Aus dem Institut für Biometrie und klinische Epidemiologie
der Medizinischen Fakultät Charité – Universitätsmedizin Berlin

DISSERTATION

Missing values of patient-reported outcome data in a longitudinal study of advanced and metastatic cancer patients

Fehlende Werte bei Patienten-berichteten Endpunkten in einer Längsschnittstudie bei fortgeschrittenem und metastasiertem Krebs

zur Erlangung des akademischen Grades
Doctor rerum medicinalium (Dr. rer. medic.)

vorgelegt der Medizinischen Fakultät
Charité – Universitätsmedizin Berlin

von

Pimrapat Gebert

Datum der Promotion: 03.03.2023

Table of contents

List of tables iii

List of figures iv

List of abbreviations v

Abstract 1

1 Introduction 5

2 Methods 8

2.1 Oncological Social Care Project (OSCAR) Study 8

2.1.1 Study population 8

2.1.2 Data measurement 8

2.1.3 Event definition 10

2.2 Objective 1: To explore missing rates and missing data patterns of longitudinal PROs data 11

2.3 Objective 2: To investigate the association between patient characteristics and time-to-dropout or time-to-death 11

2.3.1 Sub-distribution hazard function 12

2.3.2 Cause-specific hazard function 12

2.3.3 Cumulative incidence function 13

3. Results 15

3.1 Objective 1: To explore missing rates and missing data patterns of longitudinal PROs data 15

3.1.1 Missing rates of PROs data 15

3.1.2 Missing data patterns of PROs data 18

3.1.3 Patient characteristics by type and reasons for missing PROs data 21

3.2 Objective 2: To investigate the association between patient characteristics and time-to-dropout or time-to-death 23

3.2.1 Cumulative incidence of dropout 23

3.2.2 Patient characteristics associated with time-to-dropout 23
3.2.3 Patient characteristics associated with time-to-death 24
4. Discussion 25
4.1 Summary of findings 25
4.2 Implications for statistical analysis methods with missing PROs data 25
4.2.1 Dropout rate and missing data patterns 25
4.2.2 Factors associated with dropout and missing data mechanisms 26
4.2.3 Statistical approaches for analyzing longitudinal PROs with missing data .27
4.3 Strengths and Limitations 28
5. Conclusions 29
Reference list 30
Statutory Declaration 34
Declaration of your own contribution to the publications 35
Excerpt from Journal Summary List 36
Printing copy of the publication 38
Curriculum Vitae 50
Publication list 51
Acknowledgments 53

List of tables

Table 1: Patient-reported Outcome Measures (PROMs) in the OSCAR study 9
Table 2: Number of observations, missing rates and reasons for missing data at each visit in the OSCAR 19
Table 3: Missing data patterns by type of attrition in OSCAR separately by study group and outcome 20
Table 4: Patient characteristics by reasons for missing data in the EORTC QLQ-C30 .22

List of figures

Figure 1: The visualization of time-to-event in OSCAR 10
Figure 2: Competing risk model for dropout, considering death as a competing event .12
Figure 3: Cumulative missing rates for the EORTC QLQ-C30 by study group 16
Figure 4: Cumulative missing rates for the other PROMs between study groups.....17
Figure 5: Cumulative missing rates for the EORTC QLQ-C30 and other PROMS 17
Figure 6: Mean and 95%CI of GH/QoL scales of the EORTC QLQ-C30 by missing data due to dropout or death and study groups. 18

List of abbreviations

API-DM: The German modified version of the Autonomy Preference Index
CASMIN: The Comparative Analysis of Social Mobility in Industrial Nations
CI: Confidence interval
CIF: Cumulative incidence function
DCS: Decision conflict scale
EORTC QLQ-C30: The European Organization for Research and Treatment of Cancer's Quality of Life Questionnaire
GH/QoL: Global health status/quality of life
HL: Health Literacy
HLS-EU-Q6: The European Health Literacy Survey
HR: Hazard ratio
HRQoL: Health-related quality of life
IQR: Interquartile range
LOCF: Last observation carried forward
MAR: Missing at random
MCAR: Missing completely at random
MNAR: Missing not at random
OSCAR: The Oncological Social Care Project
PRA-D: Patient reaction assessment
PROs: Patient-reported outcomes
PROMs: Patient-reported outcome measures
OSSS-3: Oslo Social Support Scale
QoL: Quality of life
SD: Standard deviation
SHR: Sub-hazard ratio

Abstract

Background

Missing patient-reported outcomes (PROs) are common in follow-up visits for longitudinal studies. However, the rates of missing values are relatively high in relation to health deterioration and premature death in patients with advanced stages of cancer. This study aims to investigate the rates and patterns of missing PROs data as well as to explore the association between patient characteristics and time-to-dropout or time-to-death for a better understanding of how the missing data mechanism could be applied to missing PROs data in severely affected patients.

Methods

This is an exploratory study using data from the Oncological Social Care Project. Missing rates and missing data patterns were reported using cumulative numbers and rates. The competing-risk analysis, using Fine and Gray's proportional sub-distribution hazards model, was performed to explore factors associated with time-to-dropout when considering death as a competing event. Additionally, a Cox regression model was used to explore factors associated with time-to-death.

Results

A total of 362 patients were observed. The cumulative missing rate for assessing PROs was around 28% and 19%, caused by premature death and dropout, respectively. Being divorced or widowed (SHR=2.71; 95%CI: 1.12–6.56) and having poor social support (SHR=2.10; 95%CI: 1.01–4.35) were associated with early dropout. The presence of malignant neoplasm of pancreas cancer (HR = 2.48; 95%CI: 1.27–4.85) and a medium level of education (HR = 1.58; 95%CI: 1.02–2.45) were associated with premature death. Dropping out early from the study was associated with low baseline global health status/quality of life (GH/QoL) (SHR=1.14; 95%CI: 1.01–1.27) and low baseline role functioning, such as limited ability to do work or daily activities (SHR=1.10; 95%CI: 1.01–1.19). Furthermore, worsening scores of GH/QoL, functioning, and symptoms at baseline and at the last visit were associated with premature death.

Conclusion

In the advanced stages of cancer research, high rates of missing PROs data should be expected. The worsening of health-related quality of life (e.g., GH/QoL, physical functioning, fatigue scores) was associated with missing PROs data, suggesting that the missing data is not completely random. The investigation of patient characteristics associated with missing data is also informative and a prerequisite for further proper analysis of the data.

Zusammenfassung

Einleitung

Fehlende von den Patienten berichteten Ergebnisse (patient-reported outcomes: PROs) sind bei Nachuntersuchungen im Rahmen von Langzeitstudien üblich. Die Raten fehlender Werte sind jedoch relativ hoch, wenn es um die Verschlechterung des Gesundheitszustands und den vorzeitigen Tod von Patienten in fortgeschrittenen Krebsstadien geht. Ziel dieser Studie ist es, die Raten und Muster fehlender PRO-Daten zu untersuchen sowie den Zusammenhang zwischen Patientenmerkmalen und der Zeit bis zum Abbruch bzw. bis zum Tod zu erforschen, um besser zu verstehen, wie der Mechanismus fehlender Daten auf fehlende PRO-Daten bei schwer betroffenen Patienten angewendet werden könnte.

Methoden

Es handelt sich um eine explorative Studie, die Daten aus dem Oncological Social Care Project verwendet. Fehlende Raten und Muster fehlender Daten wurden anhand kumulativer Zahlen und Raten angegeben. Eine Analyse des konkurrierenden Risikos unter Verwendung des proportionalen Unterverteilungs-Hazards-Modells von Fine und Gray wurde durchgeführt, um Faktoren zu untersuchen, die mit der Zeit bis zum Abbruch in Verbindung stehen, wenn der Tod als konkurrierendes Risiko betrachtet wird. Zusätzlich wurde ein Cox-Regressionsmodell verwendet, um Faktoren zu untersuchen, die mit der Zeit bis zum Tod in Verbindung stehen.

Ergebnisse

Insgesamt wurden 362 Patienten beobachtet. Die kumulative Missing-rate für die Bewertung der PROs lag bei 28% bzw. 19%, verursacht durch vorzeitigen Tod bzw. Dropout. Geschieden oder verwitwet zu sein (SHR=2,71; 95%KI: 1,12 – 6,56) und geringe soziale Unterstützung (SHR=2,10; 95%KI: 1,01 – 4,35) waren mit einem frühen Abbruch verbunden. Das Vorhandensein einer bösartigen Neubildung der Pankreas (HR=2,48; 95%KI: 1,27 - 4,85) und ein mittleres Bildungsniveau (HR=1,58; 95%KI: 1,02 - 2,45) waren mit einem vorzeitigen Tod assoziiert. Ein vorzeitiger Abbruch aus der Studie war mit einem niedrigen Ausgangswert für den globalen Gesundheitszustand/die Lebensqualität (GH/QoL) (SHR=1,14; 95%KI: 1,01 - 1,27) und einem niedrigen Ausgangswert für die Rollenfunktion, wie z. B. einer eingeschränkten Fähigkeit zur Ausübung von Arbeit oder

täglichen Aktivitäten (SHR=1.10; 95%KI: 1,01 - 1,19) assoziiert. Darüber hinaus war eine Verschlechterung der Werte für GH/QoL, Funktionsfähigkeit und Symptome zu Studienbeginn und bei der letzten Visite mit einem vorzeitigen Versterben assoziiert.

Schlussfolgerung

In den fortgeschrittenen Stadien der Krebsforschung sind hohe Raten an fehlenden PRO-Daten zu erwarten. Die Verschlechterung der gesundheitsbezogenen Lebensqualität (z. B. GH/QoL, körperlichen Funktionsfähigkeit, Müdigkeit) stand in Verbindung mit fehlenden PROs-Daten. Das bedeutet, dass der Mechanismus der fehlenden Daten nicht völlig zufällig ist. Die Untersuchung von Patientenmerkmalen, die mit fehlenden Daten in Verbindung stehen, ist aufschlussreich und eine Voraussetzung für die weitere korrekte Auswertung der Daten.

1 Introduction

In oncology research, patient-reported outcomes (PROs) are often used as either primary or secondary outcomes (1). Tracking health-related quality of life (HRQoL) status in cancer patients has been implemented as routine in oncology clinics (2-4) by using electronic sources, such as web-based questionnaires and applications on smartphones and tablets. Generally, PROs are assessed as multiple domains or items, such as in the European Organization for Research and Treatment of Cancer quality of life questionnaire (EORTC QLQ-C30) (5) and the Health Literacy Survey (HLS-EU-Q6) (6). Missing data in PROs may be present in some items or, due to a drop-out, result in missing values for the whole domain or questionnaire.

Although missing data is a general problem in medical research, several problems may occur if an analysis is based solely on complete cases. Missing data in PROs reduces the study's power due to the sample size reduction (7, 8). Another potential problem when missing PROs data occurs is that the results might be biased in the estimation of intervention effects, as well as in the estimation of differential intervention effects in subgroups (7). For instance, the intervention effect on HRQoL might be overestimated because cancer patients with health deterioration are more likely to have missing data.

Missing data in PROs needs to be handled appropriately to avoid biased results. The decision on the appropriate methodology generally depends on the missing data mechanisms. Rubin has given three definitions of missing data mechanism (9), and the following describes the missing data mechanisms according to Rubin's definition, based on Carpenter and Kenward (10). Y_i is the dependent variable vector which is intended to be collected from *n* subjects $(i = 1, \dots, n)$ and X denotes the covariate variables. Thus $Y_i = \left(Y_i^O, Y_i^M\right)$ when Y_i^O is the observed data and Y_i^M is the missing data. Missing data distribution is described by R, $R_{ij} = 1$ if Y_{ij} observed and $R_{ij} = 0$ if Y_{ij} missing for subject *i* and observed time *j*, where *j*=1,…,k.

Missing PROs data could be defined as ***missing completely at random (MCAR)*** when $\Pr(R_i|Y_i, X) = \Pr(R_i)$ as the distribution of missing values are assumed to be independent of any other observed and unobserved variables. MCAR occurs if the missing value relates to neither observed nor unobserved data, such as when a patient forgets a follow-up appointment or moves to another city. If the missing data depends on the observed data (e.g., to specific characteristics at baseline, such as age, gender, and education),

this missing data is called ***missing at random (MAR)*** and can be expressed as $\Pr(R_i|Y_i, X) = \Pr(R_i|Y_i^O, X)$. The third missing data mechanism is the most problematic in terms of handling missing data: For HRQoL, if the probability of missing HRQoL relates to the patient's health status (e.g., the patient missed their follow-up visit due to disease progression), the missing data is defined as ***missing not at random (MNAR)***. In this case, the probability of missing HRQoL is associated with unobserved values of the HRQoL itself. This is expressed as $\Pr(R_i|Y_i, X) = \Pr(R_i|Y_i^M, Y_i^O, X)$. However, it is difficult to differentiate between MAR and MNAR due to the fact that the unknown dependent variable Y_i^M is involved (10).

Numerous studies have demonstrated strategies for optimizing and handling incomplete PROs data (7, 11). Unfortunately, neither reporting missing rates nor exploring missing data mechanisms in longitudinal PROs data are very common. Only around 30% of the research assessing PROs data as an outcome reported follow-up compliance rates (12, 13), and only 27% reported the strategy used for handling missing data (13). This information is particularly important in oncology research when using PROs as an outcome because it could help researchers to assess the quality of the study results and to select an appropriate statistical analysis method for their own studies, as well as to interpret the PROs data correctly (12, 14).

In addition to reporting the rates of missing PROs data, evaluating the patterns of missing PROs is also important. The information on different patterns of missing PROs can be used in the statistical analysis approach to yield unbiased estimations of intervention effects (15). Patterns of missing PROs can be classified based on the reasons for discontinuation or the type of attrition. Possible reasons for discontinuation are: 1) patient felt too ill, 2) clinician or nurse felt the patient was too ill, 3) patient felt it was inconvenient or took too much time, 4) patient felt it was a violation of privacy, 5) patient did not understand the actual language or was illiterate, 6) administrative failure to distribute the questionnaire to the patient, 7) loss of contact, 8) other reasons (15). Another way to define the missing data pattern is by using a type of attrition. This type of attrition can be categorized into two subgroups: intermittent missing data and monotone missing data patterns. In our study, follow-up visits were scheduled at 3, 6, and 12 months. Therefore, missing values due to missed visits at months 3 or 6 but not at months 12 were considered intermittent. On the other hand, if the patient completed assessments at baseline and 3 months but discontinued the study after 6 months, the missing values at months 6 and

12 were considered monotone. Intermittent missing data is less frequently compared to monotone missing data and is often assumed to be MCAR (16). The pattern of missing data is seldom reported in clinical research, although it could help researchers better understand the sources and directions of possible bias.

The probability of dropping out of the study or of premature death can be associated with baseline patient characteristics. Therefore, many studies have shown that several patient characteristics are associated with dropping out early. For example, old age was related to dropping out early (17-19), and males seemed more likely to quit the study early than females (20, 21). Baseline PROs like EORTC QLQ-C30 have been studied in association with dropping out early or premature death, and the results have shown that poor physical health was associated with premature death but was not associated with dropping out due to other reasons (17). Although there is some evidence of the relationship between a specific patient's characteristics and their relation to dropout in PROs, most literature has only focused on baseline characteristics. In this work, patient characteristics in subsequent study visits were analyzed to gain a better understanding of continuously measured patient characteristics during follow-ups and their association with dropout among severely affected cancer patients. This study will provide a better view of potential sources of bias that might impact the study results. Additionally, it will suggest strategies for handling the missing PROs data appropriately.

The main objective of this thesis is to develop a greater understanding of how missing data in longitudinal PROs study with advanced stages of cancer should be handled and reported in order to increase the quality of research. This work aims to achieve this goal by considering a number of aspects:

1. Exploring missing rates and missing patterns of longitudinal PROs data.
2. Investigating the association between patient characteristics and time-to-dropout or time-to-death.

2 Methods

2.1 Oncological Social Care Project (OSCAR) Study

2.1.1 Study population

Cancer patients were recruited for OSCAR from three study centers in Germany between January 2018 and February 2020. Eligible patients were 18 or over, with different cancer types (see the inclusion criteria in the published protocol (22)). The OSCAR is a non-randomized, controlled, multi-center intervention study. Patients who were insured by the German company health insurance fund, Pronova BKK, were enrolled in the intervention group, whereas those who were insured by the other companies were enrolled in the control group. The OSCAR was approved by the ethics committees at Charité–Universitätsmedizin Berlin (EA2/192/17) and the Medical Association of North Rhine-Westphalia (2017429). The participants were enrolled in the study after providing written informed consent. The intervention group received a monthly consultation provided by social care nurses, while the control group received standard care.

2.1.2 Data measurement

Demographic data and anamnesis were collected at baseline, e.g., age, sex, time since diagnosis, cancer diagnosis, family status, and education. Based on the Comparative Analysis of Social Mobility in Industrial Nations (CASMIN) classification of education(23), education and professional qualifications were categorized into three groups: low, medium, and high. Subjective social support was assessed using the Oslo Three-Item Social Support Scale (OSSS-3) (24). The OSSS-3 scores range from 3 to 14 points; 3 to 8 is poor, 9 to 11 is moderate, and 12 to 14 is strong. HRQoL according to the EORTC QLQ-C30 was assessed monthly in the intervention group. The EORTC QLQ-C30 for the control group and other PROMs for both groups were assessed at baseline (t0), 3 months (t1), 6 months (t2), and 12 months (t3) by study nurses (22, 25). The primary outcome of the OSCAR is the scores of GH/QoL of the EORTC QLQ-C30. An overview of the PROMs in the OSCAR is presented in Table 1.

Table 1: Patient-reported Outcome Measures (PROMs) in the OSCAR study

Questionnaires	Purpose	Score range	Interpretation
EORTC QLQ-C30 (version 3.0) (5)	to assess HRQoL in cancer patients. It is a multiple domain questionnaire, consisting of 30 questions and incorporating different functioning and symptoms.	0 to 100	Higher values indicate better functioning and GH/QoL, while higher values represent a worse health status in symptom subscales.
The Patient Reaction Assessment (PRA-D) (26)	to assess the quality of communication between physician and patient.	5 to 35	Higher values indicate better doctor-patient relationships.
The Illness perception questionnaire (IPQ-R) (27)	to assess the illness coherence subscale to represent an understanding or comprehend of patients on their illnesses.	5 to 25	Higher scores indicate better illness coherence.
The modified German version of the Autonomy Preference Index (API-DM) (28)	to measure patients' decision-making and information-seeking strategies.	0 to 100	Higher values indicate a greater preference for information seeking. The scores of decision-making close to 50 represent that decision-making is shared equally between physician and patient.
The Decision Conflict Scale (DCS) (29)	to measure patient perceptions of choosing options and decision-making	0 to 100	Higher values indicate high decision conflict.
The European Health Literacy Survey (HLS-EU-Q6) (6)	to assess how patients understand their health information in order to make a decision.	1 to 4	[1-2] scores=insufficient HL (2-3) scores=problematic HL [3-4] scores= health literacy HL

HL=Health literacy *[Own representation]*

2.1.3 Event definition

Non-participation and discontinuation from the OSCAR were documented for each follow-up visit. In this study, the focus is placed on discontinuation. Reasons for discontinuation were divided into two categories: dropout and premature death. Figure *1* presents the definition of time-to-event in the OSCAR. The terms for the analysis were defined as follows (25):

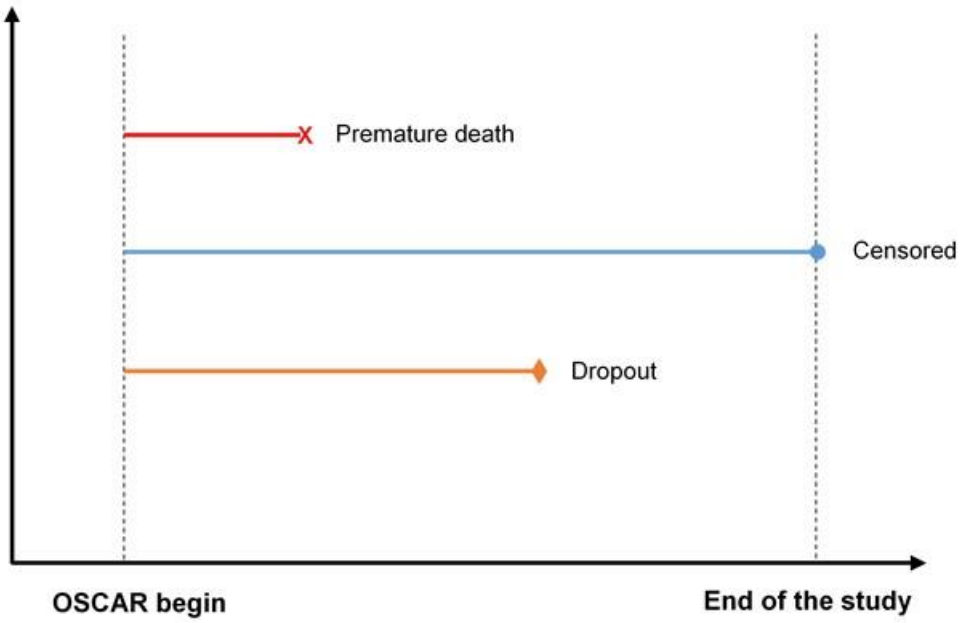
The event of interest was dropping out of the study, including dropout with and without health-related reasons, and loss to follow-up.

The competing event was premature death, which prevented the occurrence of dropout (the event of interest).

Time-to-death was defined from the enrollment date to the date of death if the patient died prematurely.

Time-to-dropout was defined from the enrollment date to the date of dropout when patients either refused to continue the study or the first date of unsuccessful contact (when study nurses could not reach the patient after trying to contact them three times).

Censored data was present if patients were alive and completed the study at 12 months.



[Own representation]

Figure 1: The visualization of time-to-event in OSCAR

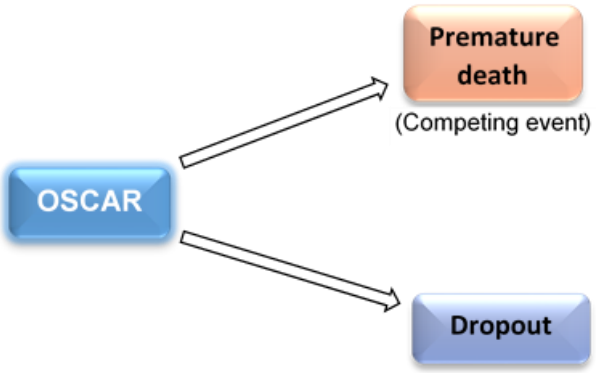
2.2 Objective 1: To explore missing rates and missing data patterns of longitudinal PROs data

Cumulative missing rates at each follow-up visit are separately presented as percentages by PROMs and study groups, as well as by visualizations using bar charts. For each reason for the missing data, cumulative frequencies in total and by study group are presented. Absolute and relative frequencies are reported for each reason in the intermittent missing data and by type of attrition. An estimated mean and a 95% confidence interval (CI) of GH/QoL are presented by a line graph, with error bars for each time point, separated by missing data due to dropout or death, and study groups.

2.3 Objective 2: To investigate the association between patient characteristics and time-to-dropout or time-to-death

Firstly, demographic data is presented by using descriptive statistics separately by subgroups defined by no dropout or reasons for dropout: fully-observed, dropout due to health-related reasons, dropout due to other reasons, and premature death. Time-to-dropout is another important outcome of interest, as it is useful to know how long patients remain in the study. In this study, a competing event is considered because three potential outcomes can be observed for each patient during the follow-up period: fully observed (censored), dropout, or premature death. It is essential to consider premature death as a competing event when evaluating time-to-dropout since a patient who dies prematurely is no longer at risk of dropping out. Unlike a general survival model with only one type of event (e.g., all-cause mortality), patients are supposed to have two types of follow-up events (dropout or premature death). For this reason, dropout can be considered an event of interest, and death “prevents” the occurrence of dropout. As the event ‘dropout’ is not observable if the patient died before the dropout occurred, premature death should be considered a competing event (Figure *2*).

In the presence of competing events, cause-specific hazard functions and sub-distribution hazard functions are widely used. In this dissertation, I reviewed both hazard functions for modelling competing risks.



[modified from Huebner et al., 2017 (30)]

Figure 2: Competing risk model for dropout, considering death as a competing event

2.3.1 Sub-distribution hazard function

The sub-distribution hazard function defines a hazard function as the probability of the event of interest after event-free time t without any event of interest or with the competing event before time t. This sub-distribution hazard function was proposed by Fine and Gray (31) and can be defined as:

$$\lambda_k(t) = \lim_{\Delta t \to 0} \frac{\Pr\left(t \leq T < t + \Delta t, K = k \middle| T \geq t \cup (T < t \cap K \neq k)\right)}{\Delta t}$$

Where *T* denotes a potential failure time, *k* is the type of event. *K*=k indicates event 1: dropout (k=1) or event 2: premature death (k=2).

The proportional sub-distribution hazards model can be written as:

$$\lambda_k(t|X) = \lambda_{0k}(t)\exp(X^t\beta_k)$$

Where $\lambda_{0k}(t)$ is the unspecified baseline sub-distribution hazards for the event of interest (dropout). *t* is the time to event, *X* is a vector of covariates, and β_k is a vector of regression coefficients where *k*=1, 2.

2.3.2 Cause-specific hazard function

The cause-specific hazard function uses all individuals at risk by considering patients who have a competing event as not having a risk for the event of interest. The function can be written as (32):

$$h_k(t) = \lim_{\Delta t \to 0} \frac{\Pr(t \leq T < t + \Delta t, K = k | T \geq t)}{\Delta t}$$

Where *T* denotes a potential failure time, *k* is the type of event. *K*=k indicates event 1: dropout (k=1) or event 2: premature death (k=2).

The cause-specific hazard model can be defined as

$$h_k(t|X) = h_{0k}(t)\exp(X^t\beta_k)$$

Where $h_{0k}(t)$ is the baseline cause-specific hazard for the event of interest. *T* is the time to event, *X* is a vector of covariates, and β_k is a vector of regression coefficients where *k*=1, 2.

2.3.3 Cumulative incidence function

Estimating the incidence of an event as a time-to-event function provides important information on the absolute risk of an event. In the presence of a competing event, the Cumulative Incidence Function (CIF) is presented for estimating the incidence of the event while accounting for the competing event without any distributional assumptions (33). The function *CIF (t)* denotes the probability of experiencing the event *k* before time *t* and before a different type of event occurs, where *K* denotes the type of event. The cumulative incidence function for the event k of interest can be written as follows (34):

$$CIF\ (t) = Pr(T \leq t, K = k)$$

In the calculation of the cumulative incidence functions, the number of individuals at risk of the event of interest is different for the cause-specific model and the sub-distribution model. The number of individuals at risk of the cause-specific function includes only those patients who have not yet experienced the event of interest and not the competing event, while the sub-distribution function includes both cases without the event of interest and those who have had the competing event. In other words, using a sub-distribution hazard function in the analysis of time-to-dropout assumes that patients who died would not have discontinued if they had survived for the whole study period. Therefore, the CIF from the sub-distribution function tends to be lower than the CIF from the cause-specific model where all cases remain at risk (32).

In this dissertation, Fine and Gray's proportional sub-distribution hazards model was used to estimate the cumulative incidence of dropout in the OSCAR and the association between patient characteristics and time-to-dropout in the presence of the competing event of premature death (25). According to one of the study purposes (estimating the time-to-dropout when considering premature death as a competing event), the calculation of this

sub-distribution hazard function, including those who have prematurely died in the risk set, makes sense when these patients are still considered at risk (meaning they still have a chance to dropout if they are alive at the end of the study). Additionally, a Cox proportional hazard model was fitted in order to understand the relationship between patient characteristics and time-to-death. For this, the outcome of interest is premature death among cancer patients. Here, a censored patient was defined as a patient who had discontinued the study or was observed until the end of the study. Effect sizes are reported as sub-hazard ratios (SHR) and hazard ratios (HR) with 95%CI for time-to-dropout and time-to-death, respectively. All of the statistical analyses were performed using Stata IC15 (StataCorp, 2017, College Station, TX, USA).

3. Results

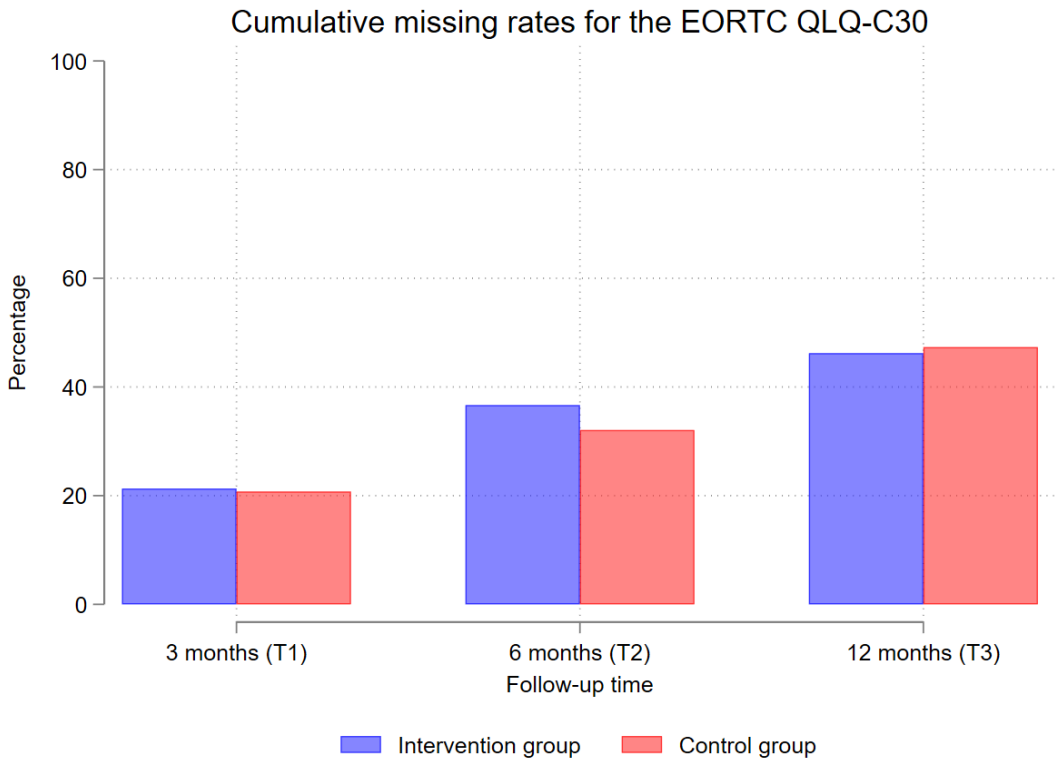
Overall, 366 patients were enrolled in the study; four cases were excluded later due to non-compliance. Therefore, 150 patients in the intervention group and 212 patients in the control group were studied. Around half of the patients dropped out or died prematurely before the study's end (12 months), with a median follow-up time of 333 days (IQR: 154, 361 days). One hundred ninety-three of 362 patients (53.3%) completed the EORTC QLQ-C30, and 189 patients (52.2%) completed the other PROMs at all four assessment time points.

3.1 Objective 1: To explore missing rates and missing data patterns of longitudinal PROs data

3.1.1 Missing rates of PROs data

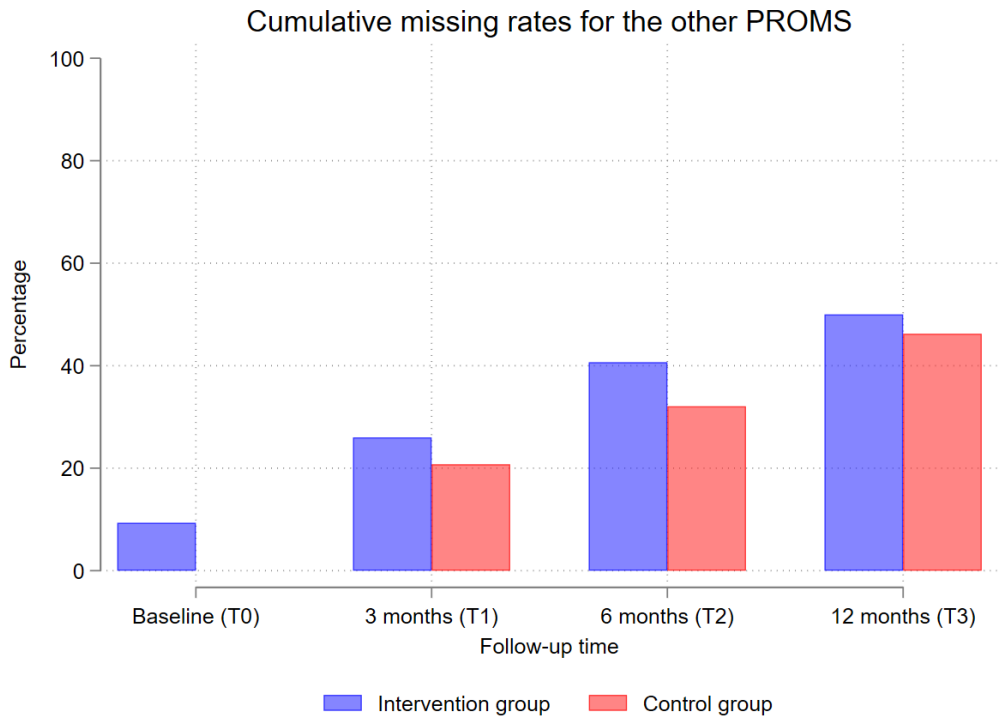
Table 2 presents the amount of missing data at each follow-up visit, the reasons for missing data, and the cumulative missing rates. The cumulative missing rate at the end of the study was 46.7% (169/362) for the EORTC QLQ-C30 and 47.8% (173/362) for the other PROMs. There was no substantial difference in the cumulative missing rates of the EORTC QLQ-C30 between study groups, with 46.0% (69/150) at 12 months for the intervention group and 47.2% (100/212) for the control group (Figure *3*). Most of this missing data was caused by premature death (102/362=28.2%), though 12% (44/362) of the patients dropped out for health-related reasons (11.3% (17/150) in the intervention group and 12.7% (27/212) in the control group). 21.0% (76/362) of missing data on the EORTC QLQ-C30 was observed after 3 months. Whereas the other PROMs had 22.9% (83/362) of missing data, 3.9% (14/362) of this missing data happened after the initial visit because the study nurses evaluated the other PROMs later after the social care nurses assessed the EORTC QLQ-C30 in the intervention group. This causes the difference in missing rates between the EORTC QLQ-C30 and other PROMs in OSCAR, such as those who had post-op mortality. For other PROMs, the cumulative missing rate was slightly higher in the intervention group than in the control group, with the missing rate at 12 months of 50.0% and 46.6%, respectively (Figure *4*). In addition, there was no difference in the cumulative missing rates for the assessment of the EORTC QLQ-C30 and the other PROMs at each follow-up visit (Figure *5*). Furthermore, the intermittent missing data was low in both groups, but the missing rates

were slightly higher in the control group compared to the intervention group for the assessment of the EORTC QLQ-C30, but this was not the case for the other PROMs, where the intermittent missing rates were 4.0%-10.0% in the intervention group and just around 4.0% in the control group.



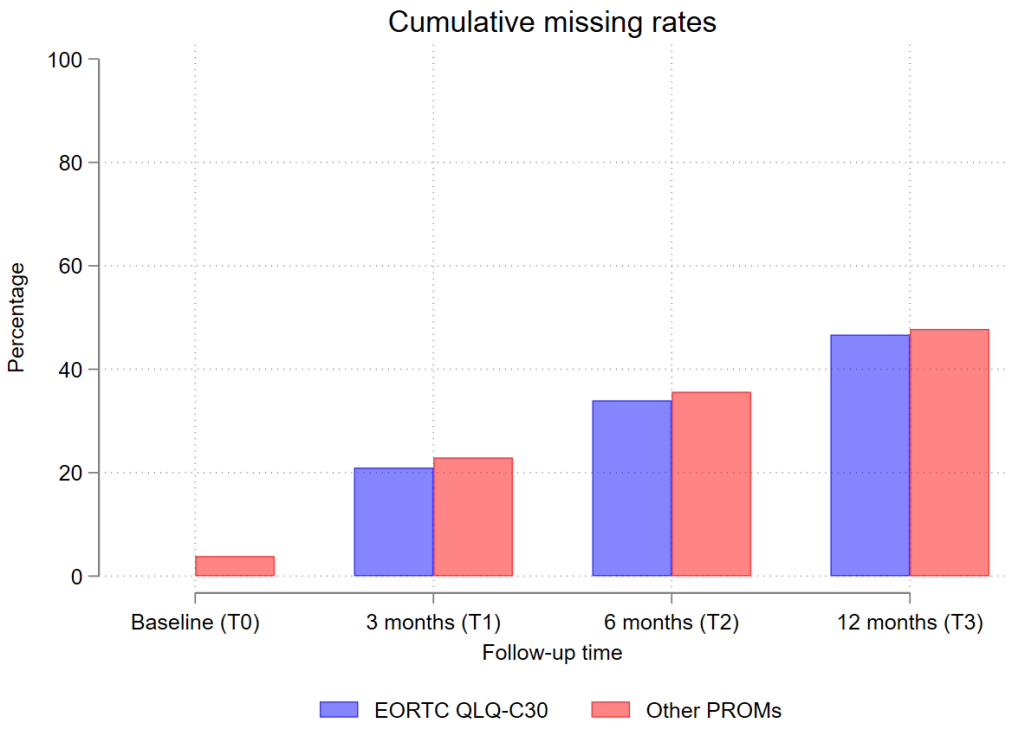
[Own representation]

Figure 3: Cumulative missing rates for the EORTC QLQ-C30 by study group



[Own representation]

Figure 4: Cumulative missing rates for the other PROMs between study groups



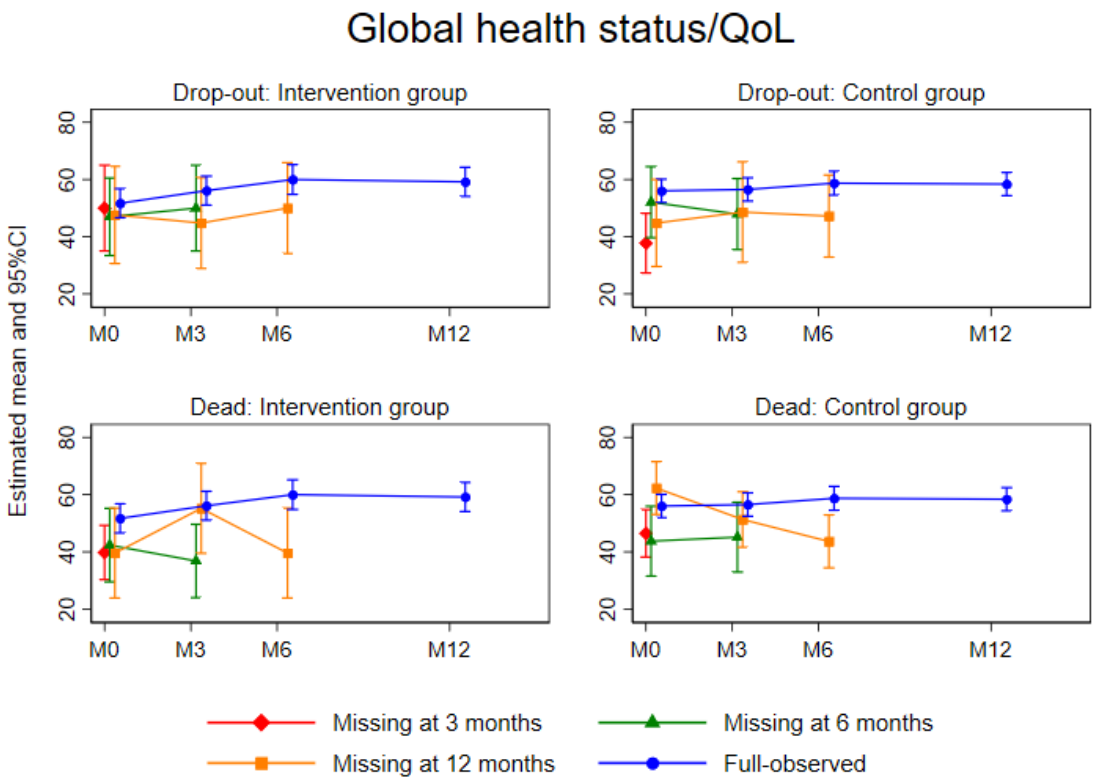
[Own representation]

Figure 5: Cumulative missing rates for the EORTC QLQ-C30 and other PROMS

3.1.2 Missing data patterns of PROs data

Missing data patterns by type of attrition in OSCAR are explored in Table 3. Overall, the completion rate for the EORTC QLQ-C30 and other PROMs was 50.8% and 46.7%, respectively. Patients in the intervention group completed the EORTC QLQ-C30 more often than the control group (52.7% vs. 49.5%), while the completion rate of other PROMs was lower in the intervention group than in the control group (42.7% vs. 49.5%). The most common missing data pattern in OSCAR was a monotone, and around 20% of the patients dropped out of the study after T1 (3 months). Intermittent or non-monotone missing data were infrequent (ca. <2%).

Figure 6 presents the missing data patterns for the GH/QoL of the EORTC QLQ-C30. It can be seen that the GH/QoL scores for patients who completed the study were higher than those who dropped out or died within the study period. Although the GH/QoL scores decreased before premature death, this trend was observed in both study groups.



[Own representation]

Figure 6: Mean and 95%CI of GH/QoL scales of the EORTC QLQ-C30 by missing data due to dropout or death and study groups.
The possible range is 0-100, with higher values indicating better GH/QoL. M0=Baseline, M3=3 months, M6=6 months, M12=12 months (end of the study).

Table 2: Number of observations, missing rates and reasons for missing data at each visit in the OSCAR

Observation times	Total (n=362)				Intervention group (n=150)				Control group (n=212)			
	T0	T1	T2	T3	T0	T1	T2	T3	T0	T1	T2	T3
EORTC QLQ-C30												
Number of patients	**362**	**362**	**286**	**239**	**150**	**150**	**118**	**95**	**212**	**212**	**168**	**144**
Cumulative number of missing data (patients)	**0**	**76**	**123**	**169**	**0**	**32**	**54**	**69**	**0**	**44**	**69**	**100**
Cumulative missing rate	0.0%	21.0%	34.0%	46.7%	0.0%	21.3%	36.0%	46.0%	0.0%	20.8%	32.5%	47.2%
Reason for missing (Cumulative numbers)												
Premature death	0	49	73	102	0	22	34	42	0	27	39	60
Dropout due to health-related reasons	0	19	35	44	0	7	12	17	0	12	23	27
Dropout due to other reasons	0	4	10	13	0	3	8	10	0	1	2	3
Loss of contact	0	4	5	10	0	0	0	0	0	4	5	10
Number of Intermittent missing data[1]	**1**	**11**	**6**	**0**	**1**	**3**	**0**	**0**	**0**	**8**	**6**	**0**
Missing rate	0.3%	2.8%	2.1%	0.0%	0.7%	2.0%	0.0%	0.0%	0.0%	3.8%	3.6%	0.0%
Reason for missing data												
Unable to contact	0	5	1	0	0	1	0	0	0	4	1	0
Refusal due to health-related reasons	1	6	5	0	1	2	0	0	0	4	5	0
Other PROMs*												
Number of patients	**362**	**348**	**279**	**233**	**150**	**136**	**111**	**89**	**212**	**212**	**168**	**144**
Cumulative number of missing data (patients)	**14**	**83**	**129**	**173**	**14**	**39**	**60**	**73**	**0**	**44**	**69**	**100**
Cumulative missing rate	3.9%	22.9%	35.6%	47.8%	9.3%	26.0%	40.0%	48.7%	0.0%	20.8%	32.5%	47.2%
Reason for missing (Cumulative numbers)												
Premature death	10	54	80	111	10[2]	27	41	51	0	27	39	60
Dropout due to health-related reasons	0	18	32	38	0	6	9	11	0	12	23	27
Dropout due to other reasons	3	4	8	10	3	3	6	7	0	1	2	3
Loss of contact	1	7	9	14	1	3	4	4	0	4	5	10
Number of Intermittent missing data	**9**	**21**	**11**	**7**	**9**	**13**	**5**	**7**	**0**	**8**	**6**	**0**
Missing rate	2.5%	6.0%	3.9%	3.0%	6.0%	9.6%	4.5%	7.9%	0.0%	3.8%	3.6%	0.0%
Reason for missing data												
Unable to contact	5	12	4	6	5	8	3	6	0	4	1	0
Refusal due to health-related reasons	4	9	7	1	4	5	2	1	0	4	5	0

[*Own representation*] *PRA-D, API-DM, IPQ-R, DCS, and HLS-EU-Q6. T0=baseline, T1=3 months, T2=6 months, T3=12 months.

[1] Intermittent missing data: patients missed a specific visit, but they come back to the visit later.

[2] The missing rate is unequal in the intervention group due to the study nurse assessing the other PROMs and the social care nurses assessing the EORTC QLQ-C30 at a different time.

Table 3: Missing data patterns by type of attrition in OSCAR separately by study group and outcome

Patterns of missing data	EORTC QLQ-C30							Other PROMs: PRA-D, API-DM,IPQ-R, DCS, HLS-EU-Q6						
	Wave				Intervention	Control	Total	Wave				Intervention	Control	Total
	T0	T1	T2	T3	(n=150)	(n=212)	(n=362)	T0	T1	T2	T3	(n=150)	(n=212)	(n=362)
Full observation	O	O	O	O	79 (52.7%)	105 (49.5%)	184 (50.8%)	O	O	O	O	64 (42.7%)	105 (49.5%)	169 (46.7%)
Monotone missing data[1]	M	M	M	M	–	–	–	M	M	M	M	20[2] (13.3%)	–	20 (5.5%)
	O	M	M	M	34 (22.7%)	44 (20.8%)	78 (21.5%)	O	M	M	M	27 (18.0%)	44 (20.8%)	71 (19.6%)
	O	O	M	M	21 (14.0%)	24 (11.3%)	45 (12.4%)	O	O	M	M	17 (11.3%)	24 (11.3%)	41 (11.3%)
	O	O	O	M	15 (10.0%)	25 (11.8%)	40 (11.0%)	O	O	O	M	14 (9.3%)	25 (11.8%)	39 (10.8%)
Non-monotone missing data	M	O	M	M	–	–	–	M	O	M	M	2 (1.3%)	–	2 (0.6%)
	M	O	O	M	1 (0.7%)	–	1 (0.3%)	M	O	O	M	–	–	–
	O	O	M	O	–	6 (2.8%)	6 (1.7%)	O	O	M	O	0 (0.0%)	6 (2.8%)	6 (1.7%)
	M	O	O	O	–	–	–	M	O	O	O	1 (0.7%)	–	1 (0.3%)
	O	M	O	M	–	5 (2.4%)	5 (1.4%)	O	M	O	M	2 (1.3%)	5 (2.4%)	7 (1.9%)
	O	M	O	O	–	3 (1.4%)	3 (0.8%)	O	M	O	O	3 (2.0%)	3 (1.4%)	6 (1.7%)

[*Own representation*] O=Observed data, M=Missing data, T0=baseline, T1=3 months, T2=6 months, T3=12 months.

[1] Monotone missing data combines missing data due to dropout and death.

[2] The missing rate occurred in the intervention group after T0 because the other PROMs were assessed by the study nurse later after the social care nurses assessed the EORTC QLQ-C30 at enrollment.

3.1.3 Patient characteristics by type and reasons for missing PROs data

Of all 362 patients, 102 cases (28.2%) died during the study period, 51 cases (14.1%) had missing data due to health-related reasons, and 25 cases (6.9%) had missing data for other reasons and loss of contact. The remaining 184 cases (50.8%) completed questionnaires at all follow-up visits (Table 4). The mean age of patients who did not complete all visits was slightly higher than those who completed the study visits. There was no remarkable difference in the rate of missing data for the intervention and control groups, with missing data due to health-related or other reasons at 11.3% and 8.0% in the intervention group and 16.0% and 6.1% in the control group, respectively. However, men were more likely to drop out for health-related or other reasons than women. By contrast, women had a higher rate of premature death than men, as 46 women (32.2%) died before their last follow-up visit, compared to 56 men (25.6%). There was a difference in the missing rate for patients with different family status: married patients were less likely to drop out but had a higher rate of premature death than divorced or widowed patients. In patients who had a longer time period between diagnosis and study entry, rates of premature death were higher than in those who had been diagnosed shortly before beginning the study.

Premature death and dropout rates were highest in patients with malignant neoplasm of the pancreas (50.0%), malignant neoplasm of the bronchus and lung (37.1%), metastatic colorectal cancer or colon carcinoma (28.2%), and acute leukemia (27.5%). Patients with a low (20.7%) or medium (15.8%) level of education had a slightly higher dropout rate for health-related reasons than those with a higher level of education (12.1%). Furthermore, patients with low social support (14.3%) were more likely to drop out for other reasons than patients with moderate (8.2%) or strong (2.2%) social support.

Table 4: Patient characteristics by reasons for missing data in the EORTC QLQ-C30

Patient's characteristics	Total (n=362)	Full-observed (n=184)	Missing due to other reasons (n=25)	Missing due to health-related reasons (n=51)	premature death (n=102)
Study centers					
Study center 1	119	74 (62.2%)	7 (5.9%)	13 (10.9%)	25 (21.0%)
Study center 2	98	40 (40.8%)	8 (8.2%)	19 (19.4%)	31 (31.6%)
Study center 3	145	70 (48.3%)	10 (6.9%)	19 (13.1%)	46 (31.7%)
Study group					
Intervention	150	79 (52.7%)	12 (8.0%)	17 (11.3%)	42 (28.0%)
Control	212	105 (49.5%)	13 (6.1%)	34 (16.0%)	60 (28.3%)
Age (years) – Mean (SD)	63 (13)	62 (13)	65 (13)	64 (13)	65 (13)
Sex – (%)					
Male	219	110 (50.2%)	18 (8.2%)	35 (16.0%)	56 (25.6%)
Female	143	74 (51.7%)	7 (4.9%)	16 (11.2%)	46 (32.2%)
Family status					
Married	226	119 (52.7%)	10 (4.4%)	29 (12.8%)	68 (30.1%)
Single	44	28 (63.6%)	3 (6.8%)	4 (9.1%)	9 (20.5%)
Divorced/Widowed	63	31 (49.2%)	7 (11.1%)	15 (23.8%)	10 (15.9%)
Missing	29	6 (20.7%)	5 (17.2%)	3 (10.3%)	15 (51.7%)
Time since diagnosis (months) - Median (IQR)	6.0 (2.0 – 22.0)	6.0 (2.0 – 21.5)	3.0 (2.0 – 25.0)	4.0 (2.0 – 20.0)	8.5 (2.0 – 22.0)
Diagnosis					
Acute leukemia	69	40 (58.0%)	3 (4.3%)	7 (10.1%)	19 (27.5%)
Lymphoma	58	33 (56.9%)	6 (10.3%)	10 (17.2%)	9 (15.5%)
Malignant neoplasm of the bronchus and lung	62	24 (38.7%)	4 (6.5%)	11 (17.7%)	23 (37.1%)
Metastatic colorectal cancer or colon carcinoma	78	42 (53.8%)	3 (3.8%)	11 (14.1%)	22 (28.2%)
Malignant neoplasm of the pancreas	32	7 (21.9%)	1 (3.1%)	8 (25.0%)	16 (50.0%)
Multiple myeloma and malignant plasma cell neoplasms	24	16 (66.7%)	3 (12.5%)	1 (4.2%)	4 (16.7%)
Metastasized malignant neoplasm of the breast	9	7 (77.8%)	1 (11.1%)	0 (0.0%)	1 (11.1%)
Others	30	15 (50.0%)	4 (13.3%)	3 (10.0%)	8 (26.7%)
Education					
Low	34	15 (51.7%)	2 (6.9%)	6 (20.7%)	6 (20.7%)
Medium	74	46 (45.5%)	6 (5.9%)	16 (15.8%)	33 (32.7%)
High	181	117 (56.8%)	13 (6.3%)	25 (12.1%)	51 (24.8%)
Missing	73	6 (23.1%)	4 (15.4%)	4 (15.4%)	12 (46.2%)
Social support (OSSS-3)					
Mean (SD)	11.0 (2.1)	11.0 (2.0)	9.9 (2.0)	10.9 (2.1)	11.2 (2.1)
Poor (3 – 8)	35	17 (48.6%)	5 (14.3%)	5 (14.3%)	8 (22.9%)
Moderate (9 – 11)	158	83 (52.5%)	13 (8.2%)	23 (14.6%)	39 (24.7%)
Strong (12 – 14)	139	79 (56.8%)	3 (2.2%)	18 (12.9%)	39 (28.1%)
Missing	30	5 (16.7%)	4 (13.3%)	5 (16.7%)	16 (53.3%)

[*Own representation*] Missing data without and with health-related included intermittent missing data n=2 and n=7, respectively. Loss of contact was included in missing data without health-related reasons.

3.2 Objective 2: To investigate the association between patient characteristics and time-to-dropout or time-to-death

In this attempt to identify patient characteristics associated with missing PROs data due to dropout, time-to-event models were performed with premature death considered as a competing event (25). The proportional sub-distribution hazard model by Fine and Gray was used to estimate the cumulative incidence of dropout and the SHR for time-to-dropout with 95%CI. The event of interest was dropout, and for a small dropout number for other reasons, dropouts due to health-related (n=51) or other reasons (n=25) were both combined. It should be noted that a multivariable analysis was not possible here because of multicollinearity[5] problems. For example, there was a correlation between the functional and symptom subscales in the EORTC QLQ C-30 (e.g., r^6=0.532 for GH/QoL and physical functioning or r=-0.633 for physical functioning and fatigue) and a correlation in the variables was measured at baseline and at the last visit (e.g., r=0.486 for physical functioning). Therefore, this work was only done in an explorative framework using bivariate analyses.

3.2.1 Cumulative incidence of dropout

The overall cumulative incidence of dropouts (including premature death) was 14.4% at 3 months, 31.5% at 6 months, and 46.8% at 12 months. The cumulative incidence of dropouts accounting for death as a competing event was 4.4% at 3 months, 12.1% at 6 months, and 18.8% at 12 months. The cumulative incidence of dropout at 12 months was slightly higher in the intervention group than in the control group (overall dropout and premature death: intervention=47.4%, control=44.2%; dropout with death as a competing event: intervention=18.8%, control=16.4%).

3.2.2 Patient characteristics associated with time-to-dropout

When considering death as a competing event, the results showed that being divorced or widowed increased the likelihood of dropping out (SHR=2.71; 95%CI: 1.12 – 6.56). Patients with poor social support were more likely to drop out early than those with strong

[5] In a statistical model, multicollinearity occurs when there are high correlations between independent variables, and it can lead to misleading results when the analysis attempts to determine how well each independent variable predicts the dependent variable.
[6] The correlation coefficient *r* ranges from -1 to 1, with value close to 1 (or -1) indicating high correlation.

social support (SHR=2.10; 95%CI: 1.01 – 4.35). Although being older, male, having a low level of education, and having malignant neoplasm of the pancreases, malignant neoplasm of the bronchus and lungs, and aggressive lymphoma were all positively associated with dropout as seen in the descriptive analyses, there was no substantial association with these characteristics for dropping out when death was considered as a competing event. The values of the PROs data were associated with dropout: patients who had low GH/QoL[7] scores at baseline had a higher likelihood of dropping out early (SHR=1.14; 95%CI: 1.01 – 1.27), as well as patients with low scores in role functioning (SHR=1.10; 95%CI: 1.01 – 1.19). Patients with low physical functioning scores on subsequent visits were also more likely to drop out (SHR=1.15; 95%CI: 1.04 – 1.27). Other functional and symptom subscales of the EORTC QLQ-C30 and other PROMs at baseline and the last visit, on the other hand, were not substantially associated with dropping out early.

3.2.3 Patient characteristics associated with time-to-death

Time-to-death was associated with cancer of malignant neoplasm of the pancreas (HR=2.48; 95%CI: 1.27 – 4.85) compared to acute leukemia and a medium level of education (HR=1.58; 95%CI: 1.02 – 2.45) compared to a high level of education. GH/QoL scores at baseline and at subsequent visits were associated with the risk of premature death (HR=0.88; 95%CI: 0.80 – 0.96 at baseline and HR=0.79; 95%CI: 0.70 – 0.88 at the last observed visit). The results were similar in the functional subscales: physical, role, and social functioning. Furthermore, the worsening of fatigue, nausea and vomiting, and appetite loss were associated with early death. Additionally, the scores at subsequent visits of emotional functioning (HR=0.90; 95%CI: 0.83 – 0.99), dyspnea (HR=1.14; 95%CI: 1.05 – 1.23), and constipation (HR=1.09; 95%CI: 1.004 – 1.19) were associated with premature death. No substantial association between other PROs and premature death was observed.

[7] Sub-hazard ratios (SHR) or hazard ratios (HR) for the EORTC QLQ C-30 change per 10-point. Low values indicate worsening of GH/QoL and functioning, while high values indicate worsening of symptoms. SHR or HR<1 indicates that a high score of GH/QoL or functioning decreases the risk of dropping out early or premature death (1/SHR (or HR) indicates that worsening of GH/QoL or functioning increases the risk of early dropout or premature death), while SHR or HR>1 indicates that worsening of the symptoms increases the risk of early dropout or premature death.

4. Discussion

4.1 Summary of findings

In the OSCAR, there was a significant percentage of missing PROs data. Premature death was the major cause of missing PROs data, followed by dropping out for health-related and non-health-related reasons. As a result, the monotone pattern was the most common missing data pattern in OSCAR. After the initial follow-up visit (3 months), around 20% of the patients dropped out of the study.

Being divorced or widowed and having poor social support were related to early dropout, whereas having cancer of malignant neoplasm of the pancreas and a medium level of education were associated with premature death.

Early dropout and premature death were found to be considerably influenced by the level of PROs data. Having low scores of GH/QoL and role functioning at baseline was related to dropping out early. However, not only PROs data at baseline, but also PROs data at the last visit were associated with early death. This was especially true for worsening scores of GH/QoL and functional subscales such as physical, role, and social functioning, as well as worsening scores on symptom subscales such as fatigue, nausea and vomiting, and appetite loss. Other PROs data such as physician-patient communication, patient preference for decision-making and information seeking, DCS, and HL were not substantially associated with either dropping out early or premature death.

4.2 Implications for statistical analysis methods with missing PROs data

4.2.1 Dropout rate and missing data patterns

This work describes the missing rate and identifies patterns of missing PROs data in a longitudinal study for advanced stages of cancer. Because of the research methodology and severity of disease, the rates and patterns of missing data for PROMs differed. For example, in the intervention group, the social care nurse evaluated patient HRQoL once a month, while the study nurses assessed other PROMs every three months. As a result, PROMs in this group had varying frequencies of missing data. Furthermore, because the social care nurses who enrolled patients assessed their HRQoL at the same time, while the study nurses evaluated the other PROMs afterwards, missing rates in other PROMs may arise in this intervention group at baseline.

The OSCAR was designed to enroll patients with an advanced stage of cancer, and the enrollment was done when patients were admitted to the hospital for therapy, such as chemotherapy and operations. Therefore, high missing rates due to premature death were expected. However, there was no considerable difference in the missing rate of dropouts between study groups, which could have caused biased results for further analysis. Although this is a descriptive result, reporting the missing rate with causes can be used as basic information for a further advanced-stage disease study. When preparing a study, for example, sample sizes should be adjusted with a 50% attrition rate, or shorter time periods between follow-up visits and post-recruitment should be scheduled.
The majority of missing data patterns in the OSCAR were monotone patterns due to early discontinuation and premature death. A small rate of intermittent patterns was reported, but this missing pattern posed no concerns in most cases because it can be ignored (16). The patterns of missing data correspond to statistical techniques for handling their impact on the quality of longitudinal data analysis.

4.2.2 Factors associated with dropout and missing data mechanisms

Identifying factors at baseline and at the last visit prior to dropout could benefit the assumption about the missing data mechanism, as well as selecting an appropriate statistical method for handling the missing PROs data. As our findings presented the relationship between worsening HRQoL (e.g., deterioration of GH/QoL, physical functioning, fatigue scores) and missing PROs data due to dropout and premature death, this suggests that the missing data mechanism is not completely random. Therefore, complete case analysis or last observation carried forward (LOCF) methods should not be applied to an analysis of the OSCAR data.
The findings revealed that various baseline patient characteristics and HRQoL were associated with missing due to dropping out, suggesting that missing due to dropping out is MAR. On the other hand, there was an association between HRQoL at the last visit and premature death. Although it is difficult to distinguish the missing mechanism between MAR and MNAR since the missing data is unobserved (35, 36), this evidence implies that the missing data due to premature death is MNAR. In this case, it is possible that the level of unobserved PROs data is influenced by the unobserved poorer health status.

In summary, patient characteristics and reasons for missing data could be useful for making an appropriate assumption about the underlying missing data mechanisms. Additionally, they could be considered to be included in the model as potential variables when applying imputation methods or performing regression models.

4.2.3 Statistical approaches for analyzing longitudinal PROs with missing data

Mixed-effect models or generalized estimating equation models (GEE) are commonly used in longitudinal data analysis to account for the correlation within and between subjects due to multiple observations (7, 16, 37). When missing data is present, however, statistical methods that ignore it may not be appropriate if the missing data is not completely random. As a result, assumptions about missing data mechanisms are important, especially when using statistical approaches for dealing with missing data based on MAR or MNAR assumptions.

For analyses of missing PROs data that assume MAR, multiple imputation is now commonly used (7, 16, 38) before performing a standard regression model, such as mixed-effect models or GEE. Additionally, patient characteristics and reasons for dropping out could be used as auxiliary data (the variables that use modelling as a covariate) in the models or in the multiple imputation models (7).

The major challenge for statisticians is how to handle the missing PROs data when it might be MNAR. The analysis methods for studies with missing data assume MNAR have increased. However, all the methods cannot be proven and it cannot be demonstrated that they are correct because the missing data is not observed, and all the methods must be based on reasonable assumptions (7, 39). There are statistical approaches to MNAR, such as pattern mixture models, selection models, and shared random effects models (7). Those models have their own assumptions, and some require special software. For example, selection models can be performed in the WinBUGS code. Pattern mixture models can be done using standard software, such as R, SAS, or Stata. Pattern mixture models estimate the PROs outcome based on missing data patterns through standard analyses such as mixed models. Following that, the overall estimates and standard errors are merged in specialized ways based on weighting in each missing pattern (7, 39). A joint model is also known as a shared parameter model. The joint model links the two separate models between time-to-event models (e.g., time-to-dropout or time-to-death) and the longitudinal outcome models by sharing random effects. There are some difficulties in

using these models, as the models fail to estimate the correlation when there is no variation in one of the random effects or when they fail to converge (7).
When missing data is present, performing a sensitivity analysis to investigate the missing data mechanism in various scenarios is always recommended, especially when unsure if the missing data is MNAR (7, 40). The consistency of results across different analysis methods based on the assumptions of missing data mechanisms could allow the researchers to be confident about their conclusions.

4.3 Strengths and Limitations

This work contributes to a better understanding of missing PROs data in a clinical trial for advanced-stage cancer by presenting the missing rates and patterns of missing data, as well as possible patient characteristics associated with dropout and premature death. This research presents the application of a time-to-event analysis when competing events have to be accounted for.
However, the OSCAR study lacked data on disease severity, such as cancer stage or the Eastern Cooperative Oncology Group (ECOG) performance status. Life expectancy was not determined prior to enrolment, and longer life expectancy was not a requirement for participation in this study, resulting in a high mortality rate during the study period.
Because multivariable analysis was not possible due to multi-collinearity within the PROs (e.g., high correlation within the EORTC QLQ-C30, between PROs at baseline, and prior to dropout), only bivariate models were used to examine how patient characteristics were associated with time-to-dropout or time-to-death. The findings were interpreted based on the effect size and 95% confidence interval, with no correction for multiple testing made.

5. Conclusions

A high rate of missing PROs data was present in the OSCAR study. Premature death was the most common reason for missing PROs data, followed by dropping out for health-related and non-health-related reasons. Some level of MNAR should be expected in advanced-stage cancer studies, and the investigation of patient characteristics associated with missing data due to dropout might reduce the degree of unexplained missing data. Before using statistical methodologies to deal with missing PROs data, the assumptions of the missing data mechanism should be investigated, and sensitivity analyses should always be undertaken.

Reference list

1. Kyte D, Retzer A, Ahmed K, Keeley T, Armes J, Brown JM, Calman L, Gavin A, Glaser AW, Greenfield DM, Lanceley A, Taylor RM, Velikova G, Brundage M, Efficace F, Mercieca-Bebber R, King MT, Turner G, Calvert M. Systematic Evaluation of Patient-Reported Outcome Protocol Content and Reporting in Cancer Trials. J Natl Cancer Inst. 2019;111(11):1170-8.

2. Basch E, Abernethy AP, Mullins CD, Reeve BB, Smith ML, Coons SJ, Sloan J, Wenzel K, Chauhan C, Eppard W, Frank ES, Lipscomb J, Raymond SA, Spencer M, Tunis S. Recommendations for incorporating patient-reported outcomes into clinical comparative effectiveness research in adult oncology. J Clin Oncol. 2012;30(34):4249-55.

3. Basch E, Deal AM, Kris MG, Scher HI, Hudis CA, Sabbatini P, Rogak L, Bennett AV, Dueck AC, Atkinson TM, Chou JF, Dulko D, Sit L, Barz A, Novotny P, Fruscione M, Sloan JA, Schrag D. Symptom Monitoring With Patient-Reported Outcomes During Routine Cancer Treatment: A Randomized Controlled Trial. J Clin Oncol. 2016;34(6):557-65.

4. Karsten MM, Speiser D, Hartmann C, Zeuschner N, Lippold K, Kiver V, Gocke P, Kirchberger V, Blohmer JU. Web-Based Patient-Reported Outcomes Using the International Consortium for Health Outcome Measurement Dataset in a Major German University Hospital: Observational Study. JMIR Cancer. 2018;4(2):e11373.

5. Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, Filiberti A, Flechtner H, Fleishman SB, de Haes JC, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst. 1993;85(5):365-76.

6. Lorini C, Santomauro F, Grazzini M, Mantwill S, Vettori V, Lastrucci V, Bechini A, Boccalini S, Bussotti A, Bonaccorsi G. Health literacy in Italy: a cross-sectional study protocol to assess the health literacy level in a population-based sample, and to validate health literacy measures in the Italian language. BMJ Open. 2017;7(11):e017812.

7. Bell ML, Fairclough DL. Practical and statistical issues in missing data for longitudinal patient-reported outcomes. Stat Methods Med Res. 2014;23(5):440-59.

8. Bell ML, Kenward MG, Fairclough DL, Horton NJ. Differential dropout and bias in randomised controlled trials: when it matters and when it may not. BMJ : British Medical Journal. 2013;346:e8668.

9. Rubin DB. Inference and missing data. Biometrika. 1976;63(3):581-92.

10. Carpenter JR KM. Introduction. Multiple Imputation and its Application. 1st ed: John Wiley & Sons, Ltd; 2013. p. 1-36.

11. Mercieca-Bebber R, King MT, Calvert MJ, Stockler MR, Friedlander M. The importance of patient-reported outcomes in clinical trials and strategies for future optimization. Patient Relat Outcome Meas. 2018;9:353-67.

12. Mercieca-Bebber R, Friedlander M, Calvert M, Stockler M, Kyte D, Kok P-S, King MT. A systematic evaluation of compliance and reporting of patient-reported outcome endpoints in ovarian cancer randomised controlled trials: implications for generalisability and clinical practice. J Patient Rep Outcomes. 2017;1(1):5-.

13. Pe M, Dorme L, Coens C, Basch E, Calvert M, Campbell A, Cleeland C, Cocks K, Collette L, Dirven L, Dueck AC, Devlin N, Flechtner HH, Gotay C, Griebsch I, Groenvold M, King M, Koller M, Malone DC, Martinelli F, Mitchell SA, Musoro JZ, Oliver K, Piault-Louis E, Piccart M, Pimentel FL, Quinten C, Reijneveld JC, Sloan J, Velikova G, Bottomley A, Setting International Standards in Analyzing Patient-Reported O, Quality of Life Endpoints Data C. Statistical analysis of patient-reported outcome data in randomised controlled trials of locally advanced and metastatic breast cancer: a systematic review. Lancet Oncol. 2018;19(9):e459-e69.

14. Mercieca-Bebber R, Palmer MJ, Brundage M, Calvert M, Stockler MR, King MT. Design, implementation and reporting strategies to reduce the instance and impact of missing patient-reported outcome (PRO) data: a systematic review. BMJ Open. 2016;6(6):e010938.

15. Curran D, Bacchi M, Schmitz SF, Molenberghs G, Sylvester RJ. Identifying the types of missingness in quality of life data from clinical trials. Stat Med. 1998;17(5-7):739-56.

16. DeSouza CM, Legedza AT, Sankoh AJ. An overview of practical approaches for handling missing data in clinical trials. J Biopharm Stat. 2009;19(6):1055-73.

17. Renovanz M, Hechtner M, Kohlmann K, Janko M, Nadji-Ohl M, Singer S, Ringel F, Coburger J, Hickmann AK. Compliance with patient-reported outcome assessment in glioma patients: predictors for drop out. Neurooncol Pract. 2018;5(2):129-38.

18. Roick J, Danker H, Kersting A, Briest S, Dietrich A, Dietz A, Einenkel J, Papsdorf K, Lordick F, Meixensberger J, Mossner J, Niederwieser D, Prietzel T, Schiefke F, Stolzenburg JU, Wirtz H, Singer S. Factors associated with non-participation and dropout among cancer patients in a cluster-randomised controlled trial. Eur J Cancer Care (Engl). 2018;27(1).

19. Spiers S, Oral E, Fontham ETH, Peters ES, Mohler JL, Bensen JT, Brennan CS. Modelling attrition and nonparticipation in a longitudinal study of prostate cancer. BMC Med Res Methodol. 2018;18(1):60.

20. Marcus MW, Raji OY, Chen Y, Duffy SW, Field JK. Factors associated with dropout in a lung cancer highrisk cohort--the Liverpool lung project. Int J Oncol. 2014;44(6):2146-52.

21. Ramsey I, de Rooij BH, Mols F, Corsini N, Horevoorts NJE, Eckert M, van de Poll-Franse LV. Cancer survivors who fully participate in the PROFILES registry have better health-related quality of life than those who drop out. J Cancer Surviv. 2019;13(6):829-39.

22. Frick J, Schindel D, Gebert P, Grittner U, Schenk L. Improving quality of life in cancer patients through higher participation and health literacy: study protocol for evaluating the oncological social care project (OSCAR). BMC Health Serv Res. 2019;19(1):754.

23. Lampert T, Hoebel J, Kuntz B, Müters S, Kroll LE. Messung des sozioökonomischen Status und des subjektiven sozialen Status in KiGGS Welle 2. Robert Koch-Institut, Epidemiologie und Gesundheitsberichterstattung; 2018.

24. Kocalevent RD, Berg L, Beutel ME, Hinz A, Zenger M, Harter M, Nater U, Brahler E. Social support in the general population: standardization of the Oslo social support scale (OSSS-3). BMC Psychol. 2018;6(1):31.

25. Gebert P, Schindel D, Frick J, Schenk L, Grittner U. Characteristics and patient-reported outcomes associated with dropout in severely affected oncological patients: an exploratory study. BMC Med Res Methodol. 2021;21(1):77.

26. Brenk-Franz K, Hunold G, Galassi JP, Tiesler F, Herrmann W, Freund T, Steurer-Stey C, Djalali S, Sönnichsen A, Schneider N, Gensichen J. Quality of the physician-patient relationship - Evaluation of the German version of the Patient Reactions Assessment (PRA-D). 2016;92:103-8.

27. Gaab J BS, Sprott H, Ehlert U. Illness Perception Questionnaire (deutsche Version). Diagnostische Verfahren in der Rehabilitation. 2008:109-11.

28. Härter M, Simon D, Kriston L. Die deutsche modifizierte Fassung des Autonomie-Präferenz-Index (API-Dm). Klin Diagn Eval. 2011;4((1E)):5-14.

29. Buchholz A, Hölzel L, Kriston L, Simon D, Härter M. Die Decisional Conflict Scale in deutscher Sprache (DCS-D)-Dimensionale Struktur in einer Stichprobe von Hausarztpatienten. Klinische Diagnostik und Evaluation. 2011;4:15-30.

30. Huebner M, Wolkewitz M, Enriquez-Sarano M, Schumacher M. Competing risks need to be considered in survival analysis models for cardiovascular outcomes. J Thorac Cardiovasc Surg. 2017;153(6):1427-31.

31. Fine JP, Gray RJ. A Proportional Hazards Model for the Subdistribution of a Competing Risk. Journal of the American Statistical Association. 1999;94(446):496-509.

32. Lau B, Cole SR, Gange SJ. Competing risk regression models for epidemiologic data. Am J Epidemiol. 2009;170(2):244-56.

33. Alipour A, Shokri A, Yasari F, Khodakarim S. Introduction to Competing Risk Model In the Epidemiological Research. International Journal of Epidemiologic Research. 2018;5(3):98-102.

34. Austin PC, Fine JP. Practical recommendations for reporting Fine-Gray model analyses for competing risk data. Stat Med. 2017;36(27):4391-400.

35. Carpenter JR, Kenward MG. Missing data in randomised controlled trials: a practical guide2007.

36. Hughes RA, Heron J, Sterne JAC, Tilling K. Accounting for missing data in statistical analyses: multiple imputation is not always the answer. Int J Epidemiol. 2019;48(4):1294-304.

37. Twisk JW, Smidt N, de Vente W. Applied analysis of recurrent events: a practical overview. J Epidemiol Community Health. 2005;59(8):706-10.

38. Jakobsen JC, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials - a practical guide with flowcharts. BMC Med Res Methodol. 2017;17(1):162.

39. Fairclough DL, Peterson HF, Chang V. Why are missing quality of life data a problem in clinical trials of cancer therapy? Statistics in Medicine. 1998;17(5-7):667-77.

40. Donaldson GW, Moinpour CM. Learning to live with missing quality-of-life data in advanced-stage disease trials. J Clin Oncol. 2005;23(30):7380-4.

Statutory Declaration

“I, Pimrapat Gebert, by personally signing this document in lieu of an oath, hereby affirm that I prepared the submitted dissertation on the topic “Missing values of patient-reported outcome data in a longitudinal study of advanced and metastatic cancer patients (Fehlende Werte bei Patienten-berichteten Endpunkten in einer Längsschnittstudie bei fortgeschrittenem und metastasiertem Krebs)”, independently and without the support of third parties, and that I used no other sources and aids than those stated.
All parts which are based on the publications or presentations of other authors, either in letter or in spirit, are specified as such in accordance with the citing guidelines. The sections on methodology (in particular regarding practical work, laboratory regulations, statistical processing) and results (in particular regarding figures, charts and tables) are exclusively my responsibility.

Furthermore, I declare that I have correctly marked all of the data, the analyses, and the conclusions generated from data obtained in collaboration with other persons, and that I have correctly marked my own contribution and the contributions of other persons (cf. declaration of contribution). I have correctly marked all texts or parts of texts that were generated in collaboration with other persons.

My contributions to any publications to this dissertation correspond to those stated in the below joint declaration made together with the supervisor. All publications created within the scope of the dissertation comply with the guidelines of the ICMJE (International Committee of Medical Journal Editors; http://www.icmje.org) on authorship. In addition, I declare that I shall comply with the regulations of Charité – Universitätsmedizin Berlin on ensuring good scientific practice.

I declare that I have not yet submitted this dissertation in identical or similar form to another Faculty.

The significance of this statutory declaration and the consequences of a false statutory declaration under criminal law (Sections 156, 161 of the German Criminal Code) are known to me.”

Date Signature

Declaration of your own contribution to the publications

Pimrapat Gebert contributed the following to the below listed publication:

Publication 1:

Gebert P*, Schindel D, Frick J, Schenk L, Grittner U. Characteristics and patient-reported outcomes associated with dropout in severely affected oncological patients: an exploratory study. BMC Med Res Methodol. 2021;21(1):77. Epub 2021/04/22. doi: 10.1186/s12874-021-01259-0.

*PG is the sole first author and corresponding author

Contribution in detail:

PG has contributed to data cleaning and preparation, reviewing literatures, performing the statistical analysis, interpretation of the results, preparation of all figures and tables, and writing an original draft preparation.

JF and DS participated in the study design, conducting critical review, and editing.

LS participated in the study design, conducting critical review, and supervising the study.

UG was involved in the statistical analysis plan, verifying the analytical methods, conducting critical review, and supervising the study.

_____

Signature of doctoral candidate

Excerpt from Journal Summary List

Journal Data Filtered By: **Selected JCR Year: 2019** Selected Editions: SCIE,SSCI
Selected Categories: **"HEALTH CARE SCIENCES and SERVICES"**
Selected Category Scheme: WoS
Gesamtanzahl: 102 Journale

Rank	Full Journal Title	Total Cites	Journal Impact Factor	Eigenfactor Score
1	BMJ Quality & Safety	5,825	6.084	0.015370
2	Implementation Science	10,777	5.531	0.020190
3	ACADEMIC MEDICINE	17,605	5.354	0.028860
4	HEALTH AFFAIRS	17,516	5.331	0.047430
5	JOURNAL OF MEDICAL INTERNET RESEARCH	16,349	5.034	0.029410
6	JOURNAL OF CLINICAL EPIDEMIOLOGY	28,878	4.952	0.028410
7	VALUE IN HEALTH	10,040	4.748	0.017370
8	JOURNAL OF GENERAL INTERNAL MEDICINE	20,229	4.597	0.026960
9	MEDICAL EDUCATION	10,598	4.570	0.011180
10	Journal of Personalized Medicine	617	4.433	0.001950
11	JMIR mHealth and uHealth	4,226	4.313	0.010020
12	MILBANK QUARTERLY	3,822	4.195	0.004300
13	JOURNAL OF THE AMERICAN MEDICAL INFORMATICS ASSOCIATION	9,959	4.112	0.017380
14	International Journal of Health Policy and Management	1,463	3.821	0.004860
15	PALLIATIVE MEDICINE	5,413	3.739	0.008460
16	PHARMACOECONOMICS	5,150	3.563	0.009120
17	Internet Interventions-The Application of Information Technology in Mental and Behavioural Health	996	3.513	0.002720
18	HEALTH TECHNOLOGY ASSESSMENT	5,573	3.370	0.009440
19	Patient-Patient Centered Outcomes Research	1,204	3.226	0.003990

Rank	Full Journal Title	Total Cites	Journal Impact Factor	Eigenfactor Score
20	MEDICAL CARE RESEARCH AND REVIEW	2,459	3.212	0.003300
21	MEDICAL CARE	20,261	3.210	0.019020
22	JOURNAL OF PAIN AND SYMPTOM MANAGEMENT	10,897	3.077	0.014840
23	JOURNAL OF MEDICAL SYSTEMS	5,695	3.058	0.007050
24	BMC Medical Research Methodology	11,581	3.031	0.018590
24	Journal of Patient Safety	1,094	3.031	0.002310
26	INTERNATIONAL JOURNAL OF MEDICAL INFORMATICS	5,368	3.025	0.007110
27	Journal of Managed Care & Specialty Pharmacy	1,667	3.021	0.005780
28	HEALTH EXPECTATIONS	3,600	3.008	0.008230
29	Health Informatics Journal	981	2.932	0.001530
30	JOURNAL OF HEALTH ECONOMICS	7,404	2.827	0.014020
31	QUALITY OF LIFE RESEARCH	14,492	2.773	0.018650
32	International Journal of Integrated Care	1,245	2.753	0.001750
33	HEALTH POLICY AND PLANNING	5,413	2.704	0.010540
34	BMJ Supportive & Palliative Care	1,309	2.681	0.003390
35	JOURNAL OF RURAL HEALTH	2,005	2.667	0.003100
36	MEDICAL TEACHER	8,633	2.654	0.009410
37	SUPPORTIVE CARE IN CANCER	12,842	2.635	0.021660
38	JOURNAL OF TELEMEDICINE AND TELECARE	2,703	2.616	0.003430
39	ADVANCES IN HEALTH SCIENCES EDUCATION	2,696	2.480	0.004210
40	Applied Health Economics and Health Policy	1,234	2.442	0.003140
41	Risk Management and Healthcare Policy	549	2.429	0.001530

Selected JCR Year: 2019; Selected Categories: "HEALTH CARE SCIENCES and SERVICES"

Printing copy of the publication

Gebert P, Schindel D, Frick J, Schenk L, Grittner U. Characteristics and patient-reported outcomes associated with dropout in severely affected oncological patients: an exploratory study. BMC Med Res Methodol. 2021;21(1):77. Epub 2021/04/22. doi: 10.1186/s12874-021-01259-0. PubMed PMID: 33879087; PubMed Central PMCID: PMCPMC8059010.

DOI Link: doi: 10.1186/s12874-021-01259-0

Impact Factor

4.615 - 2-year Impact Factor

5.260 - 5-year Impact Factor

Gebert et al. BMC Medical Research Methodology (2021) 21:77
https://doi.org/10.1186/s12874-021-01259-0

BMC Medical Research Methodology

RESEARCH ARTICLE Open Access

Characteristics and patient-reported outcomes associated with dropout in severely affected oncological patients: an exploratory study

Pimrapat Gebert[1,2*], Daniel Schindel[3], Johann Frick[3], Liane Schenk[3†] and Ulrike Grittner[1,2†]

Abstract

Background: Patient-reported outcome measures (PROMs) are commonly-used surrogates for clinical outcomes in cancer research. When researching severe diseases such as cancer, it is difficult to avoid the problem of incomplete questionnaires from drop-outs or missing data from patients who pass away during the observation period. The aim of this exploratory study was to explore patient characteristics and the patient-reported outcomes associated with the time-to-dropout.

Methods: In an Oncological Social Care Project (OSCAR) study, the condition of the participants was assessed four times within 12 months (t0: baseline, t1: 3 months, t2: 6 months, and t3: 12 months) by validated PROMs. We performed competing-risk regressions based on Fine and Gray's proportional sub-distribution hazards model for exploring factors associated with time-to-dropout. Death was considered a competing risk.

Results: Three hundred sixty-two participants were analyzed in the study. 193 (53.3%) completed a follow-up after 12 months, 67 (18.5%) patients dropped out, and 102 patients (28.2%) died during the study period. Poor subjective social support was related to a higher risk of drop-out (SHR = 2.10; 95%CI: 1.01–4.35). Lower values in health-related quality of life were related to drop-out and death. The sub-scales global health status/QoL, role functioning, physical functioning, and fatigue symptom in the EORTC QLQ-C30 were key characteristics of early drop-out.

Conclusion: Severely affected cancer patients with poor social support and poor quality of life seem more likely to drop out of studies than patients with higher levels of social support and a better quality of life. This should be considered when planning studies to assess advanced cancer patients. Methods of close continued monitoring should be actively used when patient experiences a substantial deterioration in their health-related quality of life and symptoms during the study. Results for such studies have to be interpreted with caution in light of specific drop-out mechanisms.

(Continued on next page)

* Correspondence: pimrapat.gebert@charite.de
†Liane Schenk and Ulrike Grittner contributed equally to this work.
[1]Berlin Institute of Health at Charité –Universitätsmedizin Berlin, Charitéplatz 1, 10117 Berlin, Germany
[2]Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Institute of Biometry and Clinical Epidemiology, Charitéplatz 1, 10117 Berlin, Germany
Full list of author information is available at the end of the article

(Continued from previous page)

Trial registration: OSCAR study was registered to the German Clinical Trials Register (DRKS-ID: DRKS00013640). Registered 29 December 2017.

Keywords: Cancer, Attrition, Monotone missing data, Non-compliance, Patient-reported outcome measures, Health-related quality of life, Advanced cancer

Background

Patient-reported outcome measures (PROMs) are tools for assessing a patient's physical and emotional wellbeing, satisfaction with care, symptoms, and quality of life (QoL) [1]. Patient-reported outcomes (PROs) are usually measured through questionnaires which combine several items into sub-scales or total scales. Oncology research often focuses on PROs as primary outcomes [2], and repeatedly measured PROs are typically observed for exploring and monitoring the change in the health status of cancer patients [3, 4]. As cancer patients are often severely affected by the disease, missing data from drop-outs due to deterioration in health or death are common [5].

Drop-out occurs in a longitudinal study when a participant discontinues the study completely. Rates of drop-out vary from 30 to 50% [6–8] in oncology studies, but reasons for dropping out are not usually recorded [7]. In cancer research, numerous factors have been identified as being related to dropping out. Being older, male, unmarried, having a low level of education and a low economic status are all associated with early dropout. However, the relevance of some of these factors is less consistent than others: some studies show that women are more likely to drop out, for example [6, 9]. Generally, symptom burdens and health conditions are the main factors related to discontinuing a study [6, 7, 10].

Higher rates of drop-out not only result in reduced statistical power, but also cause biased results if subpopulations are over or under-represented in the remaining sample [10]. Knowledge of a patient's characteristics related to the risk of drop-out will allow for the application of strategies for the minimization of data loss, such as continuous monitoring, reminders, and the use of modern technology (e.g. mobile apps and online questionnaires) to measure data [11] will result in it being more complete and reliable, as is the case with estimation procedures for missing data.

Investigating patient drop-out characteristics in cancer research is useful for study planning with regard to the estimation of sample size, more appropriate definitions of patient inclusion criteria, decreasing insufficient enrollment through improved inclusion criteria, increasing patient retention over study periods, improved monitoring and possible post-recruitment, and purposing appropriate statistical analysis for challenging missing data. The aim of this exploratory study was to assess patient characteristics and the patient-reported outcomes associated with time-to-dropout when accounting for death as a competing risk.

Methods

Study population

A protocol for the Oncological Social Care Project (OSCAR) has been reported previously [12]. In short, the OSCAR was developed as an intervention in oncological care by the German company health insurance fund Pronova BKK. A non-randomized, controlled, multi-center intervention study was conducted at three study sites in Germany from January 2018 to February 2020. Three hundred sixty-two participants above the age of 18 with different cancer types were included (see the inclusion criteria in the original protocol [12]). The study recruited severely affected oncological patients with advanced cancer stages, such as metastasized colorectal cancer, malignant neoplasm of the pancreas, lymphoma, or multiple myeloma and malignant plasma cell neoplasms. These patients had a high symptom burden and needed intensive supportive care. One hundred fifty patients in the intervention group and 212 patients in the control group were studied. Patients answered a monthly health-related QoL questionnaire using the European Organization for Research and Treatment of Cancer's Quality of Life Questionnaire (EORTC QLQ-C30) in the intervention group. The EORTC QLQ-C30 questionnaire for the control group and other PROMs for both groups were assessed at baseline (t0), 3 months (t1), 6 months (t2), and 12 months (t3). Refusal to participate was documented for each follow-up visit, and those who dropped out were asked whether they wanted to discontinue the study because of their health or for other reasons. In the following analysis, we will focus on drop-out (health-related or other reasons) and death.

Data collection

Demographic data was collected at baseline, e.g. age, sex, time since diagnosis, cancer diagnosis, family status, and education. Education and professional qualifications were classified based on the Comparative Analysis of Social Mobility in Industrial Nations (CASMIN) classification of education [13]. Assessment of subjective social support was based on the Oslo Three-Item Social Support Scale (OSSS-3) [14], where the total score ranges

from 3 to 14 points, and is classified as poor (3 to 8), moderate (9 to 11), or strong (12 to 14).

The following five PROMs were assessed in this study:

- The EORTC QLQ-C30 (version 3.0) is a generic tool for assessing the quality of life (QoL) for various cancer patients, and is provided by the European Organisation for Research and Treatment of Cancer [15]. It consists of 30 questions and incorporates nine multi-item scales: five functional scales (physical, role, cognitive, emotional, and social functioning); three symptom scales (fatigue, pain, and nausea/vomiting); six single item scales (dyspnea, insomnia, appetite loss, constipation, diarrhea, and financial difficulties); and a global health status/QoL scale. The total score is calculated by averaging items within scales and transforming them from 0 to 100. Higher values of a functional scale represent a high or healthy level of functioning, while higher values of a symptom scale or item represent a high level of symptomatology or problems.
- The Patient Reaction Assessment (PRA-D) [16] is an instrument used for assessing the perceived quality of the doctor-patient relationship. In OSCAR, we modified five questions using five-point instead of seven-point Likert scales. The total score was transformed using the formula (y = 1.5 * x - 0.5) [17], and ranges from 5 to 35. Higher values indicate better doctor-patient relationships.
- The modified German version of the Autonomy Preference Index (API-DM) [18] consists of two preferences: decision making and information seeking. The decision making preferences consist of four items, and information seeking preferences consist of seven. The total score for each preference is transformed to achieve scores ranging from 0 to 100. Higher values indicate a greater preference for decision making and information seeking.
- The Decision Conflict Scale (DCS) [19] is a self-reported questionnaire for evaluating decision conflicts and comprises ten items, with the sum score ranging from 0 to 100. Higher values indicate greater decision conflicts.
- The European Health Literacy Survey (HLS-EU-Q6) [20] is a shortened, six-item version of the European Health Literacy Survey (HLS-EU) for assessing health literacy (HL). The sum score is averaged and grouped into insufficient (≤2 scores), problematic (2–3 scores), and sufficient (≥3 scores) health literacy.

Statistical analysis

The study outcome was discontinuation of the study due to either dropout or death. Baseline characteristics were presented separately by participants who completed the study or who dropped out. For continuous variables, mean and standard deviation (SD) and median and interquartile ranges (IQR) are presented, depending on their distribution. For categorical data, absolute and relative frequencies were reported.

Demographic data at baseline, the PROMs at baseline and at the visit before dropping out or death were used for exploring the association with time-to-dropout. Time-to-event was defined from the enrollment date to date of death, date of drop-out, or censored at 12 months. Date of drop-out was defined as when the participant refused to continue the study or the first date of unsuccessful contact try, when the participant could not be reached after trying to contact them three times. Fine and Gray's proportional sub-distribution hazards models were performed, assuming that death was a competing risk. Cox regression models were used for assessing the association between participant characteristics and death. Multinomial logistic regression models were used for comparing characteristics related to drop-out or death as a sensitivity analysis.

Statistical testing was done within an exploratory framework at a two-sided significance level of $\alpha = 0.05$ without adjustment for multiple testing. Due to multicollinearity within PROs (such as a high correlation within the EORTC QLQ-C30 between global health status/QoL and subscales of functional or symptoms) and PROs between baseline and visit prior to drop-out, restricted by the number of observations in some variables (such as family status and HLS-EU-Q6), we only used bivariate models. All the statistical tests were performed using Stata IC15 (StataCorp, 2017, College Station, TX, USA).

Results

Participant characteristics and dropout rate

Three hundred sixty-two participants were analyzed in the study. One hundred ninety-three patients (53.3%) completed a follow-up at 12 months, 102 died during follow-up (28.2%), and 67 dropped out the study (18.5%) (Additional file 1: Table S1). Rates of drop-out and death combined were 14.4% at 3 months, 31.5% at 6 months, and 46.8% at 12 months (Fig. 1). The participants who dropped out or died were older than the participants who completed the study. However, there was almost no age difference between the participants who dropped out and those who died. The proportion of patients who died was higher in participants with malignant neoplasm of the bronchus and lungs (22.6%) and malignant neoplasm of the pancreas (15.7%) when compared to participants who dropped out and compliant participants (Additional file 1: Table S1). Participants who dropped out or died had lower values in global health status/

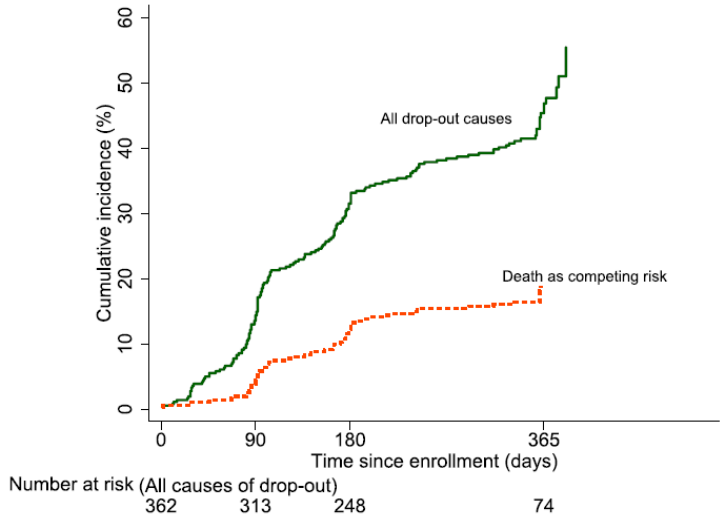


Fig. 1 Cumulative incidence of all drop-out causes (green solid line) and drop-out with death as a competing risk (orange dotted line)

QoL, physical functioning, and role functioning at baseline than participants who completed the study (Additional file 1: Table S2).

Characteristics associated with drop-out or death during follow-up

In our study, family status and poor subjective social support were related to drop-out (Table 1). Lower global health status/QoL and role functioning of the EORTC QLQ-C30 at baseline were associated with a higher risk of drop-out, as a difference in global health status/QoL and role functioning by 10 points resulted in a 12% (95%CI: 1–21%) and 9% (95%CI: 1–16%) higher risk of drop-out. In addition, low physical functioning of the EORTC QLQ-C30 at the visit before drop-out was associated with a higher likelihood of drop-out when considering death as a competing risk (Fig. 2).

Characteristics associated with shorter time to death were malignant neoplasm of the pancreas when compared to acute leukemia (HR = 2.48; 95%CI: 1.27–4.85), low levels of EORTC QLQ-C30 at baseline, low global health status/QoL, low physical functioning, low role functioning, and poorer outcomes on the symptom scale (fatigue, nausea and vomiting, dyspnea, and appetite loss) at the visit before death (Table 1 and Fig. 3). A poor doctor-patient relationship at baseline was also associated with a shorter time to death. There were also study site differences for time to death. Sensitivity analyses showed similar results, meaning that study sites, family status, and the EORTC QLQ-C30 were important factors for both drop-out and death (Additional file 1: Table S3–5).

The trajectories of the EORTC QLQ-C30 over time were substantially different for participants who dropped out, passed away during the study time, and those who completed the study visits (Fig. 4a and b). It is clear that the QoL, functionals and symptoms values decreased before the participant missed the visit due to death, whereas these values did not decrease in those who dropped out. However, patients who have low baseline values in global health status/QoL and poorer functionalities of the EORTC QLQ-C30 were more likely to drop out or die early.

Discussion

The purpose of this study was to investigate the patient characteristics and PROs associated with drop-out and death in a non-randomized intervention study for severely ill cancer patients. Family status, subjective social support, low values of global health status/QoL and role functioning of the EORTC QLQ-C30 at baseline and low value of physical functioning at the visit before drop-out were associated with time-to-dropout, while the study site, diagnosis, low values of global health status/QoL, physical functioning and role functioning, strong symptoms (e.g. fatigue, nausea, vomiting or appetite loss) of the EORTC QLQ-C30 at both baseline and the visit before death were associated with time-to-death.

Drop-out rate in the OSCAR study

Around 50% of the subjects completed all the study visits in OSCAR, meaning that there was a relatively high drop-out rate in our study when compared to prior oncological studies [6, 7, 9, 21, 22]. Surprisingly, there was no marked differential drop-out between intervention and control groups, despite our expectation of lower drop-out rates for the control group because the

Table 1 Participants' demographic data as characteristics associated with time-to-dropout, with early death as a competing risk

	n	Time-to-dropout with death as a competing risk		Death	
		SHR	(95%CI)	HR	(95%CI)
Study group					
Intervention	150	1.23	(0.76, 1.99)	1.26	(0.84, 1.88)
Control	212	1		1	
Age (years)	362	1.01	(0.99, 1.03)	1.02	(1.00, 1.03)
Study site					
Study site 1	119	1			
Study site 2	98	1.77	(0.94, 3.33)	1.69	(0.99, 2.89)
Study site 3	145	1.60	(0.87, 2.93)	1.99	(1.22, 3.26)
Sex					
Male	219	1.35	(0.82, 2.24)	0.75	(0.51, 1.11)
Female	143	1		1	
Family status					
Married	226	1.15	(0.50, 2.67)	1.68	(0.84, 3.38)
Single	44	1		1	
Divorced/Widowed	63	2.71	(1.12, 6.56)	0.89	(0.36, 2.20)
Time since diagnosis					
≤ 6 months	187	1		1	
7–12 months	49	0.43	(0.17, 1.08)	1.06	(0.58, 1.93)
13–24 months	46	0.54	(0.22, 1.32)	1.53	(0.87, 2.67)
> 24 months	80	0.97	(0.54, 1.72)	1.33	(0.81, 2.19)
Diagnosis					
Acute leukemia	69	1		1	
Aggressive lymphoma	58	1.53	(0.64, 3.65)	0.55	(0.25, 1.23)
Malignant neoplasm of the bronchus and lungs	62	1.92	(0.84, 4.38)	1.78	(0.97, 3.27)
Metastatic colorectal cancer/colon carcinoma	78	1.39	(0.60, 3.21)	1.19	(0.64, 2.21)
Malignant neoplasm of the pancreas	32	1.94	(0.79, 4.77)	2.48	(1.27, 4.85)
Multiple myeloma and malignant plasma cell neoplasms	24	1.24	(0.40, 3.83)	0.56	(0.19, 1.65)
Metastasized malignant neoplasm of the breast	9	0.91	(0.11, 7.21)	0.41	(0.05, 3.10)
Others	30	2.13	(0.79, 5.74)	1.32	(0.58, 3.03)
Education					
Low	29	1.85	(0.90, 3.81)	0.92	(0.39, 2.15)
Medium	101	1.27	(0.72, 2.25)	1.58	(1.02, 2.45)
High	206	1		1	
Social support (OSSS-3)					
Poor (3–8)	35	2.10	(1.01, 4.35)	0.83	(0.39, 1.79)
Moderate (9–11)	158	1.32	(0.74, 2.36)	0.87	(0.56, 1.36)
Strong (12–14)	139	1		1	

n Number of observation, *SHR* Sub-hazard ratio, *HR* Hazard ratio, *CI* Confidence Interval

participants would want to benefit from intervention. We found slightly higher rates of drop-out and death in the intervention group than in the control group (Additional file 1: Figure S1). The main reason for not completing the study was premature death, accounting for approximately 60% (102/169) of patient discontinuation. However, the overall attrition (drop-out) rate of 18.5% in the OSCAR (withdrawal, 15.2%; loss-to-follow up, 2.8%; other reasons, 0.5%) was modest when compared to rates reported in other oncological studies, which ranged from 18 to 31% [6, 9, 23]. A high rate of premature death was expected in this population of severely

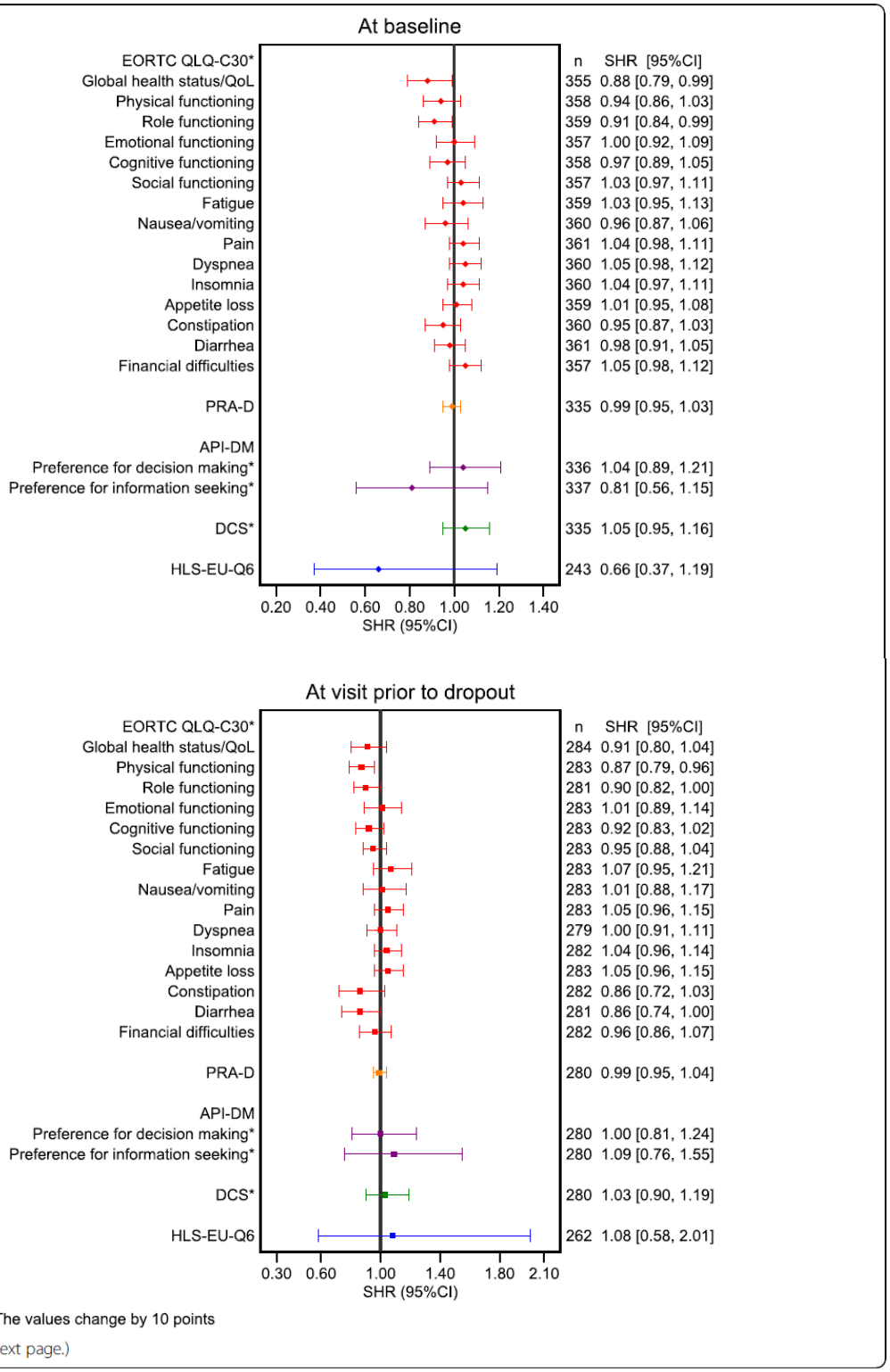


Fig. 2 (See legend on next page.)

(See figure on previous page.)
Fig. 2 Patient-reported outcomes at baseline and at the visit prior to drop out, and their association to time-to-dropout, with death as a competing risk. *n* = the observation number, SHR = Sub-hazard ratio, CI=Confidence interval, High values of global health status/QoL and functional sub-scales in the EORTC QLQ-C30 indicate a better outcome, High values of symptom sub-scales (e.g fatigue, nausea/vomiting, pain etc.) and DCS indicate a poorer outcome

affected cancer patients, in line with the study's aim of including severely affected cancer patients.

Patient characteristics associated with drop-out

Our findings show that family status (being divorced or widowed) and a lack of subjective social support were positively associated with early drop out. Similar findings were observed among cancer patients in a cluster-randomized controlled trial [9]. However, other studies could not confirm this association [6, 7]. Additionally, a lack of social support was associated with early drop-out among cancer patients. It has been reported that the family status plays a role as social support and has a positive effect on a patient's health, quality of life, and coping behavior [24]. Therefore, a lack of social support from family or friends might be one explanation for a lack of desire to continue the study.

In our study, we did not find an association between general characteristics (e.g age, gender and education) and drop-out. Older age was not strongly related to drop-out or death in our study, although there was a weak association between older age and early death. This result is in line with previous cancer studies [6, 7, 21], though some studies have reported contrary findings [9, 10]. Males seemed more likely to drop out early in our study, though this association was reversed in some other studies [22, 23]. Other studies found no association between sex and the probability of dropping out [7, 9]. Participants with lower educational status dropped out early more often, but this association was weak in our study. Similarly, Spiers et al. and Roick et al. found that low education was associated with the lack of a follow-up [9, 10].

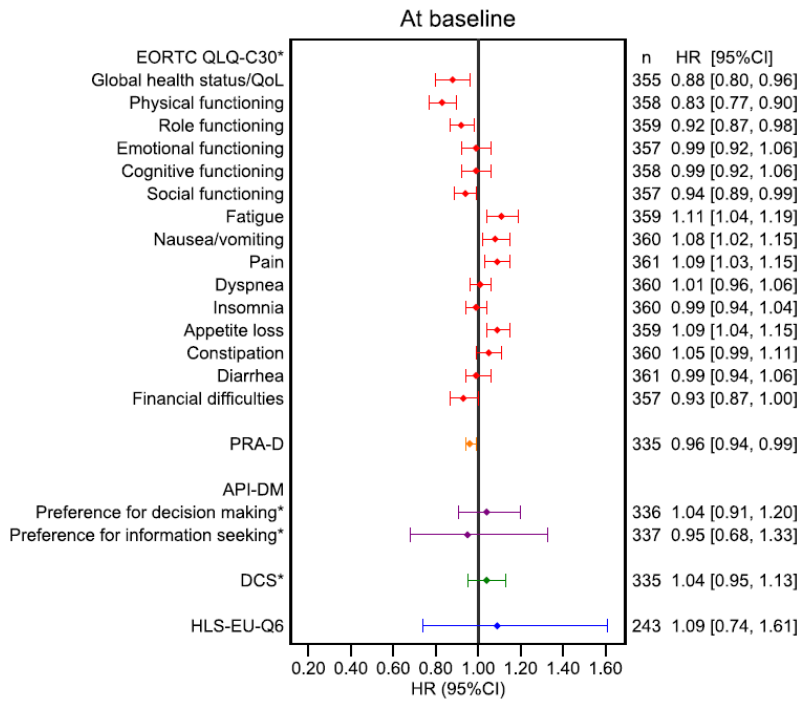
The probability of early death differed between study sites and could have been related to the difference in distribution of diagnosis between the study sites. We found that 70% of study site 3's cancer patients had malignant neoplasm of the bronchus and lungs, metastatic colorectal cancer or colon carcinoma, or malignant neoplasm of the pancreas, while study site 1 enrolled around 15% of these cancer types and about 60% with acute leukemia or aggressive lymphoma (data not shown). It is relevant to our findings that patients who were diagnosed with malignant neoplasm of the pancreas had a higher risk of early death. The prognosis for pancreatic cancer is generally poor, meaning that the overall survival rate is low: a five-year survival rate of ca. 10% compared with patients with acute lymphocytic leukemia (72.1%) [25].

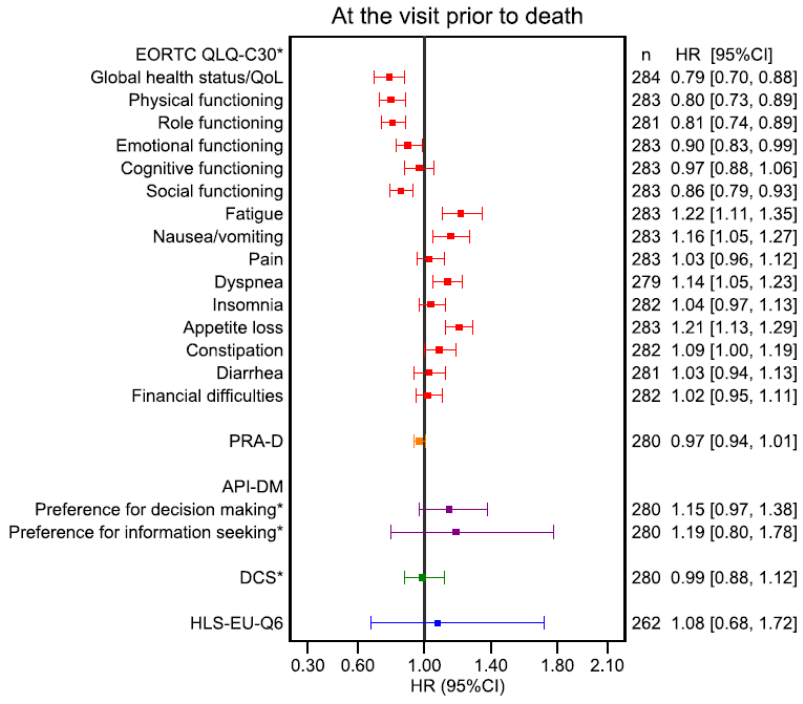
Patient-reported outcomes associated with drop-out

Our results show that lower levels of PROs are in fact associated with both a higher probability of drop-out and premature death. Cancer patients with a poor quality of life and high symptom burden at baseline and at the visit before drop-out were at high risk of early drop-out and death in our study. These findings are similar to those of other studies [6, 7, 21, 22]. In addition to this, we found that participants who dropped out due to illness or other reasons had a lower disease burden and better functionalities than those who died during the study period. We found that fatigue, nausea, vomiting, and appetite loss were associated with early death for both time points at baseline and at the visit before drop-out. These symptoms have been identified as early signs of upcoming death in cancer patients, especially fatigue symptoms [26]. In other words, our results show that patients who died prematurely tend to show a trend of progressive deterioration, indicated by a reduction in their health-related quality of life scores and lower baseline scores, while the patients who completed the study were stable throughout.

Finding application for further statistical analyses

Our findings, that baseline characteristics and QoL were associated with drop-out, may be useful for determining the potential missing data mechanism, which is a prerequisite to choosing how to handle missing PRO data, such as imputation methods. In addition to this, poor baseline QoL scores should encourage researchers to assess more auxiliary data, such as the Eastern Cooperative Oncology Group (ECOG) performance status. It should also encourage them to research the reasons for incomplete PROMs questionnaires being collected, assisting in determining what is missing and using this auxiliary data as a covariate in the model, or in multiple imputation methods [27]. In the relationship between symptom severity and missing data in health-related QoL due to drop-out or death, the missing data mechanism is clearly not completely random (MCAR). The pattern of PROs before dropping out and death may suggest that it is missing at random (MAR) or not missing at random (MNAR). The results suggest that, within the OSCAR study and similar studies, gaps in information have to be handled appropriately for a primary endpoint analysis to





*The values change by 10 points

Fig. 3 Patient-reported outcomes at baseline and at the visit prior to death, and their association to time-to-death. *n* = the observation number, HR = Hazard ratio, CI=Confidence interval, High values of global health status/QoL and functional sub-scales in the EORTC QLQ-C30 indicate a better outcome, High values of symptom sub-scales (e.g fatigue, nausea/vomiting, pain etc.) and DCS indicate a poorer outcome

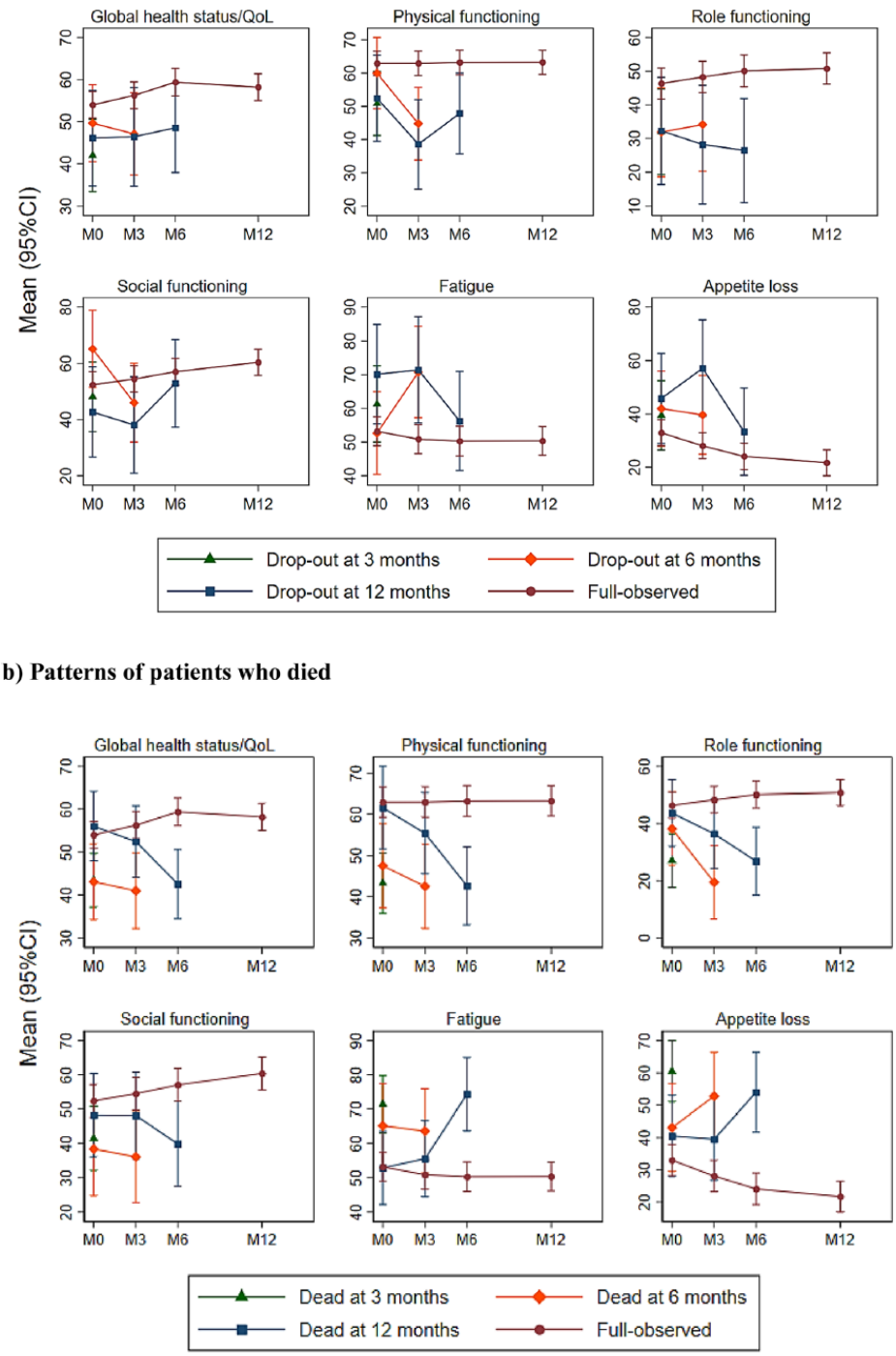


Fig. 4 Patterns of patients who (**a**) dropped out or (**b**) died: Sub-scales of the EORTC QLQ-C30 versus follow-up times, stratified by time of drop-out or death. The possible range of the EORTC QLQ-C30 is 0–100, with higher values indicating a better QoL in global health status/QoL, physical functioning, role functioning, social functioning, and lower values indicating better symptoms for fatigue and appetite loss

yield unbiased results. Complete case analysis or last observation carried forward (LOCF) methods would not be appropriate, especially in health-related QoL outcomes. However, yielding unbiased results depends on the appropriate handling of missing data within the analytical approach [28]. Alternative models (like pattern mixture models and joint models) might be used if the PROs data is MNAR [27, 29]. Independent of the particular statistical method used to handle missing PRO data, sensitivity analyses should be conducted and reported, regardless of the type of missing data mechanism [27].

Implications for further studies with advanced cancer patients

Our results show that a high drop-out rate in advanced cancer patients is to be expected, especially when due to premature death. Rates of early death are associated with certain diagnosis groups, such as malignant neoplasm of the pancreas. Therefore, the adjustment of sample sizes should have an attrition rate of up to 50% when planning a study. Inclusion criteria should be more specific (such as estimated life expectancy or ECOG performance status). An estimated life expectancy of more than 3 months is commonly used in cancer clinical trials [30]; however, choosing the estimated life expectancy depends on the subject matter of the study and their benefit from receiving an intervention. However, recruitment when patients are admitting to the intensive care unit should be avoided to reduce drop-out due to in-hospital death. Otherwise, the study assistant should do a post-recruitment if the participant dies shortly after the initial recruitment. The researcher should actively monitor baseline characteristics, as the likelihood of dropping out depends on a patient's social background (marriage, social support), gender, and educational status. Fatigue, nausea, vomiting, and appetite loss are signs of imminent death. Prompt changes in the follow-up phase study process might be necessary, such as close, continued monitoring after a patient misses an assessment and shorter time windows between follow-up visits and post-recruitment. Reasons for mis-measurement and discontinuation are very importance, especially when PROs are the primary research outcome. These reasons may help statisticians to handle the missing data and choose an appropriate statistical approach to the type of missing mechanism.

Limitations

No information on the severity of a patient's disease (such as the stage of cancer) which could affect the time until death and multivariable analyses were not possible, and the statistical power was restricted by multicollinearity within the PROs and a low number of observations in some variables, such as education, family status, and HLS-EU-Q6. This exploratory study was done with no correction for multiple testing and without a multivariable analysis.

Conclusions

Our exploratory study shows that participants with a low quality of life, poor symptoms and a lack of social support are more likely to discontinue a study earlier than patients with better values, resulting in a higher rate of missing patient-reported outcomes. This should be taken into account when planning a study of advanced cancer patients by monitoring in those who have a low baseline quality of life. Expected high mortality rates of up to 50% during the study time should already be considered in sample size calculations. Follow-up periods and study durations should be planned well. Continued monitoring is useful for characterizing the study sample and reacting quickly, thus avoiding a high drop-out rate. Methods of identifying the factors related to the drop-out similar to those presented here are useful for determining the missing data mechanism and informing choice of statistical methods for primary endpoint analysis. Methods for handling missing data might be applied appropriately, interpreting results of patient-reported outcomes should be done within caution, and the reasons for dropping out and how this may impact the findings should be discussed.

Abbreviations

API-DM: The German modified version of the Autonomy Preference Index; CASMIN: The Comparative Analysis of Social Mobility in Industrial Nations; CI: Confidence Interval; DCS: Decision Conflict Scale; EORTC QLQ-C30: The European Organization for Research and Treatment of Cancer's Quality of Life Questionnaire; HL: Health Literacy; HLS-EU-Q6: The European Health Literacy Survey; HR: Hazard Ratio; IQR: Interquartile Range; LOCF: Last Observation Carried Forward; MAR: Missing at Random; MCAR: Missing Completely at Random; MNAR: Missing Not at Random; OSCAR: The Oncological Social Care Project; PRA-D: Patient Reaction Assessment; PROs: Patient-reported Outcomes; PROMs: Patient-reported Outcome Measures; OSSS-3: Oslo Social Support Scale; QoL: Quality of Life; SD: Standard Deviation; SHR: Sub-hazard Ratio

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12874-021-01259-0.

Additional file 1.

Acknowledgements

The authors would like to thank all the participants, study nurses, and social care nurses for the information and for collecting the data. We would also like to thank Dr. Annette Aigner for her useful comments and language checks.

Authors' contributions

PG drafted the manuscript and performed the statistical analysis. JF, DS and LS participated in the design of the study and critical review. UG and LS supervised the study. UG was involved in the statistical design and verified the analytical methods. All the authors discussed the results and approved the final manuscript.

Funding

The OSCAR project received a grant from the Innovation Fund of the Federal Joint Committee (G-BA), reference number: 01NVF17016. The G-BA had no

role in the design of the study and collection, analysis and interpretation of data and in writing the manuscript. Open Access funding enabled and organized by Projekt DEAL.

Availability of data and materials

The datasets analyzed during the current study are not publicly available due to privacy and ethical concerns, and neither the data nor the source of the data can be made available. Upon request, the analysis code is available from the author.

Declarations

Ethics approval and consent to participate

The OSCAR was approved by the ethics committees at Charité – Universitätsmedizin Berlin (EA2/192/17) and the Medical Association of North Rhine (2017429). The participants were enrolled in the study after providing written informed consent.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Berlin Institute of Health at Charité –Universitätsmedizin Berlin, Charitéplatz 1, 10117 Berlin, Germany. [2]Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Institute of Biometry and Clinical Epidemiology, Charitéplatz 1, 10117 Berlin, Germany. [3]Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Institute of Medical Sociology and Rehabilitation Science, Charitéplatz 1, 10117 Berlin, Germany.

Received: 20 December 2020 Accepted: 25 March 2021
Published online: 20 April 2021

References

1. Weldring T, Smith SM. Patient-reported outcomes (PROs) and patient-reported outcome measures (PROMs). Health Serv Insights. 2013;6:61–8. https://doi.org/10.4137/HSI.S11093.
2. Osoba D. Health-related quality of life and cancer clinical trials. Ther Adv Med Oncol. 2011;3(2):57–71. https://doi.org/10.1177/1758834010395342.
3. Basch E, Deal AM, Dueck AC, Scher HI, Kris MG, Hudis C, et al. Overall survival results of a trial assessing patient-reported outcomes for symptom monitoring during routine Cancer treatment. Jama. 2017;318(2):197–8. https://doi.org/10.1001/jama.2017.7156.
4. Basch E, Deal AM, Kris MG, Scher HI, Hudis CA, Sabbatini P, et al. Symptom monitoring with patient-reported outcomes during routine cancer treatment: a randomized controlled trial. J Clin Oncol. 2016;34(6):557–65. https://doi.org/10.1200/JCO.2015.63.0830.
5. Bernhard J, Cella DF, Coates AS, Fallowfield L, Ganz PA, Moinpour CM, et al. Missing quality of life data in cancer clinical trials: serious problems and challenges. Stat Med. 1998;17(5–7):517–32. https://doi.org/10.1002/(SICI)1097-0258(19980315/15)17:5/7<517::AID-SIM799>3.0.CO;2-S.
6. Hui D, Glitza I, Chisholm G, Yennu S, Bruera E. Attrition rates, reasons, and predictive factors in supportive care and palliative oncology clinical trials. Cancer. 2013;119(5):1098–105. https://doi.org/10.1002/cncr.27854.
7. Perez-Cruz PE, Shamieh O, Paiva CE, Kwon JH, Muckaden MA, Bruera E, et al. Factors associated with attrition in a multicenter longitudinal observational study of patients with advanced cancer. J Pain Symptom Manag. 2018;55(3):938–45. https://doi.org/10.1016/j.jpainsymman.2017.11.009.
8. 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(10):899–907. https://doi.org/10.1177/0269216313486952.
9. Roick J, Danker H, Kersting A, et al. Factors associated with non-participation and dropout among cancer patients in a cluster-randomised controlled trial. Eur J Cancer Care. 2018;27:e12645. https://doi.org/10.1111/ecc.12645.
10. Spiers S, Oral E, Fontham ETH, Peters ES, Mohler JL, Bensen JT, et al. Modelling attrition and nonparticipation in a longitudinal study of prostate cancer. BMC Med Res Methodol. 2018;18(1):60. https://doi.org/10.1186/s12874-018-0518-6.
11. Basch E, Abernethy AP, Mullins CD, Reeve BB, Smith ML, Coons SJ, et al. Recommendations for incorporating patient-reported outcomes into clinical comparative effectiveness research in adult oncology. J Clin Oncol. 2012;30(34):4249–55. https://doi.org/10.1200/JCO.2012.42.5967.
12. Frick J, Schindel D, Gebert P, Grittner U, Schenk L. Improving quality of life in cancer patients through higher participation and health literacy: study protocol for evaluating the oncological social care project (OSCAR). BMC Health Serv Res. 2019;19(1):754. https://doi.org/10.1186/s12913-019-4585-0.
13. Lampert T, Hoebel J, Kuntz B, Müters S, Kroll LE. Messung des sozioökonomischen Status und des subjektiven sozialen Status in KiGGS Welle 2. In: Robert Koch-Institut, Epidemiologie und Gesundheitsberichterstattung, vol. 3; 2018.
14. Kocalevent RD, Berg L, Beutel ME, Hinz A, Zenger M, Harter M, et al. Social support in the general population: standardization of the Oslo social support scale (OSSS-3). BMC Psychol. 2018;6(1):31. https://doi.org/10.1186/s40359-018-0249-9.
15. Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, et al. The European organization for research and treatment of cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst. 1993;85(5):365–76. https://doi.org/10.1093/jnci/85.5.365.
16. Brenk-Franz K, Hunold G, Galassi JP, Tiesler F, Herrmann W, Freund T, et al. Quality of the physician-patient relationship - evaluation of the German version of the patient reactions assessment (PRA-D). Z Allg Med. 2016;92:103–8.
17. Transforming different Likert scales to a common scale https://www.ibm.com/support/pages/transforming-different-likert-scales-common-scale
18. Härter M, Simon D, Kriston L. Die deutsche modifizierte Fassung des Autonomie-Präferenz-Index (API-Dm). Klin Diagn Eval. 2011;4(1E):5–14.
19. Buchholz A, Hölzel L, Kriston L, Simon D, Härter M. Die decisional conflict scale in deutscher Sprache (DCS-D)-Dimensionale Struktur in einer Stichprobe von Hausarztpatienten. Klinische Diagnostik und Evaluation. 2011;4:15–30.
20. Lorini C, Santomauro F, Grazzini M, Mantwill S, Vettori V, Lastrucci V, et al. Health literacy in Italy: a cross-sectional study protocol to assess the health literacy level in a population-based sample, and to validate health literacy measures in the Italian language. BMJ Open. 2017;7(11):e017812. https://doi.org/10.1136/bmjopen-2017-017812.
21. Renovanz M, Hechtner M, Kohlmann K, Janko M, Nadji-Ohl M, Singer S, et al. Compliance with patient-reported outcome assessment in glioma patients: predictors for drop out. Neurooncol Pract. 2018;5(2):129–38. https://doi.org/10.1093/nop/npx026.
22. Ramsey I, de Rooij BH, Mols F, Corsini N, Horevoorts NJE, Eckert M, et al. Cancer survivors who fully participate in the PROFILES registry have better health-related quality of life than those who drop out. J Cancer Surviv. 2019;13(6):829–39. https://doi.org/10.1007/s11764-019-00793-7.
23. Marcus MW, Raji OY, Chen Y, Duffy SW, Field JK. Factors associated with dropout in a lung cancer highrisk cohort--the Liverpool lung project. Int J Oncol. 2014;44(6):2146–52. https://doi.org/10.3892/ijo.2014.2371.
24. Hosseini L, Dehghani M, Tahmasebi S, Akrami M, Shokrpour N, Ravangard R. Social support and breast cancer recurrence: is there any association? Open Public Health J. 2020;13(1):30–5. https://doi.org/10.2174/1874944502013010030.
25. SEER Cancer Statistics Review, 1975–2017,based on 2019 SEER data submission, posted to the SEER web site [https://seer.cancer.gov/csr/1975_2017/]. Accessed 7 Feb 2021.
26. Hui D, dos Santos R, Chisholm G, Bansal S, Silva TB, Kilgore K, et al. Clinical signs of impending death in cancer patients. Oncologist. 2014;19(6):681–7. https://doi.org/10.1634/theoncologist.2013-0457.
27. Bell ML, Fairclough DL. Practical and statistical issues in missing data for longitudinal patient-reported outcomes. Stat Methods Med Res. 2014;23(5):440–59. https://doi.org/10.1177/0962280213476378.
28. Bell ML, Kenward MG, Fairclough DL, Horton NJ. Differential dropout and bias in randomised controlled trials: when it matters and when it may not. BMJ. 2013;346(jan21 1):e8668. https://doi.org/10.1136/bmj.e8668.
29. Fielding S, Fayers PM, Ramsay CR. Investigating the missing data mechanism in quality of life outcomes: a comparison of approaches. Health Qual Life Outcomes. 2009;7(1):57. https://doi.org/10.1186/1477-7525-7-57.
30. Jin S, Pazdur R, Sridhara R. Re-evaluating eligibility criteria for oncology clinical trials: analysis of investigational new drug applications in 2015. J Clin Oncol. 2017;35(33):3745–52. https://doi.org/10.1200/JCO.2017.73.4186.

Publisher's Note

Curriculum Vitae

My curriculum vitae does not appear in the electronic version of my paper for reasons of data protection

Publication list

1. Hinkson L, Schauer M, Latartara E, Alonso-Espias M, Rossetti E, Gebert P, Hinkson S, Henrich W: The Charite external cephalic version for leading twin breech without regional anesthesia and tocolysis. A prospective study on feasibility, sonographic assessment and outcomes. Eur J Obstet Gynecol Reprod Biol 2022, 268:62-67. (IF=2.435)

2. Pohl J, Gebauer B, Gebert P, Tacke F, Wree A: Predicting survival after TIPS: Child Pugh score is not inferior to MELD and FIPS score - back to basics? J Hepatol 2021, 75(6):1505-1506. (IF=25.083)

3. Krobisch V, Gebert P, Gul K, Schenk L: Women bear a burden: gender differences in health of older migrants from Turkey. Eur J Ageing 2021, 18(4):467-478. (IF=0.982)

4. Karsten MM, Kuhn F, Pross T, Blohmer JU, Hage AM, Fischer F, Rose M, Grittner U, Gebert P, Ferencz J et al: PRO B: evaluating the effect of an alarm-based patient-reported outcome monitoring compared with usual care in metastatic breast cancer patients-study protocol for a randomised controlled trial. Trials 2021, 22(1):666. (IF=2.486)

5. Ghadjar P, Stritter W, von Mackensen I, Mehrhof F, Foucre C, Ehrhardt VH, Beck M, Gebert P, Kalinauskaite G, Luchte JS et al: External application of liver compresses to reduce fatigue in patients with metastatic cancer undergoing radiation therapy, a randomized clinical trial. Radiat Oncol 2021, 16(1):76. (IF=4.37)

6. Gebert P, Schindel D, Frick J, Schenk L, Grittner U: Characteristics and patient-reported outcomes associated with dropout in severely affected oncological patients: an exploratory study. BMC Med Res Methodol 2021, 21(1):77. (IF=4.402)

7. Chen Z, Wittenberg S, Auer TA, Bashkuev M, Gebert P, Fehrenbach U, Geisel D, Graef F, Maerdian S, Tsitsilonis S: The Effect of Fat Distribution on the Inflammatory Response of Multiple Trauma Patients-A Retrospective Study. Life (Basel) 2021, 11(11). (IF=3.817)

8. Keller S, Yamamura J, Sedlacik J, Wang ZJ, Gebert P, Starekova J, Tahir E: Diffusion tensor imaging combined with T2 mapping to quantify changes in the skeletal muscle associated with training and endurance exercise in competitive triathletes. Eur Radiol 2020, 30(5):2830-2842. (IF=4.101)

9. Keller S, Borde T, Brangsch J, Adams LC, Kader A, Reimann C, Gebert P, Hamm B, Makowski M: Native T1 Mapping Magnetic Resonance Imaging as a Quantitative Biomarker for Characterization of the Extracellular Matrix in a Rabbit Hepatic Cancer Model. Biomedicines 2020, 8(10). (IF=4.717)

10. Frick J, Schindel D, Gebert P, Grittner U, Schenk L: Improving quality of life in cancer patients through higher participation and health literacy: study protocol for evaluating the oncological social care project (OSCAR). BMC Health Serv Res 2019, 19(1):754. (IF=2.042)

11. Faizy TD, Broocks G, Thaler C, Rauch G, Gebert P, Sturner KH, Flottmann F, Leischner H, Kniep HC, Stellmann JP et al: Development of Cortical Lesion Volumes on Double Inversion Recovery MRI in Patients With Relapse-Onset Multiple Sclerosis. Front Neurol 2019, 10:133. (IF=2.889)

Acknowledgments

First, I would like to express my gratitude to my supervisor, ***PD Dr. Phil Ulrike Grittner***, for all her guidance, support, and constructive feedback on this dissertation. She not only supported me in my studies but also in my work. She is an excellent supervisor in my working group and my favorite colleague, inspiring me in both thought and work. I would especially like to thank ***Professor Dr. Geraldine Rauch*** for her support of my dissertation and work.

This dissertation would not have been successfully completed had I not used the data from the Oncological Social Care Project, for which I am very grateful. I would like to thank ***Professor Dr. Liana Schenk*** for giving me a chance to work on this project and allowing me to use the data, as well as ***Dr.rer.medic. Daniel Schindel*** and ***Johann Frick*** for their help and support. I would also like to thank the study nurses, the social care nurses, the participants, and the study teams for running the project.

I would also like to thank all of my colleagues in the Institute of Biometry and Clinical Epidemiology for their support. As well, a special thanks to ***Dr.rer.medic.Alice Schneider*** for her proof reading.

Last but not least, I would like to thank my family and friends here in Germany and in Thailand for their continued support throughout my academic career.