of PF between the two doses was performed at each measurement point, presenting the observed mean and standard error per measurement point. In this analysis, differences in PF per time point between both doses were based on different sets of patients. Visual inspection of the course of the means over time gave suggestions for modeling the course over time in the random effect model and pattern mixture model.

Drop-out

A large part of drop-out was due to death and censoring (e.g., patients who did not have the maximum study period follow-up). Patients able to fill in the questionnaires at a certain time point (the patients at risk) are the patients being alive and being not censored at that time point. Percentages of patients with observed HRQOL measures (i.e., responding patients) are based on these patients at risk, and are called response rates. They were calculated for each time point and for each treatment.

A distinction in drop-out due to relapse and due to death must be made. In the case of relapse, we can assume that PF exists and is affected by a relapse, but that it is not measured. The definition of PF after death is not directly obvious. Therefore, three patterns were defined: A pattern consisting of patients who died during the study period (deceased patients), a pattern with patients with relapse who were alive at the end of the study period, and a pattern with patients who remained disease free. A cross-sectional comparison of the PF assessments of these patterns was performed to analyze descriptively whether the course over time was different for the patterns or not, and whether both doses had different responses over time per pattern.

Random effects model (assuming MAR)

A random effects model was fitted on the data with ‘course over time’, ‘treatment’ and ‘interaction course over time by treatment’ as fixed explanatory variables for the PF assuming MAR. In the following, this is called the “final model”. Since we assumed that each patient had her own course over time, time effects were also considered to be random. The ‘course over time’ was specified based on the findings of the visual inspection of the descriptive analysis. *P*-values smaller than 5% were considered to be significant.

We started with fitting the empty model (Snijders and Bosker) [18], using maximum likelihood. This was a model without explanatory variables; only a fixed intercept was included (parameter for the overall mean) with two random effects. This model estimated the overall mean PF of all patients and time points and estimates the variability of the PF within patients and between patients. Then, we fitted the time model using maximum likelihood. This model included only the fixed and random intercept and fixed and random time effects. This model estimated the mean PF of all patients over time. The difference in deviance (i.e., -2 times the value of the likelihood) between the empty model and time model is an indication for the effect of time on PF, and the deviances can be compared using the likelihood ratio test. Then, the final model ‘interaction course over time by treatment’ was fitted by maximum likelihood. This model estimated the mean PF over time for the different treatment

arms. The performance of the final model was evaluated by the deviance compared to the deviance of the empty model and to the deviance of the time model. The difference between the final model and the time model is an indication of the effect of treatment over time on PF. The comparisons of different deviances between different nested models can be regarded as a relative for the fit of the model for the data.

These models assumed only MAR implying that drop-out dependent on the missing PF score itself (for example because of an unpredictable decrease caused by relapse) was not taken into account. The empty model, the time model, and final model are explained and specified in the appendix.

Pattern mixture approach (correction for informative drop-out)

In the pattern mixture approach, the drop-out process and the outcome measures conditional on the drop-out pattern are jointly modeled. The drop-out process was modeled by the probability to belong to a specific drop-out pattern for each dose separately, and estimated by the sample proportion. Patterns were based on the relapse and survival state of the patient at the end of the study period and defined as before: ‘deceased’, ‘alive with relapse’, and ‘disease free’. The outcome conditional on the drop-out pattern was modeled by the random effects model with additional explanatory variables indicating different follow-up patterns of patients. In the following, we call this model, the pattern mixture model. We implicitly assumed that each pattern has its own missing data process and own course over time with respect to PF. Moreover, we assumed that within each pattern, the drop-out is MAR. This means that conditional on being in a certain pattern, the missing PF can be predicted based on observed measurements of patients within this pattern. The pattern mixture model was specified with ‘course over time’, ‘treatment’, and ‘interaction course over time by treatment’ as explanatory variables (as before), and also with the variables ‘pattern’, and all possible interactions with pattern. This model was fitted using maximum likelihood. See the [Appendix](#) for the explanation and specification of the model.

Based on the difference in deviances (likelihood ratio test), it was decided which additional explanatory terms could be excluded from the model. This difference can be regarded as a relative measure for the fit of the model on the data.

For each pattern, the course over time per treatment was estimated. In order to compare both treatments for all patterns together, the effects of both doses were estimated by a weighted average over all patterns weighted by their sample proportions. An illustration of this is given in the appendix. With the delta-method (Bisshop et al. [19]), standard errors of the differences between the two treatments were estimated to determine the significance of the effects.

Comparison of the results of the random effects model and pattern mixture approach

The fit of the random effects model and the random effects part of the pattern mixture approach (i.e., the pattern mixture model) were compared based on the difference in deviance (likelihood ratio test). The results of the pattern mixture approach (weighting over all patterns) in terms of estimated differences in PF scores between both treatments arms per time point were compared to those of the random effects model assuming MAR.

Sensitivity analysis

Since in our data, not all patients had the maximum study period of 5 years, the patterns ‘alive with relapse’ and ‘disease free’ consisted of patients for which the disease state after 5 years was not known yet. Therefore, we implicitly assumed that censored patients and patients with maximum study period could be combined within each pattern and that this did not have consequences for the results. To check whether the definition of patterns for all patients (with complete and incomplete follow-up) leads to different results than the definition of patterns for only patients with maximum study period follow-up, we performed a sensitivity analysis by also modeling the data using only patients with complete 5-year follow-up and comparing the results.

Results

Descriptive

In Fig. 1, the observed means and standard errors of PF per arm and per measurement point are presented. The number of patients that contributes to each mean is in Table 1, in the column named ‘response’. For the high-dose arm, there was a

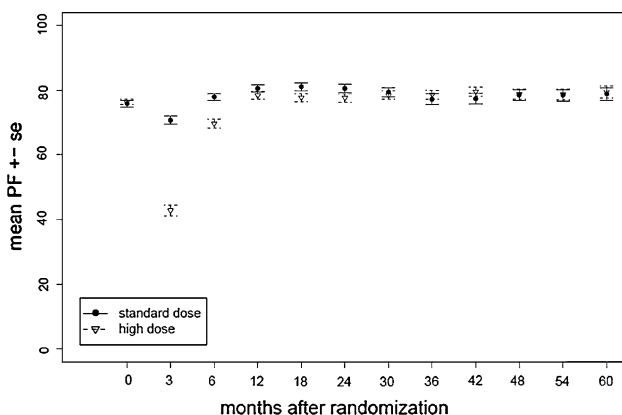


Fig. 1 Observed mean and standard error per treatment and per time point for all available PF measures

large drop in the observed mean after 3 months (just after chemotherapy). There was also a decrease in observed mean for the standard dose although much smaller. After 6 months, the means were increasing again in both treatments. After 1 year, the means in both arms were similar, and remain rather constant, and they were higher than at baseline.

Drop-out

In Table 1, the number of patients at risk (i.e., patients being alive and being not censored), the number of deceased, and the number of censored patients (with incomplete study period follow-up) are presented for each time point per chemotherapy, with also the number and percentages of observed PF scores among all patients at risk (response rates). For both schedules, the percentages were comparable for each measurement point.

In Table 2, the numbers of patients for the different patterns are presented. The majority of the patients belonged to pattern 3: ‘disease free’. About 20% died within 5 years. In Fig. 2a, the different observed courses of PF over time in terms of observed means and standard errors are given per pattern. There were small differences between the patterns during the first year, but there was an obvious difference in the observed course over time after 1 year: the observed mean of patients in the pattern ‘deceased patients’ decreased, the mean PF of patients in the pattern ‘alive with relapse’ also decreased but less steep, and the mean PF in the pattern relapse free was rather constant. In Fig. 2b–d, the observed differences between mean PF over time between the doses are presented for each pattern. During the first year, the observed course over time was similar for all patterns: a large drop at 3 months for the high dose, and a somewhat smaller drop for the standard dose. However, after 1 year, the differences between doses were not similar for the patterns. In the patterns ‘deceased’ and ‘alive with relapse’, the observed course over time suggests that patients who get recurrence or die within the study period benefitted somewhat from the high dose in the long run. In the ‘relapse free’ pattern, the observed means of both doses were similar after 1 year with a negligible difference in favor of the standard dose.

Random effects model (assuming MAR)

Based on the findings of the descriptive analysis, the course over time was modeled by four time variables, namely three dummy time variables t_0 , t_1 , t_2 on the first three time points indicating the measurements at baseline, after 3 months, after 6 months, and one variable t_c indicating the time continuously after 1 year. At each measurement point, there could be differences between both doses. Therefore, the main effect ‘dose’ was included in the model, together with the interaction terms of dose with the

Table 1 The response rates per follow-up time and per treatment arm, by number at risk, number deceased, and censored patients

Follow-up time	Standard-dose arm number			High-dose arm number				
	At risk	deceased	censored ^a	Response (%) ^b	At risk	deceased	censored ^a	Response (%) ^b
Baseline	273	0	0	251 (92)	288	0	0	271 (94)
3 months	273	0	0	244 (89)	288	0	0	257 (89)
6 months	273	0	0	248 (91)	285	3	0	250 (88)
1 year	264	9	0	242 (92)	282	6	0	251 (89)
1.5 years	251	22	0	222 (88)	279	9	0	242 (87)
2 years	248	25	0	204 (82)	268	20	0	227 (85)
2.5 years	238	35	0	195 (82)	262	26	0	213 (81)
3 years	218	41	14	187 (86)	243	30	15	205 (85)
3.5 years	183	52	38	167 (92)	216	38	34	192 (89)
4 years	158	57	58	143 (91)	187	45	56	168 (90)
4.5 years	140	63	70	128 (91)	165	49	74	143 (87)
5 years	130	68	75	102 (78)	155	53	80	127 (82)

^a Number of patients alive with incomplete follow-up due to late inclusion date

^b Number of patients with observed PF score among all patients at risk

Table 2 Number of patients per arm by survival and relapse state

Pattern	Arm		
	Standard dose	High dose	Total
Deceased	68	53	121
Alive with relapse	39	28	67
Disease free	166	207	373
Total	273	288	561

four time variables. The intercept can be considered as the estimated PF at 1 year for the standard dose. The main effect ‘dose’ can be considered as the estimated difference between both doses at 1 year. The sum of the main dose effect and the interaction effect per time dummy is the difference between doses for each time point (baseline, after 3 months, and after 6 months). The interaction effect with the time variable indicating the time continuously after 1 year indicates the difference between doses in

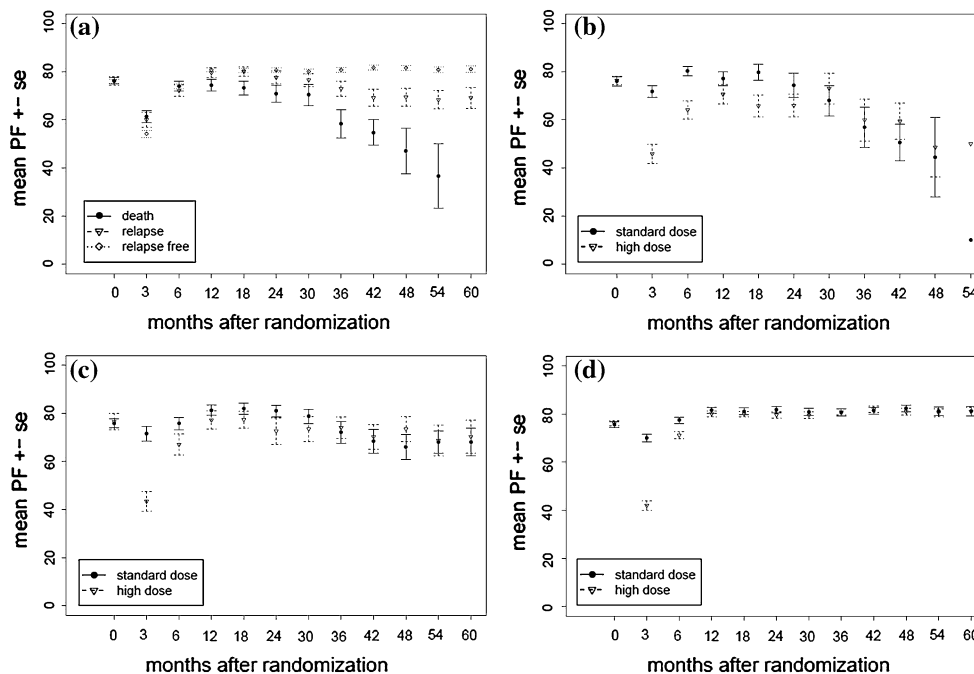


Fig. 2 **a** Observed means and standard errors per time point per pattern. **b** Observed means and standard errors per arm and per time point for deceased patients. **c** Observed means and standard errors per

arm and per time point for patients alive with relapse. **d** Observed means and standard errors per arm and per time point for relapse-free patients

Table 3 The fitted empty model, time model, and random effects model; model estimates and standard errors

Variable	Estimate (standard error)		
	Empty model	Time model	Final model
Fixed effects			
Intercept	74.79 (0.62)	79.35 (0.73)	84.21 (2.32)
t0 ^a		-3.42 (0.76)	-5.41 (1.08)
t1 ^b		-23.23 (1.19)	-10.53 (1.51)
t2 ^c		-5.72 (0.70)	-3.30 (0.99)
tc ^d		-0.13 (0.025)	-0.18 (0.04)
dose * t0 ^a			3.92 (1.51)
dose * t1 ^b			-24.67 (2.11)
dose * t2 ^c			-4.84 (1.378)
dose			-3.24 (1.46)
dose * tc ^d			0.10 (0.05)
Random part (variances level 2)			
Intercept	176.44 (12.66)	230.43 (17.62)	227.74 (17.45)
t0 ^a		130.08 (19.24)	125.53 (18.95)
t1 ^b		540.11 (45.27)	394.17 (36.05)
t2 ^c		75.64 (16.02)	69.82 (15.66)
tc ^d		0.18 (0.02)	0.18 (0.02)
Residual variance			
	260.13 (5.59)	109.98 (3.16)	109.94 (3.32)
Deviance			
	42,022	40,047	39,869
		<i>P</i> -value < 0.0001	

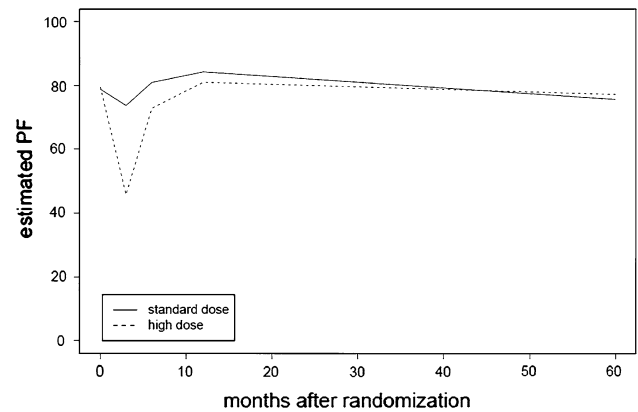
^a Time dummy indicating time point at baseline

^b Time dummy indicating time point at 3 months (just after chemotherapy)

^c Time dummy indicating time point at 6 months

^d Time variable continuously after 1 year

course over time after 1 year. (For the specification of the model, see Appendix). In Table 3, the fitted final model assuming MAR is presented, together with the empty model and the time model. The deviance of the final random effects model (i.e., 39,869) was much smaller than the deviance of the empty model (i.e., 42,022) and the time model (i.e., 40,047), indicating that the dose had a significant effect over time on the PF under the assumption of MAR ($P < 0.001$). In Fig. 3, the estimated mean PF scores for both doses based on this model are presented. In Table 4, the estimated differences between doses per time point (at baseline, after 3 months, and after 6 months) and the estimated difference in slope after 1 year are presented, together with their standard errors and *P*-values. Just after chemotherapy, after 3 months up to 1 year, there was a significant difference for PF between both arms in favor of the standard dose. After 1 year, the slope for the high-dose arm was nearly significant larger ($P = 0.055$) than the standard-dose arm.

**Fig. 3** The estimated course of PF over time per treatment based on the final random effects model assuming MAR**Table 4** The estimated difference between arms for the first time points t0 to t3 and for the slope after 1 year based on the random effects model

Time point	Estimate (standard error)	<i>P</i> -value
t0	0.68 (1.23)	0.59
t1	-27.90 (1.97)	<0.001
t2	-8.08 (1.60)	<0.001
t3	-3.24 (1.46)	0.026
After t3	0.10 (0.05)	0.055

Pattern mixture approach (correction for informative drop-out)

The fit of the pattern mixture model

In Table 5, the fitted pattern mixture model is presented. This model consisted of all explanatory variables of the final random effects model assuming MAR together with all possible interactions with ‘pattern’. Since ‘pattern’ has three categories, two dummy variables were defined indicating the pattern “relapse” and the pattern “relapse free”, respectively; the reference category is ‘deceased’. Based on the small difference in deviance ($P = 0.48$), the main effect ‘pattern’ was excluded from the model.

The estimates of the interactions between the time variables during the first year and the patterns are negative, suggesting that for the standard dose, the pattern “relapse” and the pattern “relapse free” are worse in terms of PF during the first year than the pattern “deceased”. However, these estimates are not significant.

Courses per pattern

In Fig. 4a–c, the estimated course over time for both treatments is given per pattern. It is evident that the results confirmed the findings of the descriptive analysis: the courses over time were very different for the patterns. The differences

Table 5 The fitted pattern mixture model; model estimates and standard errors

Variable	Estimate (standard error)	
Fixed effects		
Intercept	81.51 (3.32)	
t0 ^a	−5.70 (1.84)	
t1 ^b	−9.74 (2.80)	
t2 ^c	−1.53 (1.96)	
tc ^d	−1.32 (0.11)	
dose * t0 ^a	11.54 (3.22)	
dose * t1 ^b	−14.44 (4.59)	
dose * t2 ^c	−4.40 (3.19)	
dose	−11.06 (2.79)	
dose * tc ^d	0.44 (0.17)	
Interaction with patterns		
	With relapse	Relapse free
t0 ^a	−0.80 (2.76)	−0.27 (1.99)
t1 ^b	−2.10 (4.48)	−1.74 (3.24)
t2 ^c	−5.95 (3.07)	−2.32 (2.23)
tc ^d	0.84 (0.13)	1.32 (0.11)
dose * t0 ^a	−6.21 (5.16)	−8.96 (3.53)
dose * t1 ^b	−9.84 (7.38)	−11.84 (5.16)
dose * t2 ^c	0.09 (4.96)	−0.03 (3.54)
dose	8.03 (4.12)	9.04 (2.84)
dose * tc ^d	−0.30 (0.20)	−0.43 (0.17)
Random part (variances level 2)		
Intercept	225.06 (17.19)	
t0 ^a	125.21 (18.73)	
t1 ^b	378.90 (34.90)	
t2 ^c	67.16 (15.29)	
tc ^d	0.09 (0.01)	
Residual variance	108.51 (3.11)	
Deviance	39,626	
	<i>P</i> -value < 0.0001	

^a Time dummy indicating time point at baseline

^b Time dummy indicating time point at 3 months (just after chemotherapy)

^c Time dummy indicating time point at 6 months

^d Time variable continuous after 1 year

between treatments within each pattern were also in line with the descriptive analysis: during the first year, the estimated PF scores for the standard dose were larger than those for the high dose. After 1 year, the difference in slope between the high-dose arm and standard-dose arm for the patterns ‘deceased’ and ‘alive with relapse’ was estimated in favor of the high dose. For the pattern ‘relapse free’, the estimated slopes were similar for both patterns, indicating that the small difference in favor of the standard dose remained after 1 year.

All patterns together

To compare the schedules for all patterns together (weighted over all patterns), the estimated PF for both doses is shown in Fig. 5. In Table 6, the differences per time point (at baseline, after 3 months, and after 6 months) and the difference in slope after 1 year together with their standard errors are presented. During the first year, there was a significant difference in favor of the standard-dose

arm. After 1 year, the slopes differed significantly in favor of the high dose ($P = 0.001$). However, in terms of absolute differences in PF between both doses, the differences are very small (see Fig. 5).

Comparison results of both approaches

Comparing the deviance of the pattern mixture model to the final random effects model assuming MAR, we see that the fit was significantly better since the difference in deviance is large (39,626 vs. 39,869); $P < 0.0001$. This indicates that under the assumption of MAR within each pattern, the pattern mixture model has a better fit than the random effects model.

Comparing the results of the random effects model assuming MAR to the results of pattern mixture approach (weighting over all patterns), see Tables 4 and 6, we see that the estimated differences between both doses were similar for both approaches. However, the estimated difference in slope

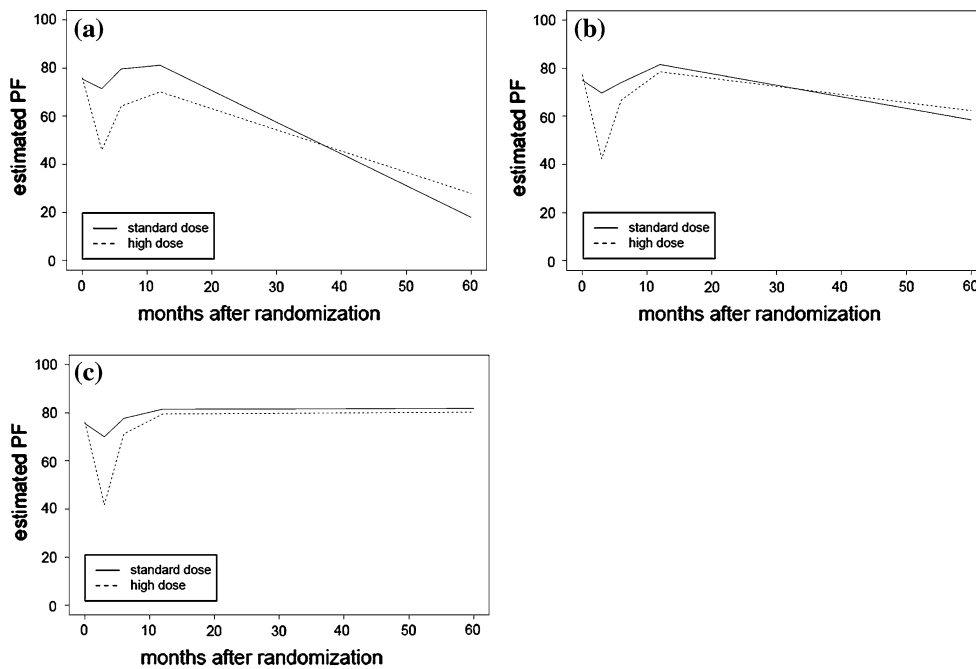


Fig. 4 **a** The estimated course of PF over time per treatment for pattern ‘deceased’. **b** The estimated course of PF over time per treatment for pattern ‘alive with relapse’. **c** The estimated course of PF over time per treatment for pattern ‘disease free’

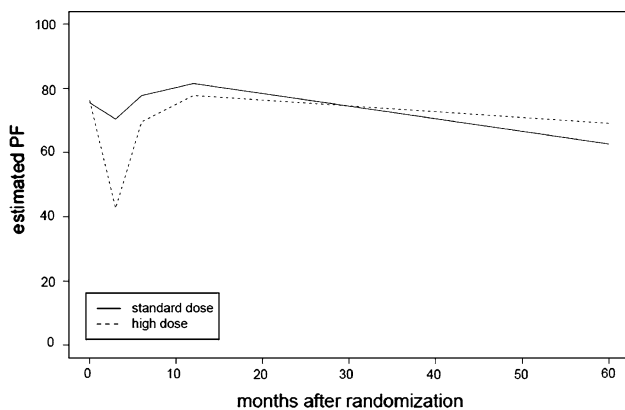


Fig. 5 The estimated course of PF over time per treatment based on the pattern mixture model

Table 6 The estimated difference between arms for the first time points t0 to t3 and for the slope after 1 year based on the pattern mixture model

Time point	Estimate (standard error)	P-value
t0	0.72 (1.26)	0.55
t1	−27.81 (2.01)	<0.001
t2	−8.17 (1.63)	<0.001
t3	−3.75 (1.47)	0.003
After t3	0.21 (0.06)	0.001

after 1 year in the pattern mixture approach (0.21) was twice the estimated difference in slope of the random effects model (0.10) and was significant. However, in both approaches, the

absolute differences in PF after 1 year were small, as can be concluded comparing Figs. 3, 4, and 5.

Sensitivity analysis

The results of the analysis for only patients for which the disease state was known after 5 years (patients with maximum study period) revealed no differences compared to the results of the analysis for all patients. Therefore, we do not present the results for this restricted number of patients. In our data, the consequences of defining patterns based on patients with different length of study period were negligible.

Discussion

In this paper, the pattern mixture approach as a method for correction for informative drop-out was studied. We illustrated this on data from a large multi-center randomized trial comparing the long-term impact on HRQOL of two different doses chemotherapy for patients with breast cancer. We only considered one dimension of the HRQOL, namely PF. We focused on a subgroup of patients with normal HER2 expression in their tumor for whom it was shown that the (disease free) survival was better for the high dose than for the standard-dose arm (Rodenhuis et al. [16]). We compared the results of the pattern mixture approach to the results of the commonly used random effects model for the subscale PF.

In this particular example, the pattern mixture approach leads to differences with respect to the estimation and

significance of the difference in decline of PF measures over time between both arms after 1 year: The estimated difference in decline is doubled from 0.10 ($P = .055$) in the random effects model to 0.21 ($P = 0.001$) in the pattern mixture approach. However, the absolute estimated differences in PF between both chemotherapy doses after 1 year are similar in both approaches, and therefore the pattern mixture approach does not lead to other conclusions than the commonly used random effects model.

In studies with a lot of missing values due to drop-out for different reasons, it is important to explore whether the drop-out is informative or not. It is possible that for some outcomes the drop-out is informative, while for other outcomes this is not. An example for this is the case where the mean blood pressure is the main outcome in a study concerning a clinical patient population with high blood pressure, and where patients are repeatedly measured over time until their blood pressure is sufficiently reduced. So, patients are not measured any more (drop-out of the study) when they have “good” blood pressure values. In this case, the drop-out is caused by the previous measurement, and therefore not informative. So if one is planning a study with different sort of outcome measures, one must consider whether the drop-out is informative for each different outcome measure. In our data, we only considered drop-out by death or relapse as the two main sources of informative drop-out.

The crucial (and maybe the hardest) point in applying the pattern mixture approach is the choice of the patterns. The validity of the model estimates is determined by the implicit assumption that within each pattern, the missing data process is MAR. This is an untestable assumption; testing MCAR vs. MAR is possible, but testing MAR vs. MNAR is not (Molenberghs et al. [20]). The choice of the patterns must be made in such a way that this assumption is plausible. The number of patterns could be extended, based on the exact survival time or time of relapse, for example to account for the fact that shortly before death, the PF often drops and the amount of missing responses increases. In our data, we defined the patterns based on HRQOL-related events as death and relapse.

In our model, we had some censored patients, and we used their last observed event state for grouping them into patterns. Compared to the patients with maximum study period, relapse-free patients who are censored could still die or get tumor relapse within 5 years, implying that they belong to the pattern ‘deceased’ or ‘alive with relapse’. Analogously, patients with tumor relapse could still die within 5 years and therefore belongs to the pattern ‘deceased’. The implication of the different length of study period follow-up might be that the assumption of MAR within each pattern is violated. To check this, we also modeled the data for only patients with complete follow-up. Comparing these results with the results of the analysis of all

patients, we found similar results with somewhat larger standard errors. This means that for our data, there is no large problem here on this point. But for the application on other data with patterns based on disease states for patients with different lengths of follow-up, one must carefully consider the point whether the data are MAR within the patterns, and check this by sensitivity analysis, if possible.

The patterns chosen here do have clinical relevance. It is clinically interesting to estimate the courses over time for relapse-free patients. This could be interesting in terms of prediction: What is the HRQOL during the next 5 years, assuming that patients remain disease free? However, our model is less appropriate for the prediction for patients alive with relapse, because the pattern is based on the state at 5 year and the estimated PF measures over time are based on PF measures for patients with relapse and patients who did not yet have a relapse at that time point. The pattern “deceased” needs also special attention. Death is different from other reasons of missing, where we can assume that the quality of life exists but is not measured. If and how quality of life after death is defined is a topic of much discussion. The random effects models used in this paper implicitly impute missing PF scores, also after a patient has died. This is done by extrapolating the patient’s individual trajectory before death. One could question this approach. Therefore, one might want to limit the analysis for patients being alive. Pauler et al. [9] proposed to use the pattern mixture model estimates conditional on not being in pattern “deceased”. Diggle et al. [14] discussed extensively the validity of different models accounting for informative drop-out, including the relatively new approaches where measures (as HRQOL) and times to events (as relapse and death) are jointly modeled. Applying the latter models to these particularly data will be a topic for new research.

In general, it is recommended in data with informative drop-out to do sensitivity analysis by using different approaches. The pattern mixture approach we propose here is rather a simple extension of the commonly used random effects models in longitudinal data, and might therefore be routinely employed in the analysis of randomized clinical trials.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

Appendix

Model specification empty model

$$Y_{ij} = \beta_0 + u_i + e_{ij},$$

where Y_{ij} is the HRQOL of patient i on time point j , u_i the random term indicating the between-person variability, and

e_{ij} the random term indicating the residual variance. The random terms are independent and assumed to be normally distributed with mean zero and constant variance, notated by

$$u_i \sim N(0, \sigma_u^2) \quad \text{and} \quad e_{ij} \sim N(0, \sigma_e^2).$$

Model specification time model

The course over time (in months after randomization) is specified using four different time variables, namely t_0 , t_1 , t_2 , and t_c defined as follows.

Let t_0 be the dummy variable for time which equals 1 at baseline and zero afterward. Let, similarly, t_1 be the dummy variable, which equals 1 at 3 months, and zero on other time points and t_2 the dummy variable, which equals 1 at 6 months, and zero at other time points. Let t_c be a variable equal to 0 in the first year and equal to time-12 thereafter. In this way, the effect of time is assumed to be continuously increasing or decreasing after 1 year.

The time model can be specified as follows

$$Y_{ij} = (\beta_0 + u_{0i}) \cdot t_0 + (\beta_1 + u_{1i}) \cdot t_1 + (\beta_2 + u_{2i}) \cdot t_2 + \beta_3 + u_{3i} + (\beta_c + u_{ci}) \cdot t_c + e_{ij},$$

where $\beta_0, \beta_1, \beta_2, \beta_3,$ and β_c are the fixed effects, and $u_{0i}, u_{1i}, u_{2i}, u_{3i}, u_{ci}$ the random effects.

Interpretation

- $\beta_0 + \beta_3$: HRQOL at baseline
- $\beta_1 + \beta_3$: HRQOL at 3 months (just after chemotherapy)
- $\beta_2 + \beta_3$: HRQOL at 6 months
- β_3 : HRQOL at 1 year
- β_c : slope of the HRQOL course over time after 1 year

The random effects $u_{0i}, u_{1i}, u_{2i},$ and u_{3i} can be grouped in a similar way, but now indicating the random variability between patients at each time point. The random effect u_{ci} indicates the random variability of the slope between patients, and e_{ij} is the residual variance.

Model specification final model

Define

$$f(\text{time}) = (\beta_0 + u_{0i}) \cdot t_0 + (\beta_1 + u_{1i}) \cdot t_1 + (\beta_2 + u_{2i}) \cdot t_2 + \beta_3 + u_{3i} + (\beta_c + u_{ci}) \cdot t_c + e_{ij}$$

indicating the time model.

The final model is the time model plus the effect of treatment arm and all interactions (as fixed effects) between treatment arm and time variables $t_0, t_1, t_2,$ and t_c .

So, the final model is specified as follows

$$Y_{ij} = f(\text{time} * \text{dose}) \\ = f(\text{time}) + \beta_d \cdot \text{dose} + \beta_{d0} \cdot \text{dose} \cdot t_0 + \beta_{d1} \cdot \text{dose} \cdot t_1 + \beta_{d2} \cdot \text{dose} \cdot t_2 + \beta_{dc} \cdot \text{dose} \cdot t_c,$$

where dose equals 1 for the high-dose arm and 0 for the standard dose.

Note that the interaction term dose: t_3 is not included in the model, since this would lead to over specification of the model. This is more evident when considering the interpretation of the different fixed effects: each extra parameter in this model reflects the difference in doses for each time variable.

Interpretation

- $\beta_d + \beta_{d0}$: difference in HRQOL between both doses at baseline.
- $\beta_d + \beta_{d1}$: difference in HRQOL between both doses at 3 months.
- $\beta_d + \beta_{d2}$: difference in HRQOL between both doses at 6 months.
- β_d : difference in HRQOL between both doses at 1 year
- β_{dc} : difference in slopes between both doses after 1 year

The fixed part of $f(\text{time})$ reflects the course over time for patients in the standard dose. Filling in 1 for dose in the fixed part of $f(\text{time} * \text{dose})$ yields the course over time for patients in the high dose.

Specification of the pattern mixture model

Let $pat1$ be the pattern dummy variable equal to 1 for patients with relapse during the study period, and zero otherwise. Let $pat2$ be the pattern dummy variable equal to 1 for relapse-free patients during study period, and zero otherwise.

The random effects model in the pattern mixture approach (called the pattern mixture model) is the final model with additional the pattern dummy variables and all interactions between the variables of the final model with these pattern dummy variables.

Let $f(\text{time} * \text{dose})$ be the final model. The pattern mixture model can then be specified as

$$Y_{ij} = f(\text{time} * \text{dose}) + \gamma_1 \cdot pat1 + \gamma_2 \cdot pat2 + \gamma_{1d} \cdot pat1 \cdot \text{dose} + \gamma_{2d} \cdot pat2 \cdot \text{dose} + \gamma_{10} \cdot pat1 \cdot t_0 + \gamma_{20} \cdot pat2 \cdot t_0 + \gamma_{11} \cdot pat1 \cdot t_1 + \gamma_{21} \cdot pat2 \cdot t_1 + \gamma_{12} \cdot pat1 \cdot t_2 + \gamma_{22} \cdot pat2 \cdot t_2 + \gamma_{1c} \cdot pat1 \cdot t_c + \gamma_{2c} \cdot pat2 \cdot t_c + \gamma_{1d0} \cdot pat1 \cdot \text{dose} \cdot t_0 + \gamma_{2d0} \cdot pat2 \cdot \text{dose} \cdot t_0 + \gamma_{1d1} \cdot pat1 \cdot \text{dose} \cdot t_1 + \gamma_{2d1} \cdot pat2 \cdot \text{dose} \cdot t_1 + \gamma_{1d2} \cdot pat1 \cdot \text{dose} \cdot t_2 + \gamma_{2d2} \cdot pat2 \cdot \text{dose} \cdot t_2 + \gamma_{1dc} \cdot pat1 \cdot \text{dose} \cdot t_c + \gamma_{2dc} \cdot pat2 \cdot \text{dose} \cdot t_c,$$

where γ_1 to γ_{2dc} are the fixed parameters.

Interpretation

The fixed part of $f(\text{time} * \text{dose})$ indicates the course over time for both doses for the deceased patients and can be interpreted as in the final model. The fixed part of $f(\text{time} * \text{dose})$ and additionally all variables including pat1 reflects the course over time for both doses for the patients with relapse. The fixed part of $f(\text{time} * \text{dose})$ and all variables including pat2 reflects the course over time for both doses for relapse-free patients (rel free).

So,

$$Y_{ij}(\text{deceased}) = f(\text{time} * \text{dose})$$

$$\begin{aligned} Y_{ij}(\text{relapse}) = & f(\text{time} * \text{dose}) + \gamma_1 + \gamma_{1d} \cdot \text{dose} \\ & + \gamma_{10} \cdot t_0 + \gamma_{11} \cdot t_1 + \gamma_{12} \cdot t_2 + \gamma_{1c} \cdot tc \\ & + \gamma_{1d0} \cdot \text{dose} \cdot t_0 + \gamma_{1d1} \cdot \text{dose} \cdot t_1 \\ & + \gamma_{1d2} \cdot \text{dose} \cdot t_2 + \gamma_{1dc} \cdot \text{dose} \cdot tc \end{aligned}$$

$$\begin{aligned} Y_{ij}(\text{rel free}) = & f(\text{time} * \text{dose}) + \gamma_2 + \gamma_{2d} \cdot \text{dose} + \gamma_{20} \cdot t_0 \\ & + \gamma_{21} \cdot t_1 + \gamma_{22} \cdot t_2 + \gamma_{2c} \cdot tc \\ & + \gamma_{2d0} \cdot \text{dose} \cdot t_0 + \gamma_{2d1} \cdot \text{dose} \cdot t_1 \\ & + \gamma_{2d2} \cdot \text{dose} \cdot t_2 + \gamma_{2dc} \cdot \text{dose} \cdot tc \end{aligned}$$

The results for patients in the standard dose are obtained for dose equal to zero (the fixed part of the time model). The results for patients in the high dose are obtained for dose equal to one.

Weighting over all patterns in the pattern mixture approach

In the pattern mixture approach, the drop-out process is modeled by the probability to belong to a specific drop-out pattern for each dose separately. Let $\pi_0 = (\pi_{00}, \pi_{01}, \pi_{02})$ be the vector of probabilities to belong to patterns 0 (deceased), 1 (relapse), or 2 (relapse free), respectively, for patients in the standard dose. Let $\pi_1 = (\pi_{10}, \pi_{11}, \pi_{12})$ be the vector of probabilities to belong to patterns 0 (deceased), 1 (relapse), or 2 (relapse free), respectively, for patients in the high dose.

The results of the pattern mixture approach are obtained by weighting the courses over time of the different patterns by their corresponding proportions.

So,

$$\begin{aligned} Y_{ij}(\text{pattern mixture approach}) = & \pi_{00} \cdot Y_{ij}(\text{deceased}, \text{dose} = 0) \\ & + \pi_{10} \cdot Y_{ij}(\text{deceased}, \text{dose} = 1) \\ & + \pi_{01} \cdot Y_{ij}(\text{relapse}, \text{dose} = 0) \\ & + \pi_{11} \cdot Y_{ij}(\text{relapse}, \text{dose} = 1) \\ & + \pi_{02} \cdot Y_{ij}(\text{rel free}, \text{dose} = 0) \\ & + \pi_{12} \cdot Y_{ij}(\text{rel free}, \text{dose} = 1). \end{aligned}$$

References

- Groenvold, M., Fayers, P. M., Petersen, M. A., et al. (2006). Chemotherapy versus ovarian ablation as adjuvant therapy for breast cancer: impact on health-related quality of life in a randomized trial. *Breast Cancer Research and Treatment*, 98, 275–284.
- Bottomley, A., & Therasse, P. (2002). Quality of life in patients undergoing systemic therapy for advanced breast cancer. *Lancet Oncology*, 3(10), 620–628.
- Fitzmaurice, G., Laird, N., & Ware, J. (2004). *Applied longitudinal analysis*. New York: Wiley-Interscience.
- Rubin, D. B. (1976). Inference and missing data. *Biometrika*, 63, 581–592.
- Hedeker, D., & Gibbons, R. D. (2006). *Longitudinal data analysis*. New Jersey: Wiley-Interscience.
- Curran, D., Molenberghs, G., Aaronson, N. K., Fossa, S. D., & Sylvester, R. J. (2002). Analysing longitudinal continuous quality of life data with drop-out. *Statistical Methods in Medical Research*, 11, 5–23.
- Michiels, B., Molenberghs, G., Bijnes, L., Vangeneugden, T., & Thijs, H. (2002). Selection models and pattern-mixture models to analyse longitudinal quality of life subject to drop-out. *Statistical Methods*, 21, 1023–1041.
- Hedeker, D., & Gibbons, R. D. (1997). Application of random-effects pattern-mixture models for missing data in longitudinal studies. *Psychological Methods*, 2, 64–78.
- Pauler, D. K., McCoy, S., & Moinpour, C. (2003). Pattern mixture models for longitudinal quality of life studies in advanced stage disease. *Statistics in Medicine*, 22, 795–809.
- Diggle, P., & Kenward, M. G. (1994). Informative drop-out in longitudinal data analysis. *Applied Statistics*, 43, 49–93.
- Little, R. J. A. (1995). Modeling the drop-out mechanism in repeated-measures studies. *Journal of the American Statistical Association*, 90, 1112–1121.
- Curran, D., Molenberghs, G., Fayers, P. M., & Machin, S. (1998). Incomplete quality of life data in randomized trials: Missing forms. *Statistics of Medicine*, 17, 697–709.
- Kurland, B. F., & Heagerty, P. J. (2005). Directly parameterized regression conditioning on being alive: Analysis of longitudinal data truncated by deaths. *Biostatistics*, 6, 241–258.
- Diggle, P., Farewell, D., & Henderson, R. (2007). Analysis of longitudinal data with drop-out: Objectives, assumptions and a proposal. *Applied Statistics*, 5, 499–550.
- Rodenhuis, S., Bontenbal, M., Beex, L. V., et al. (2003). High dose chemotherapy with hematopoietic stem-cell rescue for high-risk breast cancer. *New England Journal of Medicine*, 349, 7–16.
- Rodenhuis, S., Bontenbal, M., Van Hoesel, Q. G. C. M., et al. (2006). Efficacy of high-dose alkylating chemotherapy in HER2/neu-negative breast cancer. *Annals of Oncology*, 17, 588–596.
- Buijs, C., Rodenhuis, S., Seynaeve, C. M., et al. (2007). Prospective study of long-term impact of adjuvant high-dose and conventional-dose chemotherapy on health-related quality of life. *Journal of Clinical Oncology*, 25, 5403–5409.
- Snijders, T. A. B., & Bosker, R. J. (1999). *Multilevel analysis. An introduction to basic and advanced multilevel modeling*. London: Sage Publications.
- Bishop, Y. M. M., Fienberg, S. E., & Holland, P. (1975). *Discrete multivariate analysis: Theory and practice*. Cambridge, MA: MIT Press.
- Molenberghs, G., Goetghebeur, E. J. T., Lipsitz, S. R., & Kenward, M. G. (1999). Nonrandom missingness in categorical data: Strengths and limitations. *The American Statistician*, 53, 110–118.