# Improving the Quality of Patient Reported Outcomes by Minimizing Missing Data

By

Michael John Palmer

A thesis submitted to the Graduate Program in Public Health Sciences

in conformity with the requirements for the

Degree of Doctor of Philosophy

Queen's University

Kingston, Ontario, Canada

Final submission December, 2019

## Abstract

Patient-Reported Outcomes (PROs) are often included as an endpoint in prospective studies. Clinicians use the study results to help individuals with cancer make informed decisions about treatment. Missing PRO data threaten the validity of randomized clinical trials (RCTs) and can undermine scientific credibility of PRO-specific conclusions. The thesis goal was to use quantitative and qualitative approaches to identify causes of missing PRO data.

A systematic review of the literature identified 46 unique categories of factors reported to have an association with missing PRO data. These factors were organized into a framework having five main components: Instrument, Participant, Centre, Staff, and Study.

Cox regression examined the association of 15 factors with time to the first missing instrument in data from three RCTs: MA17 (breast cancer), PR7 (prostate cancer), and LY12 (non-Hodgkin's lymphoma). Most factors did not have significant independent association with outcome, but a few did, depending on the RCT. Centre presence of a post-graduate residency training program (MA17, PR7) and geographic location (PR7, LY12) had a stronger association. Participant age (PR7) and baseline level of well-being (LY12), and centre level of activity (LY12) had a weaker association.

Interviews were conducted with 11 Clinical Research Associates (CRAs). Ten themes were identified in analysis based on descriptive phenomenology. The CRAs described their experience with different patterns of missing PRO data. PROs for routine clinical care compete with PROs for RCTs. Several important influences of missing PRO data were inherent to the instruments

(quantity, repetition, wording of questions, and format), patients (characteristics, attributes of the cancer, and toxicity from the treatment), centres (deficiencies in research infrastructure), staff (developing relationships with patients), and RCTs (data capture in electronic versus paper format). Adapting the assessment schedule and providing sufficient time for the patient to complete the instrument were challenging for the CRAs, particularly at disease progression. There is a need for field guidance, and a motivation and knowledge base among CRAs to develop such guidance.

This dissertation provides evidence to facilitate understanding of which factors influence missing PRO data, so that steps can be taken to improve the quality of PROs by minimizing missing data.

## Co-Authorship and Financial Support

The main manuscripts were written by Michael J. Palmer in collaboration with his supervisors, Harriet Richardson and Michael Brundage, and additional co-authors. Financial support was provided by The Terry Fox Foundation Training Program in Transdisciplinary Cancer Research in partnership with CIHR (now identified as Queen's University Terry Fox Research Institute Transdisciplinary Training Program in Cancer Research), Queen's Graduate Award and Health Sciences Graduate Growth Funding. Partial support was provided by the Division of Cancer Care and Epidemiology in the Cancer Research Institute at Queen's University.

**Chapter 3 – Manuscript 1:** *A systematic review and development of a classification framework for factors associated with missing patient-reported outcome data.* This manuscript is presented in the format published in Clinical Trials in 2018. (Full citation: Palmer MJ, Mercieca-Bebber R, King M, Calvert M, Richardson R, Brundage M. A systematic review and development of a classification framework for factors associated with missing patient-reported outcome data. *Clinical Trials* 2018;15:95-106.) The systematic review was conducted by Michael J. Palmer in collaboration with Rebecca Mercieca-Bebber and supervision from Michael Brundage, Madeleine King and Melanie Calvert. The classification framework was developed by Michael J. Palmer with supervision from Harriet Richardson and Michael Brundage. The manuscript was written by Michael J. Palmer with supervision from Harriet Richardson and Michael Brundage and with additional input from Rebecca Mercieca-Bebber, Madeleine King and Melanie Calvert.

**Chapter 4 – Manuscript 2:** *Factors predicting missing instruments in three cancer randomized clinical trials*. This manuscript is presented in the format submitted to *Quality of Life Research* in 2019. Selection of variables, data analysis, interpretation of the results and writing the manuscript were done by Michael J Palmer with supervision from Harriet Richardson and Michael Brundage and with additional input from Dongsheng Tu.

**Chapter 5 – Manuscript 3:** *What is the lived experience of clinical research associates with missing patient-reported outcome data?* This manuscript is formatted for planned submission to *Quality of Life Research* in 2019. Data analysis, interpretation of the results and writing the manuscript were done by Michael J Palmer with supervision from Harriet Richardson and Michael Brundage and with additional input from Terry Krupa.

Michael J. Palmer contributed to two additional published manuscripts as co-author.

**Appendix F – Manuscript 4:** *Design, implementation and reporting strategies to reduce the instance and impact of missing patient-reported outcome (PRO) data: a systematic review.* The abstract of this manuscript is presented in the format published in *BMJ Open* in 2016. (Full citation: Mercieca-Bebber R, Palmer MJ, Brundage M, et al. 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**:e010938.) Michael J. Palmer contributed to study concept and design, data acquisition and analysis, edited and approved final manuscript.

**Appendix G – Manuscript 5:** *Guidelines for inclusion of patient-reported outcomes in clinical trial protocols: the SPIRIT-PRO extension.* The abstract of this manuscript is presented in the format published in *JAMA* in 2018. (Full citation: Calvert M, Kyte D, Mercieca-Bebber R, et al. Guidelines for inclusion of patient-reported outcomes in clinical trial protocols: the SPIRIT-PRO extension. *JAMA* 2018;319:483-494.) Michael J. Palmer, as a member of ISOQOL Best Practices for PROs in Randomized Clinical Trials Protocol Checklist Taskforce and a Delphi expert panelist, contributed to study concept and design, data acquisition and analysis, critical revision of the manuscript for important intellectual content.

Michael J. Palmer contributed to two additional published abstracts as co-author.

**Appendix B – Abstract 1**: *What Evidence Suggests a Relationship Between Patient Age and Missing PRO Data on Randomized Clinical Trials (RCTs)?* The poster is provided in the format presented at International Society for Quality of Life Research in 2017. (Full citation: Palmer MJ, Mercieca-Bebber R, King M, Calvert M, Richardson R, Brundage M. *Qual Life Res* 2017;26(Supplement 1):138-139.) The abstract was written by Michael J. Palmer with supervision from Harriet Richardson and Michael Brundage and with additional input from Rebecca Mercieca-Bebber, Madeleine King and Melanie Calvert. Michael J. Palmer prepared and presented the poster with additional input from Michael Brundage.

**Appendix H – Abstract 2:** *Missing patient-reported outcome data: understanding associated factors and strategies to reduce its instance and impact through careful design, quality assurance and reporting.* The abstract is provided in the format published in Quality of Life Research in 2018. (Full citation: Mercieca-Bebber R, Palmer MJ, Brundage M, Calvert M, Stockler MR, King MT. *Qual Life Res* 2018;27(Suppl 1):S4.) Michael J. Palmer contributed to writing the abstract, synthesizing the information, preparing the slides, and made the oral presentation with additional input from Michael Brundage.

# Acknowledgements

This dissertation reflects the positive contribution of many people who, over the years, helped me to grow as a person and to improve as a researcher. I extend a sincere 'thank you' to the following individuals for participating in my journey.

My wife, Basia, for her ongoing love and support. Your inquisitive mind helps me realize although not every question has an answer it is often worth pursuing. May the steps we have taken together begin a better future.

My 'brother', Adam Lodzinski, for his inspiring professionalism and on-going mentorship. Your leadership skills demonstrate ability with a group interview and help me understand the benefit of qualitative inquiry.

My thesis supervisors, Michael Brundage and Harriet Richardson, for their wisdom, guidance and patience. I continue to learn from you.

My thesis committee members, Dongsheng Tu and Terry Krupa, for their expertise with – and enthusiasm for – quantitative and qualitative methodology.

My advisory committee members, Melanie Calvert, Madeleine King, and Joe Pater, for their insightful comments and constructive advice.

Faculty within the Department of Public Health Sciences, especially Will Pickett, Will King, Patti Groome, and Paul Peng, for sharing their knowledge about epidemiology and biostatistics, and for challenging my thinking, during excellent courses and discussions outside of class.

My research would not be possible without the contribution of the Canadian Cancer Trials Group. I recognize several individuals, including Janice Flegg for assistance with identifying Clinical Research Associates, Liting Zhu for assistance with providing data for quantitative analyses, and 18 exceptional CRAs for participating in group or individual interviews. I thank 7 CRAs for demonstrating in a focus group that CRAs have abundant but untapped knowledge pertaining to missing PRO data. I also thank 11 CRAs for describing their lived experience with missing PRO data and offering their insight on important influences.

Fellow PhD candidates, Graham Smith and Liam Rémillard, for their enthusiasm, discussion of research, and sharing of ideas, especially while enjoying a beverage or two.

My parents, Jim and Dianne Palmer, for their ongoing belief in my abilities.

Terry Fox was an inspiration for my research. His words (“It took cancer to realize that being self-centred is not the way to live. The answer is to try and help others”) continue to motivate.

Several sources provided funding for this thesis. I am grateful to The Terry Fox Foundation Training Program in Transdisciplinary Cancer Research in partnership with CIHR (now identified as Queen’s University Terry Fox Research Institute Transdisciplinary Training Program in Cancer Research). Thank you to Lois Mulligan, for encouragement of trainees and for enthusiasm of cancer research, and to Mark Andrews for excellence in administrative support. I appreciate Queen’s Graduate Award and Health Sciences Graduate Growth Funding, and partial support provided by the Division of Cancer Care and Epidemiology in the Cancer Research Institute at Queen’s University. CCE was my academic home for a few years during which I had good fortune to meet and develop friendship with many exceptional people like Jennifer O’Donnell and Tina Dyer.

## Table of Contents

## List of Tables

**List of Figures**

## List of Abbreviations

| | |
|---|---|
| CCTG | Canadian Cancer Trials Group |
| 95% CI | 95% Confidence Interval |
| CINAHL | Cumulative Index to Nursing and Allied Health Literature |
| CONSORT | Consolidated Standards of Reporting Trials |
| CRA | Clinical Research Associate |
| FACT-G | Functional Assessment of Cancer Therapy – General |
| HR | Hazard Ratio |
| ISOQOL | International Society for Quality of Life Research |
| LY12 | The 12$^{th}$ clinical trial in the Hematology Disease Site involving non-Hodgkin's Lymphoma, as per nomenclature of the Canadian Cancer Trials Group |
| MA17 | The 17$^{th}$ clinical trial in the Breast Disease Site involving Mammary cancer, as per nomenclature of the Canadian Cancer Trials Group |
| MAR | Missing at random |
| MCAR | Missing completely at random |
| MNAR | Missing not at random |
| MEDLINE | Medical Literature Analysis and Retrieval System Online |
| MENQOL | Menopause-Specific Quality of Life Questionnaire |
| PR7 | The 7$^{th}$ clinical trial in the Genito-Urinary Disease Site involving Prostate cancer, as per nomenclature of the Canadian Cancer Trials Group |
| PRO | Patient-Reported Outcome |
| QLQ-C30 | Quality of Life Questionnaire – Core 30 |
| QOL | Health-related quality of life |
| RCT | Randomized Clinical Trial |

| | |
|---|---|
| SF-36 | Short-Form 36 |
| SISAQOL | Setting International Standards in Analyzing Patient-Reported Outcomes and Quality of Life Endpoints Data for Cancer Clinical Trials |
| SPIRIT | Standard Protocol Items Recommendations for Interventional Trials |

# Chapter 1

## General Introduction

### 1.1 Goal of the Thesis

Missing data on patient-reported outcomes (PROs) are a serious threat to the internal and external validity of a randomized clinical trial (RCT) and can undermine the scientific credibility of causal conclusions relating to PRO findings [1]. Despite different approaches to reducing or accounting for missing PRO data, the problem persists [2]. The goal of the thesis was to use quantitative and qualitative approaches to identify the causes of missing PRO data. The dissertation is manuscript-based with four parts. Part 1 identifies factors reported to be associated with missing PRO data. Part 2 examines some of the factors in secondary analyses of quantitative data. Part 3 identifies influences of missing PRO data from primary collection and analysis of qualitative data. Part 4 consolidates results from Part 3 with results from Part 2.

### 1.2 Developing Structure for the Objectives

Research to identify the causes of missing PRO data is novel and thus some groundwork was required to develop the objectives. One way of proceeding, as recommended by the reviewers of the thesis outline, was to do preliminary qualitative work with relevant stakeholders to inform the research design. The selected stakeholder group was the Canadian Cancer Trials Group (CCTG, a national program of the Canadian Cancer Society and a multi-centre cooperative oncology group). CCTG conducts clinical trials in cancer therapy, supportive care, and prevention, and has a history of collecting PRO data in phase III RCTs [3]. A focus group of research staff experienced in the collection of PROs was organized with the assistance of the

CCTG Central Operations and Statistics Office, and representative cancer patients were individually interviewed, to elicit their level of involvement in this research.

The first step was to seek guidance from the perspective of Clinical Research Associates (CRAs) who work at a cancer centre or hospital. Many CRAs have professional credentials, for example Registered Nurse or Certified Clinical Research Professional. The CRA administers an instrument (term used to describe the measuring tool which includes a collection of self-report items) to the patient on an RCT and when returned may review the instrument and ask the patient to answer missed items. Section 1.2.1 describes the focus group, organized as a window into the CRAs experience with missing PRO data to determine if CRAs are a willing source of relevant information.

The second step was to seek guidance from the perspective of cancer patients. Section 1.2.2 describes two meetings, organized to determine through informal discussions, enthusiasm for the topic and feasibility for the inclusion of cancer patients who have experience with missing PRO data on RCTs.

### 1.2.1 Focus Group with CCTG CRAs

The PhD candidate (MJP) organized the focus group [4] and attended but did not moderate, due to concern for a potential Hawthorne effect [5]. An experienced moderator with academic credentials led the focus group, held in a meeting room at a neutral location. The consent form and interview guide received Queen's ethics approval. Seven CRAs were selected from the CCTG CRA Executive Committee and from representatives on CCTG Disease Site and Standing Committees. CRAs revealed an enthusiasm for the topic, identified several factors that influence missing PRO data, and contributed their knowledge of the clinic environment. It became clear that CRAs are an untapped source of information that is relevant to identifying the causes of missing PRO data.

### 1.2.2 Interviews with the Patient Representative to the CCTG Quality of Life Committee and the Chair of the CCTG Patient Representatives Committee

The Patient Representative revealed an enthusiasm for the topic, encouraged the inclusion of cancer patients, and offered a willingness to help. The Chair revealed that although the Patient Representatives Committee has 20 members, not every member is a cancer survivor [6], many members are caregivers for a spouse or a sibling with cancer, and only two members have experience as participants on an RCT. Original data collected from cancer patients who participated on an RCT could provide a perspective on missing PRO data. However, due to several logistic challenges, the thesis committee recommended not to include interviews with cancer patients in the thesis.

## 1.3 Overview of Sources of Data

CCTG served as the source of data for quantitative analyses of missing PRO data. Available data are computerized and include information on the instruments (e.g. number of questions), participants (e.g. age), centres (e.g. location), and RCTs (e.g. type of treatment). Information on staff at centres is restricted to roster only. CCTG also served as the source for collection and analysis of qualitative data with a purposeful sample taken from the membership of 950 CRAs.

## 1.4 Objectives

1.1 To perform a systematic review of the literature.

1.2 To create a Classification Framework for factors reported to be associated with missing PRO data.

2.1 To determine which characteristics of the instruments, participants, centres, and trials predict the first missing instrument after randomization.

2.2 To determine which characteristics of the instruments, participants, centres, and trials predict missing items that preclude scoring of a scale.

3 To describe the lived experience of Clinical Research Associates with missing PRO data.

4.1 To consolidate the qualitative and quantitative results.

4.2 To revise the Classification Framework.

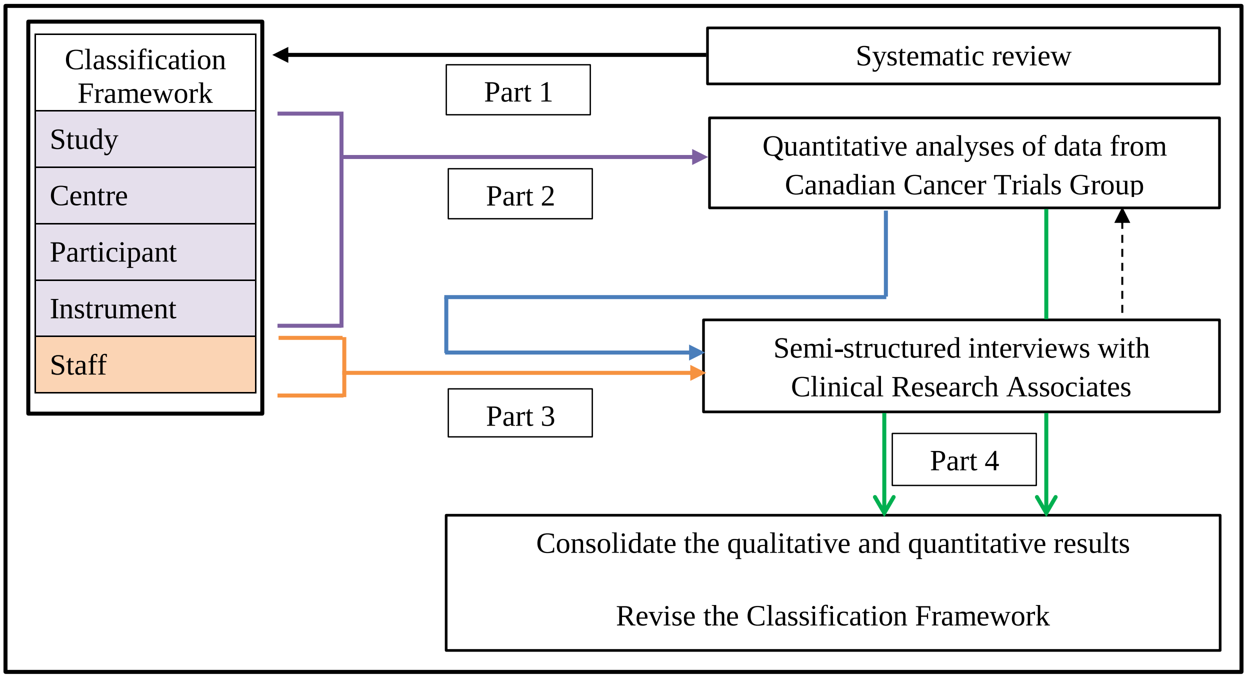
## 1.5 Study Design

Mixed-methods research involves using quantitative and qualitative approaches for the collection and analysis of data, integration of findings, and interpretation of results within a single study [7]. The sequential explanatory design begins with a quantitative component and follows with a qualitative component [8]. Despite differences between the paradigms, parallels between quantitative and phenomenological methods provide justification for combining these [9]. A sequential explanatory mixed-methods design was used to identify the important influences of missing PRO data. The combination of quantitative analyses with qualitative inquiry helped to provide a more comprehensive understanding of which factors influence missing PRO data than either in isolation.

## 1.6 Overview of the Dissertation

This dissertation consists of six chapters and eight appendices. Chapter 2 provides background information on PROs and RCTs, defines missing PRO data, and describes the types, mechanisms, and patterns. Then, Chapter 2 outlines the consequences, and summarizes the approaches to preventing and dealing with missing PRO data.

**Figure 1-1. Overview of the Thesis.**




In Part 1, a systematic review of the literature led to the creation of a Classification Framework for factors reported to be associated with missing PRO data. The manuscript was published in *Clinical Trials* in 2018 and is presented in Chapter 3. Part 2 and Part 3 focus on cancer. In Part 2, the Classification Framework informed quantitative analyses of individual participant data from three phase III cancer RCTs. The manuscript was submitted to *Quality of Life Research* in 2019 and is presented in Chapter 4. In Part 3, in-depth semi-structured interviews were undertaken with individual CRAs working at cancer centres. The manuscript was formatted for submission to *Quality of Life Research* and is presented in Chapter 5. In Part 4, the qualitative results were consolidated with the quantitative results, described in Chapter 6 with the General Discussion.

The following paragraph provides an overview of the Appendices:

Appendix A provides the research ethics board approvals of the protocol, consent forms, and interview guides, for the focus group with CRAs and for the individual interviews with CRAs. Appendix B provides the poster for the abstract, co-authored by MJP, published in *Quality of Life Research* and presented at International Society for Quality of Life Research (ISOQOL) in 2017, on preliminary work regarding evidence for an association between patient age and missing PRO data that provides justification for proceeding to Part 2. Appendix C provides the results of additional quantitative analyses performed in Part 2 after pooling data, formatted as a commentary for submission to *Translational Behavioral Medicine*. Appendix D provides the results of additional quantitative analysis performed in Part 2 involving missing items at baseline for one RCT, formatted as a research letter for submission to *Quality of Life Research*. Appendix E presents the e-supplements for the three main manuscripts presented in Chapters 3 – 5. Appendix F presents the abstract of the manuscript, co-authored by MJP and published in *BMJ Open* in 2016, for a systematic review identifying strategies for reducing missing PRO data. Appendix G presents the abstract of the manuscript, co-authored by MJP and published in *JAMA* in 2018, for an ISOQOL Protocol Checklist Taskforce and Delphi panel that developed guidelines for inclusion of PROs in protocols. Appendix H presents the abstract, co-authored by MJP, published in *Quality of Life Research*, and presented as an oral communication in symposium at ISOQOL in 2018, on preliminary work towards understanding the factors associated with missing PRO data and using strategies to reduce the negative consequences of missing PRO data.

1.7 References

[1] Little RL, D'Agostino R, Cohen ML et al. The prevention and treatment of missing data in clinical trials. *N Engl J Med* 2012; 367:1355-1361.

[2] O'Neill RT, Temple R. The prevention and treatment of missing data in clinical trials: an FDA perspective on the importance of dealing with it. *Clin Pharmacol Ther* 2012;91:550-554.

[3] Osoba D, Bezjak A, Brundage M, Pater J. Evaluating health-related quality of life in cancer clinical trials: the National Cancer Institute of Canada Clinical Trials Group experience. *Value Health* 2007;10 Suppl 2:S138-45.

[4] Krueger RA, Casey MA. *Focus groups 3rd edition. A practical guide for applied research.* SAGE Publications: Thousand Oaks, 2000.

[5] McCambridge J, Witton J, Elbourne DR. Systematic review of the Hawthorne effect: new concepts are needed to study research participation effects. *J Clin Epidemiol* 2014;67:267-277.

[6] Needham J, Stanton H, Nomikos D. Integrating patient / public involvement in the NCIC Clinical Trials Group. Society for Clinical Trials, Montreal, May 17, 2016.

[7] Creswell JW. Mixed methods research. *The SAGE encyclopedia of qualitative research methods*. SAGE Publications: Thousand Oaks, 2008.

[8] Creswell JW. *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research, Fourth Edition.* Boston: Pearson Education, 2012.

[9] Mayoh J, Onwuegbuzie AJ. Toward a conceptualization of mixed methods phenomenological research. *J Mix Methods Res* 2015;9:91-107.

# Chapter 2

# PROlogue

## 2.1 Patient-Reported Outcomes

PROs are a “measurement of any aspect of a patient’s health that comes directly from the patient without interpretation of the patient’s responses by a physician or anyone else” [1]. Health-related quality of life (QOL) is an example. PROs have three characteristics: data are collected directly from the patient, can change over time, and are often multi-dimensional [2]. PROs contrast to clinician-assessed indicators of health status, such as performance status [3]. Measurement of PROs is achieved in different ways. A visual analogue scale, for example, requires the respondent to place a tick on a ten-centimetre line between extreme response options [4, 5]. A Likert scale requires the respondent to circle a response from several categorical options [6]. A respondent may record PROs (for example, the presence of symptoms) in a diary each day for a week, or may record PROs in a questionnaire on one day concerning the past week.

Many questionnaires have been developed to measure PROs and these often include an overall question about global health status or QOL, as well as questions (typically referred to as ‘items’) about domains and dimensions of interest [7]. Generic questionnaires provide a general assessment of health status, regardless of the health condition. The SF-36, an example of a generic instrument [8], consists of 36 items organized in eight scales. Four scales measure physical health (physical functioning, role physical, bodily pain, general health). Four scales measure mental health (vitality, social functioning, role emotional, mental health). Every respondent is expected to answer all of the items in the same order. In contrast, disease-specific questionnaires provide a comprehensive assessment, with focus on key aspects of the health condition and its treatment [9]. The QLQ-C30, an example of a disease-specific questionnaire for cancer [10], consists of 30 items organized in 15 scales. Five scales measure function (physical, emotional, role, cognitive, social), nine scales measure symptoms (fatigue, nausea and vomiting, pain, dyspnea, insomnia, appetite loss, constipation, diarrhea, financial loss), and one scale measures global health status and QOL. Respondents are expected to answer every item in the same order. An adaptive version is being developed, with tailoring of the items [11]. A condition-specific module can supplement the QLQ-C30, for measurement depth relevant to the clinical setting [12]. The Functional Assessment of Cancer Therapy-General, another example of a disease-specific questionnaire for cancer, consists of 27 items organized in four scales. The four scales measure physical, social/family, emotional, and functional well-being [13].

### 2.2 Context of a Randomized Clinical Trial

The objectives of a prospective cohort study [14] or a phase III RCT [15 - 17] can include hypotheses evaluated by PROs. A phase III RCT that is well designed, properly conducted, appropriately analysed and transparently reported can contribute 1B level evidence for the efficacy of a new treatment compared with an alternative treatment [18].

Assessment of PROs is longitudinal in an RCT. Typically, the first assessment is before randomization, to measure health status at baseline before delivery of the assigned treatment. Subsequent assessments are after randomization at specific time points, to measure health status during treatment and follow-up.

PROs describe, from the patient's perspective, symptoms associated with disease, toxicity arising from treatment, the ability to function and to perform activities of daily living, and different dimensions of QOL. The level at specific time points and changes from baseline can be determined in the analysis and may demonstrate between-arm differences. The PRO results of an RCT can help individuals choose the best treatment and can enrich understanding of patient experiences [19].

### 2.3 Missing PRO Data

Missing PRO data arise when a patient fails to complete an assessment, fails to attend an evaluation, or is unavailable because of illness or death [1]. Clearly, data cannot be expected from patients who have left a study because of an event like death, so data are not regarded as missing afterward [20].

This thesis defines PRO data as missing when an assessment was expected but the patient did not complete all or part of the instrument. By this definition, if there was no expectation for a patient to complete an instrument (e.g., after progression or death, or because of non-adherence to treatment, drop-out, lost to follow-up) the PRO data are absent but are not considered to be missing.

#### 2.3.1 Types of Missing PRO Data

There are two different types of missing PRO data [21]. In one type, the entire instrument is missing, and in the second type, the instrument is partially completed and has missing items. This thesis identifies the two types as missing instruments, and missing items within non-missing instruments.

### 2.3.2 Mechanisms for Missing PRO Data

There are three different described mechanisms for missing PRO data [22, 23]. Missing completely at random (MCAR) is when the missing PRO data are independent of the observed PRO data. An example is when a staff member forgets to administer an instrument to the patient in the clinic. Missing at random (MAR) is when the missing PRO data are dependent on the observed PRO data. An example is a patient misses the first assessment after randomization following self-report of poor health status at baseline. Missing not at random (MNAR) is when the missing PRO data are dependent on the unobserved PRO data at the time of the scheduled assessment. An example is a patient with diminished QOL who does not attend the clinic because of severe symptoms. Section 2.3.4 explains the implications of these mechanisms.

### 2.3.3 Patterns of Missing PRO Data over Time

A convenient way to describe the frequency of missing instruments in an RCT is according to four patterns of missing PRO data. Participants may complete all assessments (no missing), all assessments up to the ones missed (monotone missing), some of the assessments in an irregular manner (intermittent missing), or none of the assessments (all missing). **Table 2-1**, adapted from Moinpour [24], provides the frequency and percent for each pattern of missing PRO data in an RCT of 5-fluorouracil for colorectal cancer. The shaded cells identify missing instruments. 56.9 % of 279 eligible patients completed a PRO instrument at all four scheduled assessments. Of the remainder, 32.9 % had a monotone, 9.1 % had an intermittent, and 1.1 % had an all-missing pattern of missing PRO data. A monotone pattern suggests a mechanism of MNAR (e.g., because of deteriorating health), while the intermittent pattern suggests mechanisms of MCAR and MAR.

| Baseline | 6 weeks | 11 weeks | 21 weeks | Frequency | Percent | Pattern of PRO Data |
|---|---|---|---|---|---|---|
| Yes | Yes | Yes | Yes | 159 | 56.9 | No missing |
| Yes | No | No | No | 33 | 11.8 | Monotone missing |
| Yes | Yes | No | No | 24 | 8.6 | |
| Yes | Yes | Yes | No | 35 | 12.5 | |
| Yes | No | Yes | No | 4 | 1.4 | Intermittent missing |
| Yes | No | No | Yes | 6 | 2.2 | |
| Yes | Yes | No | Yes | 8 | 2.9 | |
| Yes | No | Yes | Yes | 6 | 2.2 | |
| No | Yes | No | No | 1 | 0.4 | |
| No | No | No | No | 3 | 1.1 | All missing |

**Table 2-1. The frequency and percent for each pattern of missing PRO data in an RCT of 5-fluorouracil for colorectal cancer.** Adapted from Moinpour [24].

2.3.4 Consequences of Missing PRO Data

Missing PRO data are not without consequence. Missing PRO data are a waste of resources that if extreme reflects unethical research [25, 26]. As outlined, missing PRO data are a serious threat to the internal and external validity of RCTs, can undermine the scientific credibility of PRO-specific conclusions [27], and may impact the adoption of evidence into clinical practice.

The presence of missing items within a non-missing instrument, such as the assessment before randomization, may preclude calculation of a scale score. For some questionnaires like the QLQ-C30 [10], missing values are not imputed, so if a patient does not provide answers to more than one-half of the items that comprise the scale, the score is not calculated [28].

Missing PRO data may lead to a loss of power [29]. Power, the probability of rejecting a false null hypothesis [30], is dependent on sample size [31]. The sample size for an RCT is calculated for the primary objective, which is most often not related to PROs (e.g., disease-free survival or overall survival). Sample size is often not calculated for a secondary objective related to the comparison of PROs between arms [22]. The sample size to detect a clinically meaningful difference in PROs is usually smaller than that for the primary outcome because the magnitude of change required for PROs is usually larger than the magnitude of change required for the primary outcome. Whether the amount of missing PRO data, commensurate with a reduction in the effective sample size, results in a loss of power that is, by convention, lower than 80%, will depend on the individual RCT [32].

Estimates of outcome parameters with fewer data will lack precision from random error. The increased size of the standard errors will lead to wider confidence intervals [33], and may cause two errors of inference during hypothesis testing [34]: type I, when the investigator incorrectly concludes the data support a false alternative hypothesis, or type II, when the investigator incorrectly concludes the data do not support a true alternative hypothesis.

The main problem with missing PRO data is biased estimates of outcome parameters when the mechanism of missing PRO data is MNAR, which is common in cancer RCTs [35]. Randomization will balance known and unknown confounders [36] but will not protect against potential bias from missing PRO data [37]. The risk of bias depends on the mechanism [38], not on the magnitude of the response proportion [39]. With MCAR, there are no systematic differences in missing and observed values. With MAR, any systematic differences are explainable by differences in the observed values. With MNAR, systematic differences persist between the missing and observed values. Bias, then, may arise when the PRO data are MAR or MNAR. Of these, MNAR is considered to be the more serious problem [40, 41] and may demonstrate survivor bias [42] due to missing PRO data from individuals with diminished health status, increased symptoms from the disease, or toxicity from the assigned treatment. Individuals who do complete the PRO instrument are often healthier than those with missing PRO data which may compromise the external validity of the RCT. This is because analyses of between-group mean change scores include participants who provide PRO data both before randomization and also during treatment or follow-up, so the results would be limited to a subset and not generalizable to all patients with similarity to those having entered the RCT [42].

Missing PRO data may impact the adoption of evidence into clinical practice. Two reviews indicated reporting of RCTs with PROs as an endpoint was suboptimal. Fourteen percent of 97 RCTs that used the QLQ-C30 or SF-36 discussed the potential influence of missing PRO data on the results in the paper [43]. Eighteen percent of 66 RCTs that used the QLQ-C30 or another PRO instrument explained the reasons for missing PRO data, and 29% of the RCTs discussed the clinical significance of the PRO findings, in the paper [44]. Together, these data suggest that clinical interpretation of the published results of RCTs is made challenging by the presence of missing PRO data, and evidence to support decision making is lacking. Of note, only two RCTs were of high quality and low risk of bias and provided information in support of clinical decision making [44].

To summarize, in the presence of missing PRO data, allowing more participants to enter the RCT will increase the power and will improve the precision of parameter estimates, reflected in narrower confidence intervals, but will not eliminate bias from missing PRO data due to systematic error [29].

### 2.3.5 Approaches to Preventing and Dealing with Missing PRO Data

One of the ways of dealing with missing PRO data is in the analysis of the trial findings. Statistical analysis techniques, however, depend on unverifiable assumptions and will not resolve bias that is present from missing PRO data when MNAR. No technique is definitive, and different models provide different results [45 - 47]. Readers with an interest in statistical analysis techniques for dealing with missing PRO data may refer to other papers as this topic is outside the scope of the thesis. Of note, an international consortium – Setting International Standards in Analyzing Patient-Reported Outcomes and Quality of Life Endpoints Data for Cancer Clinical Trials (SISAQOL) – is identifying appropriate statistical methods to analyze PRO data and to evaluate missing data [48].

Consistent with an evolving realization that missing PRO data are a serious and complex problem, there is an emerging consensus from the plethora of publications within the last two decades that preventing missing PRO data is the best approach [49], but how? The following paragraphs summarize the historical approach and highlight the limitations of different strategies.

Early on, a lack of commitment from scientists and funding bodies was considered to be the crux of the problem with missing PRO data [50]. Staff required training, support, and resources to conduct PRO research.

More recently, an expert panel gathered information at a one day workshop and provided a report [51] that intended for use by the Food and Drug Administration to influence phase III RCTs. The panel did not include representation from Canada, clinicians or cooperative oncology groups. The panel made 18 recommendations [52], emphasized that RCTs should be designed and conducted in a way that limits missing data and published ideas how this may be done. For example, RCTs should be designed for a population that is not adequately served by current treatments, and conducted by investigators with prior competence in collecting complete data from participants [27]. These recommendations appear to be based on experience and may be difficult to implement in practice.

Evidence-based guidance for study design and conduct has been established. Li *et al* performed a systematic review followed by a Delphi process for consensus and proposed four mandatory standards: define research question, take steps in design and conduct to minimize missing data, continue collecting information on key outcomes, and monitor missing data [53]. Mercieca-Bebber *et al* performed a systematic review and identified several strategies to reduce missing data: incorporate PRO assessment schedules and termination rules in the protocol, reduce patient burden, identify individuals to coordinate PRO activities, provide staff training and ensure adequate resources [54].

Further, evidence-based guidance for study protocols and publications has been established by authors who performed a systematic review followed by a Delphi process for consensus that resulted in checklists for clinical trialists. The checklists provide a PRO extension to Standard Protocol Items Recommendations for Interventional Trials (SPIRIT), and to Consolidated Standards of Reporting Trials (CONSORT) [55, 56]. SPIRIT [57] recommends protocols contain methods for data collection and a plan to promote participant retention, but is not specific for the design and conduct of phase III RCTs with PRO. The SPIRIT-PRO extension [55] includes guidance for creating an RCT protocol with PROs. For example, the protocol should specify strategies for minimizing missing PRO data, and the protocol should state how missing PRO data will be handled in the analysis. The CONSORT-PRO extension [56] recommends that publications include a clear statement of the extent of missing PRO data as well as the statistical approaches used to deal with missing PRO data, but does not identify preferred methodology [58], nor does it specify what should have been done to prevent missing PRO data.

In summary, improving study design and conduct are considered the best approach to preventing missing PRO data. The evidence-based guidance documents are laudable but provide limited empirical evidence on which factors influence missing PRO data. The PRO results of RCTs can benefit the increasing numbers of people living with cancer [59, 60]. As described in Chapter 3, most (57%) of the 100 papers identified in the systematic review reported factors for missing PROs in cancer patient populations. Accordingly, Chapter 4 focuses on quantitative analyses of data from patients with cancer recruited onto one of three selected RCTs conducted by the CCTG at Canadian member-centres. There were several reasons for this decision. CCTG has a history of collecting PRO data in RCTs, mandated for all CCTG-led phase III RCTs for over three decades within Canadian member-centres. Twenty-three RCTs were eligible for analysis. Three RCTs were selected from the 10 having curative intention of treatment to provide variety in context and large sample size: from the breast disease site (MA17), the genito-urinary disease site (PR7), and the hematologic disease site (LY12). Chapter 5 focuses on qualitative inquiry of CRAs having experience with missing PRO data in cancer RCTs from working with patients at different CCTG centres in Canada.

2.4 References


[1] Mayo NE. *Dictionary of quality of life and health outcomes measurement, first edition.* International Society for Quality of Life Research: Lexington, 2015.

[2] Food and Drug Administration. *Guidance for industry. Patient-reported outcome measures: use in medical product development to support labeling claims.* 2009.

[3] Zubrod CG, Schneiderman M, Frei E et al. Appraisal of methods for the study of chemotherapy of cancer in man: comparative therapeutic trial of nitrogen mustard and triethylene thiophosphoramide. *J Chron Dis* 1960;11:7-33.

[4] Freyd M. The graphic rating scale. *J Educ Psychol* 1923;14:83-102.

[5] Wewers ME, Lowe NK. A critical review of visual analogue scales in the measurement of clinical phenomena. *Res Nurs Health* 1990;13:227-236.

[6] Likert R. A technique for the measurement of attitudes. *Arch Psychol* 1932;140:1-55.

[7] Fayers PM, Hand DJ, Bjordal K, Groenvold M. Causal indicators in quality of life research. *Qual Life Res* 1997;6:393-406.

[8] Ware J, Snow KK, Kosinksi M et al. *SF-36 health survey manual and interpretation guide*. The Health Institute: Boston, 1993.

[9] Fayers PM, Machin D. *Quality of life: the assessment, analysis and interpretation of patient-reported outcomes, second edition*. John Wiley & Sons Ltd: West Sussex, 2007.

[10] Aaronson NK, Ahmedzai S, Bergman B et al. The European Organisation for Research and Treatment of Cancer QLQ-C30: A quality-of-life instrument for use in international clinical trials in oncology. *J Ntl Cancer Inst* 1993;85:365-376.

[11] Petersen MA et al. Development of computerised adaptive testing (CAT) for the EORTC QLQ-C30 dimensions – general approach and initial results for physical functioning. *Eur J Cancer* 2010;46:1352-1358.

[12] Bottomley A, Aaronson NK. International perspective on health-related quality-of-life research in cancer clinical trials: the European Organisation for Research and Treatment of Cancer experience. *J Clin Oncol* 2007;25:5082-5086.

[13] Webster K, Cella D, Yost K. The functional assessment of chronic illness therapy (FACIT) measurement system: properties, applications, and interpretation. *Health Qual Life Outcomes* 2003;1:79.

[14] Lamberts MP et al. Prospective multicentre cohort study of patient-reported outcomes after cholecystectomy for uncomplicated symptomatic cholecystolithiasis. *Brit J Surg* 2015;102:1402-1409.

[15] Priestman TJ, Baum M. Evaluation of quality of life in patients receiving treatment for advanced breast cancer. *Lancet* 1976;1:899-901.

[16] Sedgwick P. Phases of clinical trials. *BMJ* 2011;343:d6068 doi: 10.1136/bmj.d6068.

[17] Stockler MR et al. Patient-reported outcome results from the open-label phase III Aurelia trial evaluating bevacizumab-containing therapy for platinum-resistant ovarian cancer. *J Clin Oncol* 2014;32:1309-1316.

[18] Szklo M, Nieto FJ. *Epidemiology beyond the basics, second edition.* Jones and Bartlett Publishers: Sudbury, 2007.

[19] Au H-J, Ringash J, Brundage M, Palmer M, Richardson H, Meyer RM. Added value of health-related quality of life measurement in cancer clinical trials: the experience of the NCIC CTG. *Expert Rev Pharmacoecon Outcomes Res* 2010;10:119-128.

[20] Osoba D, Zee B. Completion rates in health-related quality-of-life assessment: approach of the National Cancer Institute of Canada Clinical Trials Group. *Stat Med* 1998;17:603-612.

[21] Curran D, Bacchi M, Hsu Schmitz SF, Molenberghs G, Sylvester RJ. Identifying the types of missingness in quality of life data from clinical trials. *Stat Med* 1998;17:739-756.

[22] Fairclough DL. *Design and analysis of quality of life studies in clinical trials, second edition*. Chapman & Hall / CRC, 2010.

[23] Rubin DB. Inference and missing data. *Biometrika* 1976;63:581-592.

[24] Moinpour CM et al. Challenges posed by non-random missing quality of life data in an advanced-stage colorectal cancer clinical trial. Psycho-Oncology 2000;9:340-354.

[25] CIOMS. International ethical guidelines for biomedical research involving human subjects. Council for International Organizations of Medical Sciences. Geneva, 2002.

[26] White E, Armstrong BK, Saracci R. *Principles of exposure measurement in epidemiology. Collecting, evaluating, and improving measures of disease risk factors. Second edition*. Oxford University Press, 2008.

[27] Little RL, D'Agostino R, Cohen ML et al. The prevention and treatment of missing data in clinical trials. *N Engl J Med* 2012;367:1355-1361.

[28] Fayers PM, Aaronson NK, Bjordal K, Groenvold M, Curran D, Bottomley A. *The EORTC QLQ-C30 Scoring Manual (3rd Edition)*. Published by: European Organisation for Research and Treatment of Cancer, Brussels, 2001.

[29] Little RJ, Cohen ML, Dickersin K et al. The design and conduct of clinical trials to limit missing data. *Stat Med* 2012;31:3433-3443.

[30] Porta M. *A dictionary of epidemiology, fifth edition.* Oxford University Press: New York, 2008.

[31] Sedgwick P. Randomised controlled trials: the importance of sample size. *BMJ* 2015;350:h1586 doi: 10.1136/bmj.h1586.

[32] Sloan JA, Dueck AC, Erickson PA, Guess H, Revicki DA, Santanello NC. Analysis and interpretation of results based on patient-reported outcomes. *Value Health* 2007;10(Supplement 2):S106-S115.

[33] McIsaac MA, Cook RJ, Poulin-Costello M. Incomplete data in randomized dermatology trials: consequences and statistical methodology. *Dermatology* 2013;226:19-27.

[34] Newman DA. Missing data: five practical guidelines. *Organ Res Meth* 2014;17:372-411.

[35] Dueck AC, Frost MH, Halyard MY, Sloan JA. Handling missing data. *Curr Probl Cancer* 2005;November/December:317-325.

[36] Sedgwick P. Controlled trials: allocation concealment, random allocation, and blinding. *BMJ* 2015;350:h2633.

[37] Ross L, Thomsen BL, Boesen EH, Johansen C. In a randomized controlled trial, missing data led to biased results regarding anxiety. *J Clin Epidemiol* 2004;57:1131-1137.

[38] Sterne JAC, White IR, Carlin JB et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009;339(18 July):157-160.

[39] Stang A. Nonresponse research – an underdeveloped field in epidemiology. *Eur J Epidemiol* 2003;18:929-931.

[40] Neymark N et al. Methodological and statistical issues in quality of life and economic evaluation in cancer clinical trials: report of a workshop. *Eur J Cancer* 1998;34:1317-1333.

[41] Rock EP, Scott JA, Kennedy DL, Sridhara R, Pazdur R, Burke LB. Challenges to use of health-related quality of life for food and drug administration approval of anticancer products. *J Natl Cancer Inst Monogr* 2007;37:27-30.

[42] Brundage M et al. Lessons learned in the assessment of health-related quality of life: selected examples from the National Cancer Institute of Canada Clinical Trials Group. *J Clin Oncol* 2007;25:5078-5081.

[43] Rombach I, Rivero-Arias O, Gray AM, Jenkinson C, Burke O. The current practice of handling and reporting missing outcome data in eight widely used PROMs in RCT publications: a review of the current literature. *Qual Life Res* 2016;25:1613-1623.

[44] Rees JR et al. Patient-reported outcomes in randomised controlled trials of colorectal cancer: an analysis determining the availability of robust data to inform clinical decision-making. *J Cancer Res Clin Oncol* 2015;141:2181-2192.

[45] King MT et al. Analyzing patient-reported outcome endpoints in advanced cancer clinical trials: responder analysis versus mixed model for repeated measures illustrated with results from the AURELIA ovarian cancer trial. *Qual Life Res* 2013;22:15.

[46] Moinpour CM et al. Evaluating health-related quality-of-life therapeutic effectiveness in a clinical trial with extensive nonignorable missing data and heterogeneous response: results from a phase III randomized trial of gemcitabine plus paclitaxel versus paclitaxel monotherapy in patients with metastatic breast cancer. *Qual Life Res* 2012;21:765-775.

[47] Peyre H, Leplege A, Coste J. Missing data methods for dealing with missing items in quality of life questionnaires. A comparison by simulation of personal mean score, full information maximum likelihood, multiple imputation, and hot deck techniques applied to the SF-36 in the French 2003 decennial health survey. *Qual Life Res* 2011;20:287-300.

[48] Bottomley A, Pe M, Sloan J et al. Moving forward toward standardizing analysis of quality of life data in randomized clinical trials. *Clinical Trials* 2018 doi:10.1177/1740774518795637.

[49] Vickers AJ, Altman DG. Statistics notes: missing outcomes in randomised trials. *BMJ* 2013;346:f3438 doi:10.1136/bmj.f3438.

[50] Moynihan C. Patient 'non-compliance' and 'missing data' in quality of life research: where does the problem lie? *Eur J Cancer* 1998;34:9-11.

[51] National Research Council. *The Prevention and Treatment of Missing Data in Clinical Trials.* Panel on Handling Missing Data in Clinical Trials. Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press, 2010.

[52] O'Neill RT, Temple R. The prevention and treatment of missing data in clinical trials: an FDA perspective on the importance of dealing with it. *Clin Pharmacol Ther* 2012;91:550-554.

[53] Li T,  Hutfless S, Scharfstein DO et al. Standards should be applied in the prevention and handling of missing data for patient-centred outcomes research: a systematic review and expert consensus. *J Clin Epidemiol* 2014;67:15-2.

[54] 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**:e010938.

[55] Calvert M, Kyte D, Mercieca-Bebber R, et al. Guidelines for inclusion of patient-reported outcomes in clinical trial protocols: the SPIRIT-PRO extension. *JAMA* 2018;319:483-494.

[56] Calvert M et al. Reporting of patient-reported outcomes in randomized trials: the CONSORT PRO extension. *JAMA* 2013;309:814-822.

[57] Chan A-W et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. *BMJ* 2013;346:e7586

[58] Altman DG. Missing outcomes in randomized trials: addressing the dilemma. *Open Medicine* 2009;3:51-53.

[59] Canadian Cancer Statistics Advisory Committee. *Canadian Cancer Statistics 2019.* Toronto, ON: Canadian Cancer Society, 2019. Available at: cancer.ca/Canadian-Cancer-Statistics-2019-EN

[60] Canadian Partnership Against Cancer. Living with Cancer: A Report on the Patient Experience. Toronto, ON: Canadian Partnership Against Cancer; 2018.

# Chapter 3

## A Systematic Review and Development of a Classification Framework for Factors Associated with Missing Patient-Reported Outcome Data

Michael J Palmer, Rebecca Mercieca-Bebber, Madeleine King, Melanie Calvert, Harriet Richardson, Michael Brundage, A systematic review and development of a classification framework for factors associated with missing patient-reported outcome data, *Clinical Trials* (volume 15 and issue 1) pp. 95-106.  doi: 10.1177/1740774517741113 journals.sagepub.com/home/ctj

3.1 Abstract

**Background/Aims:** Missing patient-reported outcome (PRO) data can lead to biased results, to loss of power to detect between-treatment differences, and to research waste. Awareness of factors may help researchers reduce missing PRO data through study design and trial processes. The aim was to construct a Classification Framework of factors associated with missing PRO data in the context of comparative studies. The first step in this process was informed by a systematic review.

**Methods:** Two databases (MEDLINE and CINAHL) were searched from inception to March 2015 for English articles. Inclusion criteria were: a) relevant to PRO; b) discussed missing data or compliance in prospective medical studies; c) examined predictors or causes of missing data, including reasons identified in actual trial datasets and reported on cover sheets. Two reviewers independently screened titles and abstracts. Discrepancies were discussed with the research team prior to finalizing the list of eligible papers. In completing the systematic review, four particular challenges to synthesizing the extracted information were identified. To address these challenges, operational principles were established by consensus to guide the development of the Classification Framework.

**Results:** 6027 records were screened. 100 papers were eligible and included in the review. Of these, 57% focused on cancer, 23% did not specify disease, and 20% reported for patients with a variety of non-cancer conditions. 40% of the papers offered a descriptive analysis of possible factors associated with missing data, but some papers used other methods. 663 excerpts of text (units), each describing a factor associated with missing PRO data, were extracted verbatim. Redundant units were identified and sequestered. Similar units were grouped, and an iterative process of consensus among the investigators was used to reduce these units to a list of factors that met the guiding principles. The list was organized on a framework, using an iterative consensus-based process. The resultant Classification Framework is a summary of the factors associated with missing PRO data described in the literature. It consists of five components (Instrument, Participant, Centre, Staff, and Study) and 46 categories, each with one or more sub-categories or examples.

**Conclusions:** A systematic review of the literature revealed 46 unique categories of factors associated with missing PRO data, organized into five main component groups. The Classification Framework may assist researchers to improve the design of new RCTs and to implement procedures to reduce missing PRO data. Further research using the Classification Framework to inform quantitative analyses of missing PRO data in existing clinical trials and to inform qualitative inquiry of research staff is planned.

## 3.2 Introduction

Patient-reported outcomes (PROs) include measures from the patient's perspective of symptoms associated with disease, toxicity arising from treatment, the ability to function and to perform activities of daily living, and different dimensions of health-related quality of life [1]. PROs have three characteristics: the measurement is directly from the patient, can change over time, and is often multi-dimensional [2]. In contrast are clinician-assessed indicators of health status, such as performance status [3].

Many questionnaires have been developed to measure PROs. These often include questions comprising scales about global health status and health-related quality of life, and about different domains and dimensions of interest [4]. Generic questionnaires like the SF-36 [5] provide a general assessment of health status regardless of the health condition, whereas disease-specific questionnaires like the QLQ-C30 [6] provide a comprehensive assessment with focus on specific aspects of the health condition and its treatment [7]. Sometimes, additional questions are used in a PRO instrument to supplement the questionnaire for measurement depth relevant to the clinical setting [8].

In the context of a randomized clinical trial (RCT) that includes the assessment of PROs as an objective, typically the first assessment is before randomization, to measure health status at baseline before delivery of the assigned treatment. Subsequent assessments after randomization measure health status during treatment and follow-up. Health-status scores at specific time points and changes from baseline can be determined and may demonstrate between-arm differences, thus providing important information on treatment efficacy and toxicity. The PRO results of an

RCT can thus help individuals choose the best treatment and enrich understanding of patient experiences when considered with other trial outcome data [9].

A particular challenge with collecting and analyzing PROs in RCTs is missing data, which arise when a study participant fails to complete an assessment, fails to attend an evaluation, or is unavailable because of illness or death [1]. Operationally, PRO data are missing when an assessment was expected but the patient did not complete all or part of the instrument. By this definition, if there was no expectation for a patient to complete an instrument (e.g., after progression or death) the PRO data are absent but are not considered to be missing. Missing PRO data may indicate the entire instrument is missing, or that the instrument is partially completed and has missing items, at a given assessment [10].

Three different mechanisms for missing PRO data have been described [11]. Missing completely at random occurs when the missing PRO data are independent of the observed PRO data, for example, when a staff member forgets to administer an instrument. Missing at random occurs when the missing PRO data are dependent on the observed PRO data (e.g., a participant misses the first assessment after randomization following self-report of poor health status at baseline) or other observed covariate (e.g., employment status, such as when an employed participant, compared with a retiree, is too busy to complete the instrument). Missing not at random occurs when the missing PRO data are dependent on the unobserved PRO data at the time of the scheduled assessment, for example, when a patient with diminished health-related quality of life does not attend the clinic because of severe symptoms.

Missing PRO data are not without consequence. PROs that are missing waste resources and, if extreme, reflect unethical research [12, 13]. Missing PRO data are a serious threat to the internal and external validity of RCTs, can undermine the scientific credibility of causal conclusions [14] and may impact the adoption of evidence into clinical practice.

The presence of missing items may preclude calculation of a score for some scales within an instrument. For example, if the questionnaire scoring strategy requires half or more of the scale items to be complete in order to produce that scale score, missing items may result in the scale for a patient to be not included in the analysis.

Missing PRO data may lead to a loss of power [15]. Sample size for an RCT is typically calculated for the primary objective. In symptom management trials, including pain trials, the PRO endpoint may be the primary endpoint and thus forms the basis of the sample size calculation. However, in the cancer setting, PROs are often included as secondary endpoints, and primary endpoints such as disease-free survival or overall survival form the basis for the sample size calculation [16]. Whether the amount of missing PRO data, commensurate with a reduction in the effective sample size, results in a loss of power that is, by convention, lower than 80%, will depend on the individual RCT [17]. Estimates of outcome parameters with fewer data will lack precision from random error and will lead to wider confidence intervals [18] which may cause type I or type II errors [19]. Allowing more participants to enter the RCT will increase the power and will improve the precision of parameter estimates but will not eliminate bias [20].

Randomization balances known and unknown confounders [21] but will not protect against potential bias from missing PRO data [22] whose risk depends on the mechanism [23] not on the magnitude of the response proportion [24]. With missing completely at random, there are no systematic differences in missing and observed values. With missing at random, any systematic differences are explainable by differences in the observed values. With missing not at random, systematic differences persist between the missing and observed values. Bias, then, may arise when the PRO data are missing at random or missing not at random. Of these, missing not at random is considered to be the more serious problem [25, 26], such as survivor bias in cancer RCTs [27]. In this context, patients who complete the PRO instrument are often healthier than those with missing PRO data who may compromise the external validity of the RCT by limiting the results to a subset and not generalizable for all patients with similarity to those having entered the RCT [27].

Missing PRO data may impact the adoption of evidence into clinical practice. Two recent reviews indicated reporting of RCTs with PROs as an endpoint was suboptimal [28, 29]. Together, these data suggest that clinical interpretation of the published results for RCTs is made challenging by the presence of missing PRO data, and evidence to support decision making is lacking.

Clearly, missing PRO data are an important problem. An early initiative to understand which factors associated with missing PRO data in multi-centre RCTs involved representatives of 12 cooperative oncology groups who described their experience at a landmark international workshop. The reports were published in a special issue of *Statistics in Medicine* (volume 17, 1988) and sporadic publications since then have described missing PRO data. Several factors were thought to associate with missing instruments but less was known with missing items [30]. Collectively, three studies [31-33] demonstrated that characteristics of the instruments, patients, institutions and RCTs were significant factors associated with missing instruments. A few studies provided examples illustrating specific questions in an instrument such as regarding feelings of self-consciousness [34], religious issues [35] or sexual relationships [36] may be the source of missing items. Although individuals with older age, poorer physical functioning and lower education were thought to have more missing PRO data [30], evidence for age and missing instruments was contradictory [31-33] while evidence for age and missing items was scant [32].

The literature addressing missing PRO data is difficult to interpret, in part due to differences in study design, sources of data, selection and definition of variables (factors), assessment time points, approaches to modelling, the mix of RCTs and international context. Some predictors of missing PRO data, like patient age, have been identified by regression analyses, and some causes of missing PRO data, like symptom burden, have been advanced by expert opinion, but a comprehensive inventory of factors associated with missing PRO data has not yet been constructed. Such an inventory would have potential utility in preventing missing PRO data and in guiding further research. The objectives of this study were to perform a systematic review of the literature, and to create a Classification Framework for factors reported to be associated with missing PRO data.

## 3.3 Methods

The purpose of the systematic review was to identify reported factors associated with missing PRO data. MEDLINE and CINAHL databases were searched from inception to March 2015 for all records on criteria of PRO terms and missing data terms that were unrestricted with respect to disease and defined through an iterative process of development with an index paper [14] and with guidance from health librarians (Table 1 and Table 2).

Titles and abstracts were independently screened in parallel against *a priori* eligibility criteria by two reviewers (MP, RM-B). The first 200 titles and abstracts were reviewed in a pilot phase. Results were discussed to clarify interpretation and to standardize the screening procedures. The remaining titles and abstracts were then screened against the final eligibility criteria. Discrepancies were resolved through discussion among the research team in regular meetings and every decision was documented. The majority of decisions were based on consensus. Sometimes, arbitration by the senior member (MB) was required.

Records were included if they: a) were relevant to PROs; b) discussed missing data or compliance in prospective health care or medical research studies, including RCTs; c) examined predictors or causes of missing data, including reasons identified in actual trial datasets and reported on cover sheets. A record was excluded if it was a: a) non-English publication; b) case-control study, c) conference presentation, d) cross-sectional report noting missing data, e) non-patient sample, f) protocol; or if the focus was on: g) response shift, h) instrument development, field testing or translation, i) proxy assessment, j) statistical handling of missing data.

Articles were obtained and the full text was assessed against the same eligibility criteria. A hand-search of references was performed for included articles to identify additional records and the full text was also assessed against the same eligibility criteria. Each paper was reviewed in its entirety and every reported factor associated with missing PRO data was identified, extracted verbatim with key information about the source paper and entered into a database.

In completing the systematic review, four particular challenges to synthesizing the extracted information were identified. First, 30 authors suggested descriptive elements for missing PRO data, but there was a lack of consistency on how these authors grouped similar units. For example, one author [37] reported staff burden consisted of time constraints, unwillingness to collect PRO data and inability to collect PRO data, while another author [31] reported staff factors consisted of staff limitations only. Second, authors sometimes used different wording to convey the same idea (redundancy). For example, 'patient refusal' was reported in 27 different ways. Third, ambiguity in some units was evident. Fourth, the collective information had limited application in the absence of structure.

Operational principles were established by consensus, to address each of the four identified challenges and to guide the development of the Classification Framework. First, similar units should be grouped together in a manner that demonstrated cohesiveness of logical types. Second, each unit should be unique and enter only one category, and each category should be unique and enter only one component, to reflect mutual exclusivity. Third, unambiguous wording of content was essential for clarity. Fourth, the final product had to be fit for purpose.

Taxonomy was used to create explicit classification within a hierarchical and multidimensional structure [38]. A familiar approach, involving the answers to five questions – what, who, where, when, why – helped us to determine key components for classifying missing PRO data within a framework. The instrument (consisting of a series of questions with options for answers) required completion. The participant (individual with a health condition) completed the instrument. The centre (hospital, clinic) was the facility where the participant completed the instrument. Staff (research personnel at the centre) informed the participant when to complete the instrument. Why this was necessary was because of the study. In general, an instrument was expected to be completed by a participant at a centre when requested by staff, according to the study [16].

Iterative development ensured robustness. Redundant units and units that were not consistent with the operational definition of missing PRO data were removed. Similar units were organized in groups. An iterative process of consensus reduced these units to a list of factors that met the guiding principles. The list was organized on the framework, using an iterative consensus-based process.

3.4 Results

Figure 1 provides a flow diagram [39] illustrating how the included papers were identified. After screening 6027 records, 100 papers from 70 journals met the eligibility criteria (the on-line appendix provides a list of the included papers and a summary table of the journals). In all, 57% of the 100 papers focused on cancer, 23% did not specify disease, and 20% reported for patients with a variety of non-cancer conditions. These included: acquired immune deficiency syndrome, anxiety, astigmatism, brain aneurysm, cataract, chronic obstructive pulmonary disease, chronic pruritis, coronary artery bypass graft, critical illness, depression, insomnia, knee pain, liver morbidity, lumbar spine disorders, multiple sclerosis, prostate morbidity, rheumatoid arthritis, substance abuse, temporomandibular disorders, total hip replacement or total joint arthroplasty.

Different methods were used in the papers to identify possible factors associated with missing PRO data (Table 3). Most of the papers (40%) just provided a descriptive analysis of missing PRO data, 19% were solely based on expert opinion, and 10% were solely based on regression analysis. Six hundred sixty three excerpts of text (units), each describing a factor associated with missing PRO data, were extracted verbatim from the 100 papers.

Figure 2 illustrates the cumulative factors associated with missing PRO data reported by the papers. Prior to 1998, in general each paper contributed one factor. Nine reports from the international workshop contributed 27 factors in 1998. Subsequently, the yield diminished with no new factors identified since 2012. Overall, 38 papers contributed at least one factor associated with missing PRO data (range one to seven, median one).

The resultant Classification Framework (Figure 3) is a summary of the factors associated with missing PRO data in the literature, consisting of five components (Instrument, Participant, Centre, Staff, and Study) and 46 categories, each with one or more sub-categories or examples. Some categories within Participant and within Centre, but no categories within Instrument, Staff or Study, were based on regression analyses.

3.5 Discussion

This study was the first to perform a systematic review of the literature, to extract verbatim the reported factors associated with missing PRO data (units) from the included papers, and to classify the factors within a hierarchical and multidimensional framework according to the presumed relationships among components and between categories. The Classification Framework is a robust inventory of the factors associated with missing PRO data identified in the literature.

The Classification Framework organizes the factors associated with missing PRO data within five key components. These in turn are described by 46 different categories with each having one or more sub-categories. For example, missing PRO data may occur because of characteristics of the instrument (e.g. clarity), participant (e.g. sociodemographic characteristics), centre (e.g. level of activity), staff (e.g. continuity in workload) and study (e.g. intention of treatment).
We chose not to assign weights to each factor because the majority of papers had no statistical testing, thus the Classification Framework presents an inventory, not a ranking, of factors. Accordingly, each entry should be considered relevant to missing PRO data in the context of one or more settings reported in the literature.

How each factor may lead to missing PRO data, and how certain factors may be linked, requires further study. These mechanisms of action are largely unknown and are likely context specific. For example, missing PRO data may be associated with older age or with younger age depending on the context of the RCT. Additional work needs to be done to determine the mechanisms of action and modifiers of causal effect for missing PRO data.

It is naïve to believe that an RCT with assessment of PROs included as an objective will have zero missing PRO data, given the ethical principle of respect for autonomy [19]. Every individual who consents to an RCT has the capacity to complete the PRO instrument but may choose not to. The reason why individuals do not complete the entire PRO instrument, or why individuals provide responses for some but not for all of the questions, may or may not be known. Missing PRO data will occur on every RCT. The Classification Framework raises awareness by providing concise information on the specific factors associated with missing PRO data in a convenient format.

One of the ways of dealing with missing PRO data is by statistical analysis techniques that may depend on unverifiable assumptions and will not resolve bias that is present when the mechanism is missing not at random. No technique is definitive, and different models provide different results [40-42]. Guidelines for analysis are insufficient and a consortium has been created with intention to set international standards [43]. Readers with an interest in statistical analysis techniques for dealing with missing PRO data may refer to other papers as this topic is outside the scope of the current research.

Consistent with an evolving realization that missing PRO data are a serious and complex problem, there is an emerging consensus that preventing missing PRO data is the best approach [44], but to comprehend this is not as clear. Historically, a lack of commitment from scientists and funding bodies was considered to be the crux of the problem with missing PRO data [45]. Staff required training, support and resources to conduct PRO research. More recently, improving study design and conduct were considered the best approach to preventing missing PRO data [46]. An expert panel made 18 recommendations [47], emphasized that RCTs should be designed and conducted in a way that limits missing data, and published ideas how this may be done. For example, RCTs should be designed for a population that is not adequately served by current treatments, and conducted by investigators with prior competence in collecting complete data from participants [14]. These recommendations appear to provide guidance based on experience but may be difficult to implement in practice.

Evidence-based guidance for study design and conduct has been established. Li *et al* performed a systematic review followed by a Delphi process of consensus and proposed four mandatory standards: define research question, take steps in design and conduct to minimize missing data, continue collecting information on key outcomes and monitor missing data [48]. Subsequently, Mercieca-Bebber *et al* performed a systematic review and identified several strategies for reducing missing data including: incorporate PRO assessment schedules and termination rules in the protocol, reduce patient burden, identify individuals to coordinate PRO activities, provide staff training, ensure adequate resources [49]. This review points out that in some settings missing PRO data are preventable through rigorous study design and conduct, whereas in other settings, missing PRO data are unavoidable because of deteriorating health status.

Further, evidence-based guidance for protocols and publications has been established by authors who performed a systematic review followed by a Delphi process of consensus. One checklist [50] recommends protocols contain methods for data collection and a plan to promote participant retention. An extension provides guidance specific for the design and conduct of RCTs with PRO endpoints [51]. Another checklist [52] recommends publications include a clear statement of the statistical approaches for dealing with missing PRO data, but does not identify good methodology [53] nor does it specify what should be done to prevent missing PRO data.

Collectively, the previous evidence-based guidance documents are laudable for reinforcing the need to address missing PRO data from the design through reporting stages. The limited empirical evidence in these documents on the factors associated with missing PRO data is an impediment to researchers, especially to researchers having little or no experience with the collection of PROs from patients in RCTs. An important step towards being able to prevent missing PRO data is first understanding specific factors are associated with missing PRO data. Our research builds on existing knowledge and the product contributes evidence-based guidance that can facilitate understanding. Of benefit for all personnel who design and conduct RCTs, the Classification Framework has potential use in preventing avoidable missing PRO data. This would involve consideration of the factors provided within the Classification Framework, then implementation of an approach to reduce missing PRO data [49], with recognition of the context. For example, in an RCT of palliative treatment for cancer, missing PRO data from staff error in processing the instrument may be prevented through staff training, but missing PRO data from progressing disease may be unavoidable.

The Classification Framework has potential as a valuable research tool. Most categories within the five components were not based on regression analysis. Our research team will systematically assess each component with regression analyses of pooled individual participant data from RCTs to determine which, if any, characteristics of the instruments, participants, centres and studies predict missing PRO data. The quantitative analyses will potentially provide 1A level of evidence for three criteria [54] (strength, consistency, biological gradient) and one additional criterion (temporality) of causality, but these are limited by the availability of data, especially regarding staff factors.

Of note, only one paper [55] used a qualitative method to identify reasons for missing PRO data and reinforced the notion that staff are an important – but under-utilised – source for insight on missing PRO data. Our pilot focus group with research personnel confirmed staff, when asked, will provide their personal knowledge of the clinic environment and of patients who did not complete instruments on RCTs.  We are planning interviews with staff and anticipate their experience will reveal which factors are most important and what aspects of each component contribute to missing PRO data.

We expect the combination of quantitative analyses with qualitative inquiry will identify priorities for action. This may involve simple strategies, like ensuring the instrument length does not exceed a certain number of questions, or may require more effort, like development of an educational intervention to improve the knowledge of staff regarding the value of PROs.

Strengths of this study include (a) using the approach of the Canadian Institutes of Health Research for knowledge creation: the inquiry of existing knowledge led to the synthesis of information and creation of an original product [56]. (b) The search criteria for the systematic review were carefully defined through an iterative process of development with an index paper [14] and with guidance from health librarians, and applied to two electronic databases having extensive coverage of targeted records. (c) Specific procedures were implemented to enhance the accuracy and reliability of screening which increased the internal validity of the study. Two reviewers independently screened in parallel the titles and abstracts of all records identified through database searching against *a priori* eligibility criteria. (d) A hand-search of references was performed for included articles to identify additional records and the full text was also assessed against the same eligibility criteria. (e) The Classification Framework was developed through an iterative process of consensus based on guiding principles that addressed the particular challenges to synthesizing the extracted information identified in completing the systematic review.

A few limitations are recognized. (a) Our pragmatic approach restricted the systematic review to English-language records and these may be susceptible to publication bias. (b) Databases other than MEDLINE and CINAHL may contain eligible records and therefore our list of factors associated with missing PRO data for the Classification Framework may be incomplete. We are confident in its current content. (c) The quality of included papers was not assessed for the pragmatic reason that the majority did not report original research from comparative studies and our intention was not to extract information for a meta-analysis. Thus, indicators of quality commonly used in traditional systematic reviews were not applicable in this study.

In summary, the Classification Framework provides concise information on the factors associated with missing PRO data and will likely help researchers in three ways. First, to appreciate a finite number of specific characteristics existing within five components (instrument, participant, centre, staff and study) may be linked to missing PRO data. Second, if the content is used to improve the design of RCTs it should be possible to prevent some missing PRO data. Third, increased awareness will help researchers to implement procedures to reduce missing PRO data.

### 3.6 Acknowledgements

Preliminary work was presented at International Society for Quality of Life Research in Vancouver on October 23, 2015 and at Society for Clinical Trials in Montreal on May 17, 2016. The authors would like to thank Shawn Hendrikx, Clinical Outreach Librarian, Queen's University and, Rebecca Goldsworthy and Jeremy Cullis, Faculty Liaison Librarians, University of Sydney for their contribution in developing the search strategy.

**Table 3-1. Search criteria for MEDLINE.**

| # | PATIENT REPORTED OUTCOME TERMS | Results |
|---|---|---|
| 1 | "Quality of Life"/ | 122099 |
| 2 | (qualit* adj2 life).tw | 143391 |
| 3 | Self Report/ | 10061 |
| 4 | ((self or patient) adj1 report*).tw | 92411 |
| 5 | Health Status/ | 60612 |
| 6 | health status.tw | 35551 |
| 7 | (health adj2 level*).tw | 5799 |
| 8 | patient reported outcome*.tw | 3546 |
| 9 | Data Collection/st [standards] | 2924 |
| | **MISSING DATA TERMS** | |
| 10 | missing data.tw | 4920 |
| 11 | missing value*.tw | 1145 |
| 12 | missing response*.tw | 187 |
| 13 | missing item*.tw | 154 |
| 14 | (compliance adj3 ('quality of life' or 'patient reported outcome*' or QOL or QL or HRQOL or HRQL or PRO)).tw | 448 |
| 15 | (response rate adj1 ('quality of life' or 'patient reported outcome*' or QOL or QL or HRQOL or HRQL or PRO)).tw | 27 |
| 16 | completion rate*.tw | 2073 |
| 17 | attrition.tw | 6722 |
| 18 | drop out.tw | 3295 |
| 19 | los* to follow up.tw | 13520 |
| 20 | quality assurance.tw | 17751 |
| 21 | quality control.tw | 26823 |
| 22 | missing quality of life data.tw | 10 |
| 23 | missing patient reported outcome* data.tw | 2 |
| 24 | missing QOL data.tw | 5 |
| 25 | missing QL data.tw | 0 |
| 26 | missing HRQOL data.tw | 1 |
| 27 | missing HRQL data.tw | 0 |
| 28 | missing PRO data.tw | 1 |
| | **FINAL SEARCH RESULTS** | |
| 29 | or/1-9 | 339743 |
| 30 | or/10-28 | 74422 |
| 31 | 29 and 30 | 4516 |
| 32 | limit 31 to English language | 4152 |

**Table 3-2. Search criteria for CINAHL.**

| # | PATIENT REPORTED OUTCOME TERMS | Results |
|---|---|---|
| 1 | (MH "Quality of Life+") | 48,312 |
| 2 | TI quality of life OR AB quality of life | 44,283 |
| 3 | (MH "Self Report") | 36,134 |
| 4 | TI Self-report* or AB Self-report* | 27,531 |
| 5 | (MH "Health Status+") | 46,799 |
| 6 | TI health status OR AB health status | 14,952 |
| 7 | TI health level OR AB health level | 4605 |
| 8 | TI level of health OR AB level of health | 8612 |
| 9 | (MH "Outcome Assessment") | 17,816 |
| 10 | TI patient-report* or AB patient-report* | 3,020 |
| | **MISSING DATA TERMS** | |
| 11 | TI missing data OR AB missing data | 1444 |
| 12 | TI missing value* OR AB missing value* | 213 |
| 13 | TI missing response* OR AB missing response* | 95 |
| 14 | TI missing item* OR AB missing item* | 91 |
| 15 | TI completion rate* OR AB completion rate* | 926 |
| 16 | TI attrition OR AB attrition | 1991 |
| 17 | TI drop out OR AB drop out | 635 |
| 18 | TI los* to follow up OR AB los* to follow up | 1,720 |
| 19 | TI quality assurance OR AB quality assurance | 3434 |
| 20 | TI quality control OR AB quality control | 2987 |
| 21 | TI "missing quality of life data" OR AB" missing quality of life data" | 4 |
| 22 | TI “missing patient-reported outcome* data” OR AB “missing patient-reported outcome* data” | 1 |
| 23 | TI "missing QOL data" OR AB "missing QOL data" | 0 |
| 24 | TI "missing QL data" OR AB "missing QL data" | 0 |
| 25 | TI "missing HRQOL data" OR AB "missing HRQOL data" | 0 |
| 26 | TI "missing HRQL data" OR AB "missing HRQL data" | 0 |
| 27 | TI "missing PRO data" OR AB "missing PRO data" | 1 |
| | **FINAL SEARCH RESULTS** | |
| 28 | or/S1-S10 | 179,277 |
| 29 | or/S11-S27 | 12,978 |
| 30 | S28 and S29 | 2,452 |

**Table 3-3. How possible factors associated with missing PRO data were identified in 100 source papers.**

| Method | % |
|---|---|
| Descriptive analysis | 40 |
| Descriptive analysis and regression analysis | 8 |
| Descriptive analysis and statistical analysis | 5 |
| Descriptive analysis and focus group | 1 |
| Expert opinion | 19 |
| Expert opinion, and descriptive and statistical analyses | 1 |
| Regression analysis | 10 |
| Statistical analysis | 7 |
| Reference to another paper | 7 |
| Systematic review (with focus other than causes of missing PRO data) | 2 |

**Figure 3-1. Flow diagram.**

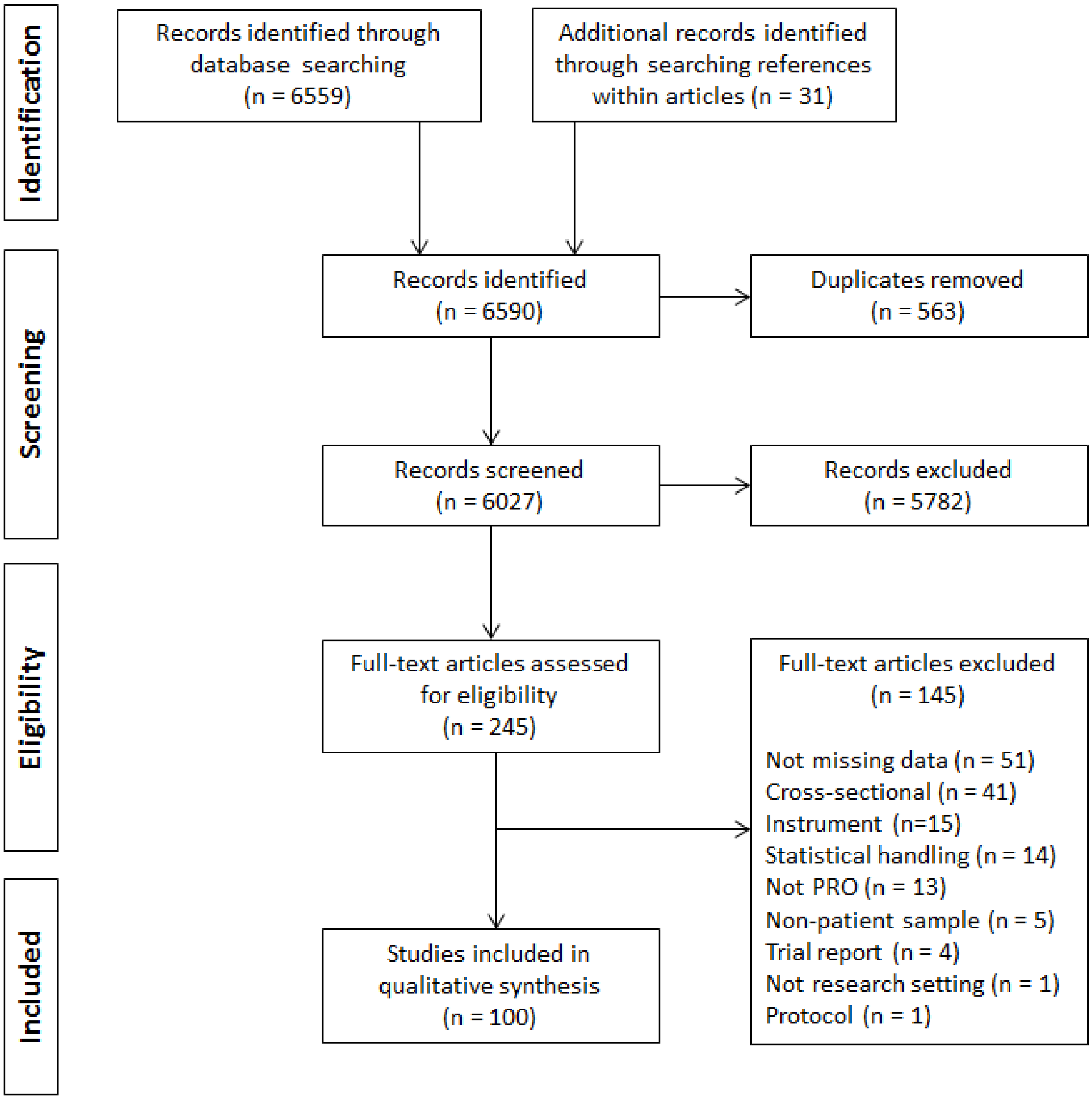


**Figure 3-2. Cumulative factors associated with missing patient-reported outcome data identified from papers in the systematic review.**

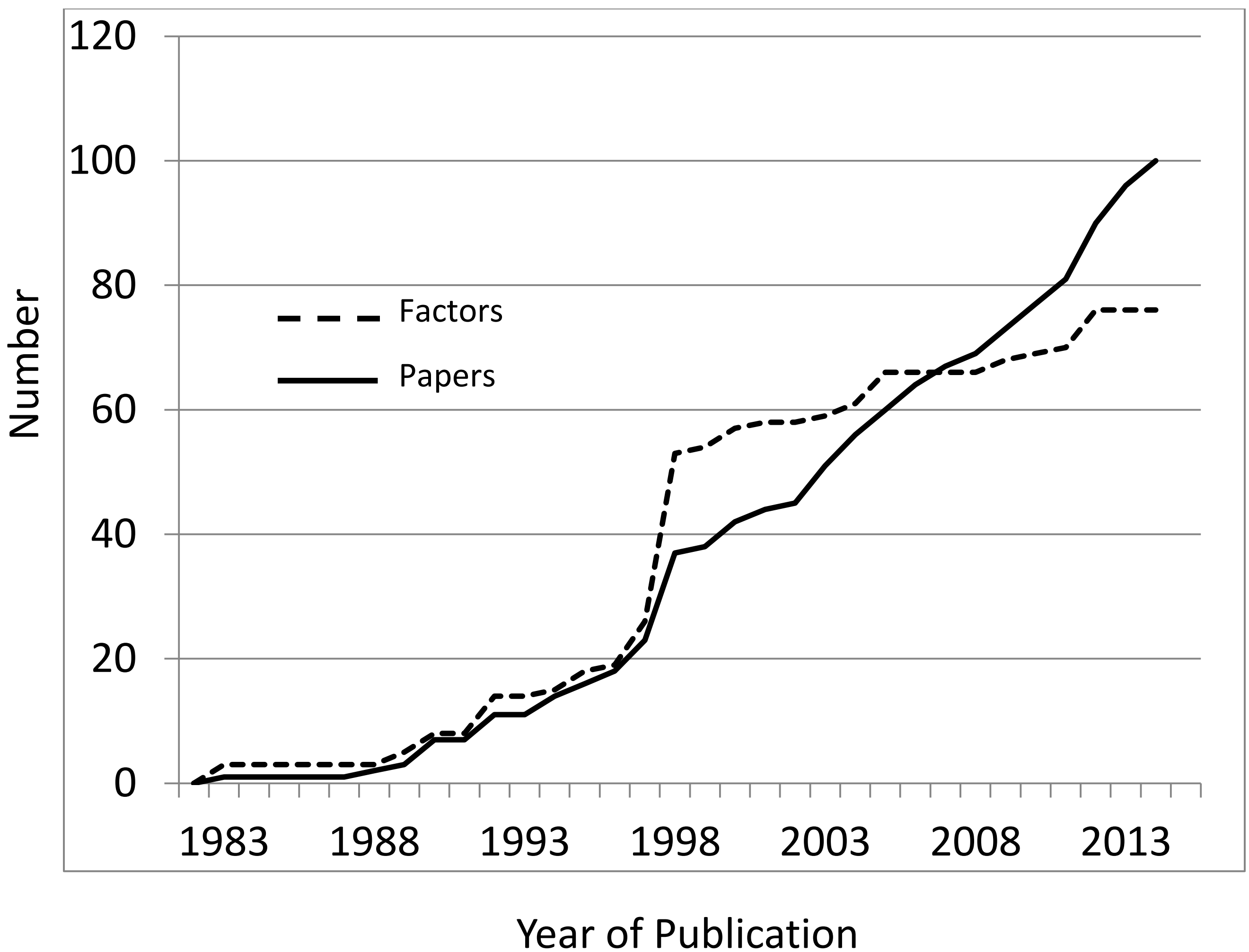


**Figure 3-3. The Classification Framework for factors associated with missing patient-reported outcome data.** In the column identified by #, factors that were based on regression analysis are denoted by Yes.

[**Figure 3-3** is provided on the next two pages]

| Component | Category | Sub-Category or Example | # | Author |
|---|---|---|---|---|
| Instrument | Availability in different languages | used non-validated translations | | Kramer 2000 |
| | Clarity | instrument has unrecognized ambiguities | | Fallowfield 1998 |
| | Inappropriate content | instrument is inappropriate to patients | | Fallowfield 1998 |
| | Unacceptable content | instrument is unacceptable to patients | | Fallowfield 1998 |
| | Length | the instrument was felt to be too long | | Groenvold 1999 |
| Participant | Sociodemographic | age | Yes | Marcus 1983 |
| | | total family income | Yes | Marcus 1983 |
| | | employment status | Yes | Marcus 1983 |
| | | education | Yes | Guadagnoli 1992 |
| | | race/ethnicity | | Psaty 1994 |
| | | culture | | Bernhard 1998 |
| | | gender | | Kaasa 1998 |
| | | living alone | Yes | Zunzunegui 2001 |
| | | place of birth | Yes | Garcia 2005 |
| | | did not live in vicinity to hospital | | Williams 2011 |
| | | no previous research participation | Yes | Lamers 2012 |
| | | retired | Yes | Mein 2012 |
| | | married | Yes | Mein 2012 |
| | | homeowner | Yes | Mein 2012 |
| | | no insurance | Yes | Zhivan 2012 |
| | Diagnosis | chronic disease | Yes | Guadagnoli 1992 |
| | Stage | four or more positive nodes at diagnosis | Yes | Hahn 1998 |
| | Comorbidity | concurrent chronic disease | Yes | Hahn 1998 |
| | Intellectual capacity | incapable of completing an instrument | | Aaronson 1990 |
| | Motivation | require encouragement or effort to complete the instrument | | Bernhard 1998 |
| | Caregiver | partner took instrument home but patient did not complete | | Moinpour 2000 |
| | Emotional status | patient was upset or overwhelmed | | Yancik 1989 |
| | Emotional response | patient feels inquiry is a violation of privacy | | Yancik 1989 |
| | Error in completing the instrument | patient provides unclear answers | | Hurny 1992 |
| | Error in returning the instrument | patient loses the instrument | | Huntington 2005 |
| | Refusal to complete the instrument | refuse to complete the instrument unless it was read to them | | Ganz 1998 |
| | Cognitive status | exacerbation of psychiatric conditions | | Sherman 2005 |
| | Physical impairment | worsening of performance status | | Ganz 1998 |
| | Progressing disease | compliance decreases during follow-up primarily because of disease progression | | Hopwood 1996 |
| | Toxicity | Grade: 3+ | | Bernhard 1998 |
| | | Type: diarrhea | | Bernhard 1998 |
| | Level of well-being | High: patients experiencing fewer problems may not be diligent in returning instruments | | Fayers 1997 |
| | | Low: baseline FACT-G score | Yes | Hahn 1998 |

| | | | | |
|---|---|---|---|---|
| Centre | Level of activity | High: many patients entered on a trial in a short period of time | | Sadura 1992 |
| | | Low: few patients entered on a trial | Yes | Osoba 1998 |
| | Research infrastructure | lack of staff | | Hopwood 1997 |
| | Location | linguistic region | Yes | Courvoisier 2012 |
| | Refusal to collect PRO data | participating centre refuses to collect PRO data | | Hurny 1992 |
| Staff | Conceptual knowledge | misunderstandings about the value of PRO | | Hahn 1998 |
| | Procedural knowledge | staff unfamiliar with procedures | | Hopwood 1997 |
| | Commitment | unwillingness of staff to collect PRO data | | Aaronson 1990 |
| | Perceived patient distress | patient perceived as too distressed to be approached | | Hopwood 1997 |
| | Error in administering the instrument | failure to approach the patient | | Hurny 1992 |
| | | patient told instrument is not needed | | Moinpour 2000 |
| | | the wrong instrument is administered to the patient | | Bernhard 1998 |
| | | the instrument is administered to the patient without instructions | | Taphoorn 2010 |
| | | failure to collect the instrument | | Hopwood 1997 |
| | Error in processing the instrument | failure to check the instrument for completeness | | Hopwood 1997 |
| | | staff may lose or misplace completed instruments | | Bernhard 1998 |
| | | incorrect data entry may lead to missing data | | Bernhard 1998 |
| | Error in sending the instrument | staff may not send completed instruments to the central office | | Bernhard 1998 |
| | Continuity in workload | frequent changes in the nursing staff | | Bernhard 1995 |
| | General workload | time constraints of medical or nursing staff in the typical clinical setting | | Aaronson 1990 |
| | Trial-specific workload | priority given to other matters | | Hopwood 1997 |
| | | conflicts with other studies | | Bernhard 1995 |
| | Personal skills | communication at time of initial recruitment | | Moynihan 1998 |
| | | continued engagement at follow-up | | Moynihan 1998 |
| Study | Frequency of assessment | frequent assessment over a long time interval | | Bernhard 1998 |
| | Timepoint of assessment | assessment of PRO after treatment failure | | Bernhard 1998 |
| | Data capture format | Electronic: technical difficulties | | Aaron 2004 |
| | | Electronic: electronic format versus paper format of instrument | | Gaertner 2004 |
| | | Paper: errors in hardcopy of instrument due to photocopy machine | | Huntington 2005 |
| | | Paper: self-report format versus interview format of instrument | | De Sousa 2009 |
| | | Paper: allowing patient to complete and return instrument by mail, rather than on-site | | Fallowfield 1998 |
| | | Verbal: no phone access for follow-up | | Sherman 2005 |
| | Research publicity | negative publicity about clinical trial research | | Ganz 1998 |
| | PRO outcome status | PRO specified as a secondary outcome | | Pinchon 2009 |
| | Location of treatment | treatment by an oncologist in a private practice | | Bernhard 1995 |
| | Intention of treatment | receipt of palliative rather than curative treatment | | Kopp 2003 |
| | Type of treatment | radical surgery | | Mosconi 1998 |
| | Allowable time window for assessment | instrument completion outside of allowed interval | | Ganz 1998 |

3.7 References


1. Mayo NE. *Dictionary of quality of life and health outcomes measurement, first edition.* International Society for Quality of Life Research: Lexington, 2015.

2. Food and Drug Administration. *Guidance for industry. Patient-reported outcome measures: use in medical product development to support labeling claims.* 2009.

3. Zubrod CG, Schneiderman M, Frei E, et al. Appraisal of methods for the study of chemotherapy of cancer in man: comparative therapeutic trial of nitrogen mustard and triethylene thiophosphoramide. *J Chron Dis* 1960;11:7-33.

4. Fayers PM and Hand DJ. Factor analysis, causal indicators and quality of life. *Qual Life Res* 1997;6:139-150.

5. Ware J, Snow KK, Kosinksi M, et al. *SF-36 health survey manual and interpretation guide*. The Health Institute: Boston, 1993.

6. Aaronson NK, Ahmedzai S, Bergman B, et al. The European Organisation for Research and Treatment of Cancer QLQ-C30: A quality-of-life instrument for use in international clinical trials in oncology. *J Ntl Cancer Inst* 1993;85:365-376.

7. Fayers PM and Machin D. *Quality of life: the assessment, analysis and interpretation of patient-reported outcomes, second edition*. John Wiley & Sons Ltd: West Sussex, 2007.

8. Bottomley A and Aaronson NK. International perspective on health-related quality-of-life research in cancer clinical trials: the European Organisation for Research and Treatment of Cancer experience. *J Clin Oncol* 2007;25:5082-5086.

9. Au H-J, Ringash J, Brundage M, et al. Added value of health-related quality of life measurement in cancer clinical trials: the experience of the NCIC CTG. *Expert Rev Pharmacoecon Outcomes Res* 2010;10:119-128.

10. Curran D, Bacchi M, Hsu Schmitz SF, et al. Identifying the types of missingness in quality of life data from clinical trials. *Stat Med* 1998;17:739-756.

11. Rubin DB. Inference and missing data. *Biometrika* 1976;63:581-592.

12. CIOMS. International ethical guidelines for biomedical research involving human subjects. Council for International Organizations of Medical Sciences. Geneva, 2002.

13. White E, Armstrong BK and Saracci R. *Principles of exposure measurement in epidemiology. Collecting, evaluating, and improving measures of disease risk factors. Second edition*. Oxford University Press, 2008.

14. Little RL, D'Agostino R, Cohen ML, et al. The prevention and treatment of missing data in clinical trials. *N Engl J Med* 2012; 367:1355-1361.

15. Little RJ, Cohen ML, Dickersin K, et al. The design and conduct of clinical trials to limit missing data. *Stat Med* 2012;31:3433-3443.

16. Fairclough DL. *Design and analysis of quality of life studies in clinical trials, second edition*. Chapman & Hall / CRC, 2010.

17. Sloan JA, Dueck AC, Erickson PA, et al. Analysis and interpretation of results based on patient-reported outcomes. *Value Health* 2007;10(Supplement 2):S106-S115.

18. McIsaac MA, Cook RJ and Poulin-Costello M. Incomplete data in randomized dermatology trials: consequences and statistical methodology. *Dermatology* 2013;226:19-27.

19. Newman DA. Missing data: five practical guidelines. *Organ Res Meth* 2014;17:372-411.

20. Dueck AC, Frost MH, Halyard MY, et al. Handling missing data. *Curr Probl Cancer* 2005;November/December:317-325.

21. Sedgwick P. Controlled trials: allocation concealment, random allocation, and blinding. *BMJ* 2015;350:h2633.

22. Ross L, Thomsen BL, Boesen EH, et al. In a randomized controlled trial, missing data led to biased results regarding anxiety. *J Clin Epidemiol* 2004;57:1131-1137.

23. Sterne JAC, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009;339(18 July):157-160.

24. Stang A. Nonresponse research – an underdeveloped field in epidemiology. *Eur J Epidemiol* 2003;18:929-931.

25. Neymark N, Kiebert W, Torfs K, et al. Methodological and statistical issues in quality of life and economic evaluation in cancer clinical trials: report of a workshop. *Eur J Cancer* 1998;34:1317-1333.

26. Rock EP, Scott JA, Kennedy DL, et al. Challenges to use of health-related quality of life for food and drug administration approval of anticancer products. *J Natl Cancer Inst Monogr* 2007;37:27-30.

27. Brundage M, Osoba D, Bezjak A, et al. Lessons learned in the assessment of health-related quality of life: selected examples from the National Cancer Institute of Canada Clinical Trials Group. *J Clin Oncol* 2007;25:5078-5081.

28. Rees JR Whale K, Fish D, et al. Patient-reported outcomes in randomised controlled trials of colorectal cancer: an analysis determining the availability of robust data to inform clinical decision-making. *J Cancer Res Clin Oncol* 2015;141:2181-2192.

29. Rombach I, Rivero-Arias O, Gray AM, et al. The current practice of handling and reporting missing outcome data in eight widely used PROMs in RCT publications: a review of the current literature. *Qual Life Res* 2016;25:1613-1623.

30. Bernhard J, Cella D, Coates AS, et al. Missing quality of life data in cancer clinical trials: serious problems and challenges. *Stat Med* 1998;17:517-532.

31. Hahn HA, Webster KA, Cella D, et al. Missing data in quality of life research in Eastern Cooperative Oncology Group (ECOG) clinical trials: problems and solutions. *Stat Med* 1998;17:547-559.

32. Osoba D, and Zee B. Completion rates in health-related quality-of-life assessment: approach of the National Cancer Institute of Canada Clinical Trials Group. *Stat Med* 1998;17:603-612.

33. Hussain JA, White IR, Langan D, et al. Missing data in randomized controlled trials testing palliative interventions pose a significant risk of bias and loss of power: a systematic review and meta-analyses. *J Clin Epidemiol* 2016;74:57-65.

34. Kiebert GM, Curran D and Aaronson NK. Quality of life as an endpoint in EORTC clinical trials. *Stat Med* 1998;17:561-569.

35. Kaasa S, Hjermstad MJ, Jordhoy MS, et al. Compliance in quality of life data: a Norwegian experience. *Stat Med* 1998;17:623-632.

36. Mosconi P, Torri V, Cifani S, et al. The multi-centre assessment of quality of life: the Interdisciplinary Group for Cancer Care Evaluation (GIVIO) experience in Italy. *Stat Med* 1998;17:577-585.

37. Aaronson NK. Quality of life research in cancer clinical trials: a need for common rules and language. *Oncology* 1990;4:59-66.

38. Rogers DP. The philosophy of taxonomy. *Mycologia* 1958;50:326-332.

39. Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009;6(6):e1000097.

40. King MT, Freudensprung U, Stockler M, et al. Analyzing patient-reported outcome endpoints in advanced cancer clinical trials: responder analysis versus mixed model for repeated measures illustrated with results from the AURELIA ovarian cancer trial. *Qual Life Res* 2013;22:15.

41. Moinpour CM, Donaldson GW, Liepa AM, et al. Evaluating health-related quality-of-life therapeutic effectiveness in a clinical trial with extensive nonignorable missing data and heterogeneous response: results from a phase III randomized trial of gemcitabine plus paclitaxel versus paclitaxel monotherapy in patients with metastatic breast cancer. *Qual Life Res* 2012;21:765-775.

42. Peyre H, Leplege A and Coste J. Missing data methods for dealing with missing items in quality of life questionnaires. A comparison by simulation of personal mean score, full information maximum likelihood, multiple imputation, and hot deck techniques applied to the SF-36 in the French 2003 decennial health survey. *Qual Life Res* 2011;20:287-300.

43. Bottomley A, Pe M, Sloan J, et al. Analysing data from patient-reported outcome and quality of life endpoints for cancer clinical trials: a start in setting international standards. *Lancet Oncol* 2016;17:e510-514.

44. Vickers AJ and Altman DG. Statistics notes: missing outcomes in randomised trials. *BMJ* 2013;346:f3438 doi:10.1136/bmj.f3438.

45. Moynihan C. Patient 'non-compliance' and 'missing data' in quality of life research: where does the problem lie? *Eur J Cancer* 1998;34:9-11.

46. National Research Council. *The Prevention and Treatment of Missing Data in Clinical Trials.* Panel on Handling Missing Data in Clinical Trials. Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press, 2010.

47. O'Neill RT and Temple R. The prevention and treatment of missing data in clinical trials: an FDA perspective on the importance of dealing with it. *Clin Pharmacol Ther* 2012;91:550-554.

48. Li T, Hutfless S, Scharfstein DO, et al. Standards should be applied in the prevention and handling of missing data for patient-centred outcomes research: a systematic review and expert consensus. *J Clin Epidemiol* 2014;67:15-32.

49. Mercieca-Bebber R, Palmer MJ, Brundage M, et al. 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**:e010938.

50. Chan A-W, Tetzlaff JM, Gotzsche PC, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. *BMJ* 2013;346:e7586.

51. Calvert M, Kyte D, Duffy H, et al. Patient-reported outcomes (PRO) assessment in clinical trials: a systematic review of guidance for trial protocol writers. *PLOS One* 2014;9:e110216.

52. Calvert M, Blazeby J, Altman DG, et al. Reporting of patient-reported outcomes in randomized trials: the CONSORT PRO extension. *JAMA* 2013;309:814-822.

53. Altman DG. Missing outcomes in randomized trials: addressing the dilemma. *Open Medicine* 2009;3:51-53.

54. Weed DL. Interpreting epidemiological evidence: how meta-analysis and causal inference methods are related. *Int J Epidemiol* 2000;29:387-390.

55. Sherman DW, McSherry CB, Parkas V, et al. Recruitment and retention in a longitudinal palliative care study. *Appl Nurs Res* 2005;18:167-177.

56. Graham ID, Logan J, Harrison MB, et al. Lost in knowledge translation: time for a map? *J Contin Educ Health Prof* 2006;26:13-24.

# Chapter 4

## Factors predicting missing instruments in three cancer randomized clinical trials

## 4.1 Abstract

**Purpose:** Missing patient-reported outcome (PRO) data can seriously threaten the validity of randomized clinical trials (RCTs). Identifying which factors predict missing instruments may help researchers develop strategies to prevent it from happening. This study examined the association of factors with time to the first missing instrument after randomization in three cooperative group RCTs.

**Methods:** We performed descriptive analyses and Cox proportional hazards regressions for three RCTs selected from the Canadian Cancer Trials Group: MA17 (breast cancer), PR7 (prostate cancer), and LY12 (non-Hodgkin's lymphoma). The outcome was the time from randomization to the first missing instrument. Variables for 15 factors were used as covariates based on availability and previously-reported putative associations with missing PRO data.

**Results:** Nine percent of 1352 subjects on MA17, 37% of 923 subjects on PR7, and 59% of 477 subjects on LY12 had a missing instrument. Twenty-five percent of subjects on MA17 had first missing instrument within 4.6 years. The median time to first missing instrument was: not observed for MA17, 7.3 years for PR7, 0.12 years for LY12. Cox regression revealed statistically significant independent associations with outcome for only five factors: baseline age (PR7) and level of well-being (LY12), and centre level of activity (LY12), presence of post-graduate residency training program (MA17, PR7), and centre geographic location (PR7, LY12).

**Conclusion:** Many factors reported to have association with missing instruments do not seem to predict time to the first missing instrument after randomization in RCTs. Context is important in understanding the few that may.

4.2 Introduction

Missing data in comparative studies for patient-reported outcome (PRO) endpoints are problematic. Although several factors are believed to be related to missing PRO data [13], few studies have provided quantitative evidence of associations of these factors with missing PRO data. Identifying which factors predict it may help researchers develop strategies for prevention.

A phase III randomized clinical trial (RCT) [1-3] can include PRO-specific hypotheses, and if well designed and properly conducted, appropriately analysed, and transparently reported, can contribute 1B level evidence to the evaluation of a new treatment compared with an alternative treatment [4]. Unfortunately, data for PROs are often missing and this common occurrence can have a deleterious effect. For example, the missing PRO data may preclude the planned analysis and reporting of the primary objective [5].

Two decades ago, representatives of cooperative oncology groups attending a workshop discussed missing PRO data in cancer clinical trials and analytical strategies. A special issue of *Statistics in Medicine* (volume 17, 1998) was devoted to their reports. Missing PRO data was identified as an important issue and over the years has been a serious practical problem [6]. One decade ago, the Food and Drug Administration established an expert panel to report recommendations for reducing and analysing missing data in clinical trials. The panel recommended reducing the frequency of missing during data collection rather than using statistical techniques to account for the missing data, since the choice of statistical method requires unverifiable assumptions concerning the mechanism causing the missing data [7].

Despite these recommendations, missing PRO data continue to threaten the validity of RCTs and potentially undermine the scientific credibility of causal conclusions [8, 9]. Recognizing the negative consequences of sub-optimal design and conduct, analysis, and reporting of RCTs, evidence-based guidelines were developed for including PROs in protocols [10], analyzing data [11] and reporting results [12]. Each of these documents provide guidance for dealing with missing PRO data but provide limited empirical evidence on factors associating with missing PRO data.

Previously, we performed a systematic review of the literature and created a Classification Framework for factors reported to have association with missing PRO data [13]. This inventory consists of five major components – Instrument, Participant, Centre, Staff, and Study – and 46 categories of factors. The extent to which each factor predicts missing instruments is largely unknown, in part because few were derived from regression models. Most factors were based on literature providing a general description of missing PRO data or offering expert opinion regarding factors believed to have association with missing PRO data.

We examined the evidence for an association between the age of subjects and missing instruments based on studies identified in the systematic review [13]. Three articles presented a statistical analysis including age and missing instruments within oncology RCTs [14-16]. We found variation in context, conceptualization of age, identification of missing instruments and analytic approaches. Consequently, the role of age in predicting missing instruments was difficult to determine, because evidence was inconsistent and contradictory [17].

Further research confirming the direction and strength of association for baseline age and other factors with missing instruments during treatment and follow-up would provide evidence for identifying which factors predict missing instruments and may help researchers develop strategies to prevent it from happening [18]. This would benefit stakeholders who have interest in minimizing missing PRO data. For example, the Canadian Cancer Trials Group (CCTG) conducts multi-centre phase III RCTs in cancer therapy, supportive care and prevention and has history collecting PRO data [19]. The prevalence of missing instruments at last follow-up was found to be more than 10% for nine of 10 RCTs examined in a convenience sample, and was lower in the curative setting [20-22] with range 7% at one month to 45% at 36 months, than in the palliative setting [23-27] with range 18% at six weeks to 42.4% at 50 weeks.

The objective of this study was to determine which factors had significant independent association with time to the first missing instrument after randomization in the setting of cooperative oncology group RCTs. Our approach was to perform descriptive and regression analyses of individual participant data from subjects in three RCTs that were led by the CCTG.

## 4.3 Methods

### 4.3.1 Selection of RCTs

Phase III RCTs from the CCTG were identified according to *a priori* eligibility criteria (**Figure 4-1**). A subset of 44 trials was identified having PROs and activation date after January 1, 1998. Subsequently, several trials were excluded because they were phase II, had low accrual, involved prevention, were not led by CCTG, were not open to accrual, had not been analyzed, or had activation date after January 1, 2013 which was when participating centres began to enter PRO data from paper. Afterwards, 23 phase III RCTs were eligible for selection.

Subjects receiving palliative treatment differ considerably from subjects receiving curative treatment, including severity of their disease, length of follow-up, and number of assessments. As such, factors associating with missing data may be quite different. To account for clinical heterogeneity [28] the 23 RCTs were stratified according to treatment intent (**Figure 4-1**) as palliative (n = 13, including treatment for an incurable metastatic cancer) or curative (n = 10, including adjuvant treatment for a potentially curative cancer). We elected to focus on RCTs having curative intention of treatment, and selected three providing variety in context and large sample size: from the breast disease site (MA17 [29]), the genito-urinary disease site (PR7 [30]), and the hematologic disease site (LY12 [31]).

4.3.2 Selection of variables

The content of the Classification Framework was compared with the list of variables in the CCTG database for each of the three RCTs. Variables for the categories of Staff, and variables for some categories of the other components (Instrument, Participant, Centre, Study) were not available. Accordingly, the approach to selecting covariates for analyses involved three steps. 1) Available variables were chosen for Participant (several sociodemographic characteristics, stage, caregiver, physical impairment) and for Study (treatment). 2) PRO data provided by subjects before randomization were selected for cognitive status, for toxicity, and for level of well-being. 3) Three new covariates were created for Centre: the level of activity described the number of subjects entered on study, the presence of a post-graduate residency training program in surgical oncology, gynecologic oncology, medical oncology, radiation oncology or hematology distinguished university-affiliated (academic) centres from community (non-academic) centres, and location identified emplacement in one of four national geographic regions (the Atlantic provinces, Quebec, Ontario, or the western provinces of Canada). **Table 4-1** provides the covariates that were included in the analyses.

The CCTG provided individual participant data for subjects from Canada on the three RCTs, which facilitated initial data cleaning and assessment of the nature, shape and quality of the data. Baseline scores for scales within the Short Form-36 Health Survey [32], Quality of Life Questionnaire Core-30 [33], and Functional Assessment of Cancer Therapy-General [34] were calculated according to published criteria.

4.3.3 Definition of outcome

The outcome was the time to first missing instrument after randomization, defined for each eligible subject with a non-missing instrument at the time of randomization as the time from randomization to the earliest time that the instrument was not completed when expected after the date of randomization. A dichotomous variable of missing instrument was also created to measure whether an eligible subject had one or more instruments missing. Allowable time windows for assessments were established *a priori.*

4.3.4 Approach for descriptive analysis

Each covariate was represented in a continuous or categorical form, and the association with missing instrument was evaluated. The t-test compared the mean of two groups when continuous representation satisfied the normality assumption. The Wilcoxon rank-sum test compared the two groups when the distribution was skewed. The Chi-squared test compared the observed frequency with the expected frequency in each cell of a contingency table. Fisher's exact test was used when a cell count was less than five. The level of significance for each test was $p < 0.05$.

4.3.5 Approach for regression analysis

Cox proportional hazards regression was the statistical model for time to first missing instrument after randomization. Follow-up time was censored at the date of death, at the date of recurrence, at date off protocol, and at the date of last attendance if a missing instrument was not previously observed. Ties were handled with the exact method. Cox regression identified covariates having significant independent association with the outcome [35, 36]. First, within a given component, univariate models examined the association between each covariate and the outcome. Then, parsimonious multivariate models were built with covariates having $p < 0.20$ for each of the components, using a backward selection approach and cut-point of $p = 0.20$. Finally, the covariates from the component models were included together in a model, and a backward selection approach with the same cut-point was used to build a final multivariate model. Hazard ratios were interpreted as relative risks for missing instrument [37]. An increase in relative risk above one represents shorter time to first missing instrument, while a decrease in relative risk below one represents longer time to first missing instrument. Separate analyses were performed for each of the three RCTs with Statistical Analysis System version 9.4 (Cary, NC).

4.4 Results

4.4.1 General description of the three RCTs (**Table 4-2**)

The 1352 subjects on MA17 were females with post-menopausal breast cancer, allocated to an aromatase inhibitor or placebo. The 923 subjects on PR7 were males with prostate cancer, allocated to intermittent or continuous androgen suppression combination therapy. The subjects on LY12 were a mix of 189 females and 288 males with non-Hodgkin’s lymphoma, and with younger median age, allocated to different chemotherapy, followed by autologous stem cell transplant. Approximately half of the centres on MA17, two-thirds of the centres on PR7, and three-quarters of the centres on LY12 had a post-graduate residency training program. The majority of centres were in the province of Ontario. The median level of centre activity was lowest for centres on LY12.

Nine percent of subjects on MA17, 37% of subjects on PR7, and 59% of subjects on LY12 had a missing instrument. Twenty-five percent of subjects on MA17 had first missing instrument within 4.6 years. The median time to the first missing instrument was not observed for MA17, was 7.3 years for PR7, and was 0.12 years for LY12. **Appendix E.2** provides graphs for Kaplan-Meier estimates of survival function.

4.4.2 <u>MA17</u> (**Table 4-3**)

In descriptive analysis, significant differences were observed for Centre level of activity, presence of a post-graduate residency training program, and location. There was a trend towards a lower percentage of subjects having a missing instrument at centres with an increased level of activity. A higher percentage of subjects with a missing instrument were from centres with a post-graduate residency training program. The lowest percentage of subjects with a missing instrument was from centres in western Canada. No significant differences were observed for baseline bodily pain, vitality or general health scale scores.

In univariate Cox regression, significant associations with time to the first missing instrument were observed for presence of a post-graduate residency training program and location, and these two covariates for Centre remained in the final multivariate model. Subjects at centres with a post-graduate residency training program, compared to subjects at centres without a post-graduate residency training program, had a 1.66-fold increase in the relative risk of missing instrument, controlling for centre location. Subjects at centres in Quebec, compared to subjects at centres in western Canada, had a 2.14-fold increase in the relative risk, controlling for presence of a post-graduate residency training program.

4.4.3 PR7 (**Table 4-4**)

In descriptive analysis, significant differences were observed for Participant age, and for Centre presence of a post-graduate residency training program and location. Subjects with a missing instrument were younger (95% confidence interval 72.0 - 73.4 years), in comparison to those with no missing instrument (95% confidence interval 73.7 - 74.7 years). No significant differences were observed for baseline cognitive function, pain, fatigue, nausea and vomiting, or global scale scores.

In univariate Cox regression, significant associations with time to the first missing instrument were observed for Participant age, and for Centre level of activity, presence of a post-graduate residency training program, and location. These four covariates remained in the final multivariate model. Subjects with a one-year increase in age had a three percent decrease in the relative risk. Subjects at centres with a one-unit increase in level of activity had a null increase in the relative risk. Subjects at centres with a post-graduate residency training program had a 1.81-fold increase in the relative risk. Subjects at centres in Quebec had a 26% decrease in the relative risk, but the 95% confidence interval bounded unity.

4.4.4 LY12 (**Table 4-5**)

In descriptive analysis, significant differences were observed for Centre level of activity and location. No significant differences were observed for baseline physical, social/family, emotional or functional well-being scale scores, total score, or trial outcome index.

Univariate Cox regression revealed significant associations of new covariates for Participant with time to the first missing instrument. These include stage, caregiver, performance status, and baseline scores for the PRO scales. Multivariate analysis revealed functional well-being had the strongest association with time to the first missing instrument when the four well-being scales were included in a model. To further develop a model for Participant, stage, caregiver and performance status were included in separate models with functional well-being, total score, or trial outcome index from Functional Assessment of Cancer Therapy-General. Results were similar for the different models. Caregiver and functional well-being remained significant after performance status and then stage were removed. Centre level of activity and location also remained significant when included together in a component model. Remaining in the final multivariate model were functional well-being, level of activity and location. Subjects with a one-unit increase in functional well-being had a three percent decrease in the relative risk. Subjects at centres with a one-unit increase in level of activity had a one percent decrease in the relative risk. In comparison to subjects from centres in western Canada, subjects from centres in eastern Canada (HR = 2.20), Ontario (HR = 1.61) or Quebec (HR = 1.33) had an increase in the relative risk, but the 95% confidence interval for Quebec bounded unity.

## 4.5 Discussion

For an RCT including the evaluation of PROs, the first evaluation is typically before randomization measuring health status at baseline, and subsequent assessments measure health status during treatment and follow-up. Ideally, the subject should complete an instrument at each longitudinal assessment. Time to the first missing instrument is an important consideration because long times imply that more PRO data are available for determining the health status during the course of the RCT.

We examined the association of 15 factors (11 within Participant, three within Centre, and one within Study) with time to the first missing instrument after randomization in three CCTG RCTs. MA17 included female subjects with breast cancer, PR7 included male subjects with prostate cancer, and LY12 included female and male subjects with non-Hodgkin’s lymphoma. Our results for the final multivariate Cox regression models revealed most factors did not have significant independent association with outcome, but a few did, depending on the RCT. Baseline age (PR7) and level of well-being (LY12) were found to have significant independent association with time to the first missing instrument, as were level of activity (LY12), presence of a post-graduate residency training program (MA17, PR7), and geographic location (PR7, LY12).

Our systematic review of the literature identified 46 categories of factors reported to have association with missing PRO data that we organized into five major components: Instrument, Participant, Centre, Staff, and Study [13]. The extent to which each factor predicts missing instruments is largely unknown, in part because only 15 factors within Participant and only two factors within Centre were derived from regression models. Our analysis of three RCTs provides additional evidence from regression for four factors (age, level of well-being, level of activity, location) having association with missing instruments, but not for three factors (employment status, insurance, stage). Our results do not provide evidence for seven factors having association with missing instruments (race/ethnicity, gender, caregiver, cognitive status, physical impairment, toxicity, or type of treatment).

Hopwood [38] reported that organizational challenges including lack of staff within centres may lead to missing data. Because the database did not identify whether centres had research infrastructure, we searched university websites and used the available information to create a new covariate for each centre, describing the presence or absence of a post-graduate residency training program in oncology or hematology, to distinguish university-affiliated (academic) centres from community (non-academic) centres. One may expect subjects from centres with a post-graduate residency training program – presumably with more research support – would have longer time to the first missing instrument than subjects from centres without a post-graduate residency training program, as manifested in a decreased relative risk. We found, however, that subjects on MA17 and subjects on PR7 from centres with a post-graduate residency training program had an increased relative risk, indicating shorter time to the first missing instrument. This result is surprising and the reason is not clear, perhaps due to additional activity – centres were likely involved with more RCTs than the three selected – which could not be assessed in the current study. Three-quarters of LY12 centres had a post-graduate residency training program and the unbalance across centres may explain the null association on LY12. Our results for MA17 and PR7 contribute new evidence that presence of a post-graduate residency program within centres may be an important factor leading to missing instruments on RCTs.

The results for geographic location are intriguing. Subjects on MA17 from centres in Quebec, and subjects on LY12 from centres in the Atlantic provinces or in Ontario, had increased relative risk in comparison to subjects from centres in the western provinces, while for subjects on PR7, the relative risks fluctuated with location, and confidence intervals bounded unity. The reason is not easily apparent. Interestingly, however, Courvoisier [39], in a study of electronic collection of patient-reported mood via cell phone, found the probability of answering a given call was significantly lower for subjects from the German-speaking region compared to the French-speaking region in Switzerland. Our results agree with Courvoisier [39] that location may contribute to missing PRO data.

The level of activity described the number of subjects entered on study at the centres, and as such was a measure of accrual. Entry of many subjects on a trial in a short period of time [20], and entry of few subjects overall on a trial [16], have been reported in association with missing PRO data. In our study, a one-unit increase in level of activity was associated with a one percent decrease in the relative risk on LY12.

The literature suggests older individuals may have a higher chance of missing instruments. In our systematic review, 23 articles examined the relationship between age and missing PRO data in different health conditions [17]. Most of the 11 articles providing expert opinion or a general description reported an association between older age and an *increase* in missing PRO data, four of the 12 articles providing a statistical analysis also reported an *increase* (one RCT [16] and three cohort studies), but the other eight articles reported an association between older age and a *decrease* in missing PRO data (two RCTs [14, 15] and 6 cohort studies). In our study, for men on

PR7, the relative risk decreased by three percent for every one-year increase in age. This result refutes the stereotype that older men are not compliant with instrument assessment [40].

Two factors reported to have association with missing PRO data are baseline level of well-being [16] as measured with the Functional Assessment of Cancer Therapy-General [34], and worsening of performance status [41] as an example of physical impairment. In our study, on LY12, although worsening of performance status was significantly associated with increased relative risk in univariate Cox regression, performance status did not remain in the final multivariate model. A one-unit increase in functional well-being was significantly associated with a three percent decrease in the relative risk, and remained the sole factor for Participant, in the final multivariate model. Interestingly, significant associations with time to the first missing instrument were not observed for baseline scale scores of MA17 or PR7 subjects, or for their performance status, likely because they were healthier with less variability than LY12 subjects.

The main strength of the study is its internal validity. (a) We identified RCTs according to *a priori* eligibility criteria. (b) To account for clinical heterogeneity, we stratified the RCTs and chose three having curative intention of treatment. (c) We used previously-reported putative associations to guide the selection of covariates. For each RCT, we developed a component-specific model for covariates within Participant and for covariates within Centre, then combined these models and developed a final trial-specific model. (d) The RCTs had long duration and multiple points of expected assessment. The measurement of outcome was novel. Our approach to analyses of missing instruments improves on regression at an arbitrary cut-point [16] or without careful consideration of time [15]. (e) We chose Cox regression for the statistical model because the shape of the underlying hazard was unknown, and a semi-parametric method is more robust than parametric methods, due to their vulnerability for misspecification of the baseline hazard. (f) Multivariate modeling efficiently estimated the measures of association while simultaneously controlling for all covariates. The statistical models were valid as the number of events per covariate surpassed ten [42, 43].

The study had limitations. (a) We were not able to examine factors within Instrument due to collinearity with RCTs, and variables for Staff were not available in the database. (b) We did not proceed with analyses pooling individual participant data. There were several reasons for this decision. The intention of treatment was curative but the type of treatment was different. Furthermore, careful inspection revealed the mix of subjects, the severity of their disease, the scales comprising the core questionnaire, and the schedules of assessment after randomization, were different on each RCT. Researchers having interest using regression models to identify factors which may predict missing instruments should be aware that separate analyses of individual participant data, as we have demonstrated, can contribute evidence and pooling data from multiple RCTs may not be appropriate. (c) The results may not generalize to RCTs of subjects with cancer receiving palliative treatment.

Identifying which factors predict missing instruments may help researchers develop strategies to prevent it from happening. Our results suggest missing instruments are a phenomenon whose occurrence is context-specific. Two baseline factors of Participant (age and level of well-being) and one factor of Centre (level of activity) had weak association, and two factors of Centre (presence of a post-graduate residency training program and location) had stronger association, with time to the first missing instrument, but are not convincing in predicting the outcome with certainty. Presently, it is not sensible to use these factors for developing specific strategies for prevention, given the instability of results across RCTs. General strategies for reducing missing PRO data through design and implementation [44] may help in a given context.

Further regression analyses to examine factors in other RCTs are unlikely to provide most benefit for an investment of resources, for reason that the availability of covariates limits a deductive approach, because the RCTs were not specifically designed to evaluate missing instruments. Inductive evidence through qualitative inquiry is essential. Clinical Research Associates (CRAs) are responsible for recruiting subjects [45], and for the collection, organization, and submission of their data [46]. Consequently, we are interviewing CRAs, from centres in different geographic locations, and from centres with and without a post-graduate residency training program, about their experience and we anticipate gaining further insights into important influences of missing PRO data.

### 4.6 Conclusion

Many factors reported to have association with missing instruments do not seem to predict time to the first missing instrument after randomization in RCTs. Context is important in understanding the few that may.

4.7 Acknowledgements

Melanie Calvert, Madeleine King and Joseph Pater offered constructive advice for the analytical strategy. Liting Zhu helped in providing the requested data. Preliminary work was presented at International Society for Quality of Life Research in Philadelphia on October 21, 2017. MJP was supported by Queen's University Terry Fox Research Institute Transdisciplinary Training Program in Cancer Research, Queen's Graduate Award, and Health Sciences Graduate Growth Funding. Partial support was provided by the Division of Cancer Care and Epidemiology in the Cancer Research Institute at Queen's University.

**Figure 4-1. Selection of Phase III Randomized Clinical Trials for Quantitative Analyses**

n = 391 Canadian Cancer Trials Group trials

Exclude if

- PRO is not an endpoint, n = 317
- activation date is before 1998-JAN-01, n = 30 (old trials, *Statistics in Medicine* special volume)

n = 44 CCTG trials
with PRO and activation date after 1998-JAN-01

Exclude if

- phase II, n = 6
- low accrual, n = 6
- prevention trial, n = 3
- CCTG is not the lead group, n = 2
- trial is open to accrual, n = 2
- trial has not been analyzed, n = 1
- activation date is after 2013-JAN-01, n = 1 (participating centres enter PRO data from paper)

n = 23 RCTs
closed phase III RCTs led by CCTG with PRO, accrual,
activation date 1998-JAN-01 – 2012-DEC-31

Eligible RCTs for quantitative analyses

| Disease Site | Palliative (n = 13) | Curative (n = 10) |
|---|---|---|
| Brain | | CE6 |
| Breast | MA19, 31 | MA17, 20, 21 |
| Gastro-Intestinal | CO17, 20, PA3 | |
| Genito-Urinary | | PR7, 11 |
| Gynecologic | OV12, 16 | |
| Head and Neck | | HN3, 6 |
| Hematologic | MY10 | LY12 |
| Lung | BR21, 26, 29 | |
| Symptom Control | SC20, 23 | SC19 |

Selected for quantitative analyses[1]

MA17

PR7

LY12

[1] From the list of randomized clinical trials with curative intention of treatment, three were selected: from the breast (MA17), genito-urinary (PR7) and hematologic (LY12) disease sites for quantitative analyses.

**Table 4-1. Covariates for factors from the Classification Framework that were included in the quantitative analyses.**

| Classification Framework | | | Covariate | Randomized Clinical Trial | | |
|---|---|---|---|---|---|---|
| Component | Category | Sub-Category | | MA17 | PR7 | LY12 |
| Participant | Sociodemographic | Age | Age | Yes | Yes | Yes |
| | | Employment status | Employment status | | | Yes |
| | | Race / ethnicity | Race / ethnicity | | Yes | Yes |
| | | Gender | Gender | | | Yes |
| | | Insurance | Insurance | | | Yes |
| | Stage | | Stage | Yes | | Yes |
| | Caregiver | | Caregiver | | | Yes |
| | Physical impairment | | Physical impairment | Yes | Yes | Yes |
| | Cognitive status | | Cognitive function [1] | | Yes | |
| | Toxicity | | Pain [1] | Yes | Yes | |
| | | | Fatigue [1] | Yes | Yes | |
| | | | Nausea and vomiting [1] | | Yes | |
| | Level of well-being | | General health [1] | Yes | | |
| | | | Global health / quality of life [1] | | Yes | |
| | | | Physical well-being [1] | | | Yes |
| | | | Social / family well-being [1] | | | Yes |
| | | | Emotional well-being [1] | | | Yes |
| | | | Functional well-being [1] | | | Yes |
| | | | Total FACT-G score [1] | | | Yes |
| | | | Trial outcome index [1] | | | Yes |
| Centre | Level of activity | | Level of activity – patients | Yes | Yes | Yes |
| | Research infrastructure | | Post-graduate residency training program | Yes | Yes | Yes |
| | Location | | Location | Yes | Yes | Yes |
| Study | Type of treatment | | Type of treatment | Yes | Yes | Yes |

[1] Scale score for patient-reported outcome, from baseline assessment

**Table 4-2. General description of the three randomized clinical trials that were included in the quantitative analyses.**

| MA17 | PR7 | LY12 |
|---|---|---|
| Study design | | |
| Women (median age 63.4 years) with post-menopausal breast cancer who completed five or more years of adjuvant tamoxifen were allocated to double-blind treatment with the aromatase inhibitor letrozole or placebo | Men (median age 74.4 years) with prostate cancer who completed radiotherapy to the prostate more than a year ago and had a rising prostate specific antigen without other evidence of metastasis were allocated to non-blinded intermittent or continuous androgen suppression combination therapy with a leuteinizing hormone-releasing hormone agonist and a nonsteroidal antiandrogen | Women or men (median age 55.0 years) with non-Hodgkin's lymphoma who had disease that was refractory to or relapsed after one prior first-line anthracycline-containing chemotherapy regimen were allocated to non-blinded salvage chemotherapy with GDP (Gemcitabine, Dexamethasone, and Cisplatin) or DHAP (Dexamethasone, Cytarabine, and Cisplatin), followed by autologous stem cell transplant |
| Number of subjects who were entered on the randomized clinical trial (start and end of recruitment) | | |
| 5187 (August 1998 to September 2002) | 1386 (January 1999 to November 2005) | 619 (August 2003 to November 2011) |
| Number of subjects from Canada who were included in this study | | |
| 1352 | 923 | 477 |
| Patient-Reported Outcome instrument | | |
| Short Form-36 Health Survey + Menopause-Specific Quality of Life Questionnaire | Quality of Life Questionnaire Core-30 + additional questions | Functional Assessment of Cancer Therapy-General + neurologic toxicity subscale + lymphoma subscale + additional questions |
| Number of questions in Patient-Reported Outcome instrument | | |
| 65 | 52 | 66 |
| Scales comprising core questionnaire | | |
| Physical Functioning, Role Physical, Role Emotional, Social Functioning, Bodily Pain, Vitality, Mental Health, General Health | Physical Function, Role Function, Dyspnea, Pain, Fatigue, Insomnia, Appetite Loss, Nausea and Vomiting, Constipation, Diarrhea, Cognitive Function, Emotional Function, Social Function, Financial Difficulties, Global Health and Quality of Life | Physical Well-Being, Social/Family Well-Being, Emotional Well-Being, Functional Well-Being |

| | | |
|---|---|---|
| Schedule of assessment with Patient-Reported Outcome instrument after randomization | | |
| at 6, 12, 24, 36, 48 and 60 months | at 4, 8, 12, 16, 20 and 24 months then every eight months | at the end of cycle 1, middle and end of cycle 2, and one-month after autologous stem cell transplant |
| Participating centres in Canada | | |
| Six in East, 12 in Quebec, 26 in Ontario, and 10 in West. Twenty-six (48.1%) had a post-graduate residency training program. Median level of activity was 16.5. | Four in East, 4 in Quebec, 12 in Ontario, and 10 in West. Nineteen (63.3%) had a post-graduate residency training program. Median level of activity was 22.5. | Five in East, 6 in Quebec, 9 in Ontario, and 6 in West. Nineteen (73.1%) had a post-graduate residency training program. Median level of activity was 9.0. |

**Table 4-3. Descriptive analysis and Cox proportional hazards regression of individual participant data from MA17.**

| Component | Covariate | Level | Descriptive analysis | | | | | Cox proportional hazards regression | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | One or more missing | | None missing | | p-value | Univariate | p-value | Multivariate | p-value |
| | | | n | % | n | % | | HR (95% CI) [1] | | HR (95% CI) [1] | |
| Participant | age [2a, 2b] | Mean (SD) | 64.3 (9.2) | | 63.8 (8.9) | | 0.503 | 1.01 (0.99, 1.03) | 0.482 | | |
| | | 30.0 – 59.9 | 45 | 9.1 | 449 | 90.9 | 0.567 | reference | 0.496 | | |
| | | 60.0 – 69.9 | 39 | 8.2 | 439 | 91.8 | | 0.88 (0.57, 1.34) | | | |
| | | 70.0 – 89.9 | 39 | 10.3 | 341 | 89.7 | | 1.14 (0.75, 1.76) | | | |
| | stage [3] | Negative | 59 | 9.4 | 569 | 90.6 | 0.684 | reference | 0.745 | | |
| | | Positive | 58 | 9.2 | 575 | 90.8 | | 1.02 (0.71, 1.47) | | | |
| | | Unknown | 6 | 6.6 | 85 | 93.4 | | 0.74 (0.32, 1.70) | | | |
| | physical impairment [4] | 0 | 108 | 9.2 | 1072 | 90.8 | 0.858 | reference | 0.974 | | |
| | | 1 | 14 | 8.6 | 149 | 91.4 | | 1.01 (0.58, 1.76) | | | |
| | | 2 | 1 | 11.1 | 8 | 88.9 | | 1.26 (0.18, 9.02) | | | |
| | pain [5] | Median (IR) | 84.0 (42.5) | | 84.0 (38.0) | | 0.811 | 1.00 (0.99, 1.01) | 0.993 | | |
| | fatigue [6] | Median (IR) | 70.0 (25.0) | | 70.0 (25.0) | | 0.267 | 1.00 (0.99, 1.01) | 0.419 | | |
| | general health [7] | Median (IR) | 82.0 (25.0) | | 82.0 (23.0) | | 0.499 | 1.00 (0.99, 1.01) | 0.886 | | |
| Centre | level of activity – patients [8a, 8b] | Median (IR) | 11.0 (17.0) | | 17.0 (29.0) | | **0.001** | 1.00 (0.99, 1.01) | 0.392 | | |
| | | ≤ 9 | 56 | 12.5 | 391 | 87.5 | **0.005** | reference | 0.643 | | |
| | | 10 – 19 | 30 | 9.5 | 285 | 90.5 | **0.0004** (trend) | 0.91 (0.58, 1.43) | | | |
| | | 20 – 39 | 22 | 7.2 | 284 | 92.8 | | 0.81 (0.49, 1.34) | | | |
| | | ≥ 40 | 15 | 5.3 | 269 | 94.7 | | 0.70 (0.39, 1.26) | | | |
| | post-graduate residency training program [9] | Yes | 98 | 10.5 | 839 | 89.5 | **0.009** | 1.82 (1.17, 2.82) | **0.008** | 1.66 (1.06, 2.60) | **0.027** |
| | | No | 25 | 6.0 | 390 | 94.0 | | reference | | reference | |
| | location [10] | East | 7 | 6.0 | 109 | 94.0 | **0.042** | 1.08 (0.44, 2.64) | **0.035** | 1.19 (0.48, 2.93) | **0.112** |
| | | Quebec | 27 | 12.4 | 190 | 87.6 | | 2.37 (1.26, 4.47) | | 2.14 (1.13, 4.04) | |
| | | Ontario | 74 | 9.8 | 684 | 90.2 | | 1.57 (0.90, 2.74) | | 1.49 (0.86, 2.61) | |
| | | West | 15 | 5.8 | 246 | 94.2 | | reference | | reference | |
| Study | type of treatment [11] | 1 | 65 | 9.3 | 632 | 90.7 | 0.765 | 1.07 (0.75, 1.52) | 0.721 | | |
| | | 2 | 58 | 8.8 | 597 | 91.2 | | reference | | | |

[1] Hazard Ratio (95% Confidence Interval).

Age at randomization in years: [2a] continuous, and categories based on [2b]decades.

[3] Nodal status.

[4] Performance status as measured by ECOG scale.

[5] Transformed scale score for Bodily Pain, calculated as ((BP_SCORE-2) / 10)*100 from questions 7, 8. N Miss = 3,16 : 1,0.

[6] Transformed scale score for Vitality, calculated as ((VT_SCORE-4) / 20)*100 from questions 9A,E, G, I. N Miss = 2,2 : 1,0.

[7] Transformed scale score for General Health, calculated as ((GH_SCORE-5) / 20)*100 from questions 1, 11A, B, C, D. N Miss = 3,6 : 1,0.

Number of patients entered on study: [8a] continuous, and [8b] categories.

[9] Post-graduate residency training program in surgical oncology, gynecologic oncology, medical oncology, radiation oncology or hematology. Each centre was assigned Yes or No. Yes included 26 centres: CAAJ, CABN, CAEJ, CAGH, CAGQ, CAHA, CAHC, CAHF, CAHL, CAHM, CAHN, CAKK, CAKO, CALC, CALM , CAMG, CAMM, CAMN, CAMP, CAMS, CAMX, CANL, CARM, CAVA, CAVF and CAVV. No included 28 centres: CACC, CAEM, CAER, CAGL, CAGS, CAGV, CAJV, CAKP, CALO, CALS, CALY, CAME, CAMH, CAMJ, CAMR, CAMT, CAMU, CAMV, CANW, CAPN, CAPS, CAPT, CASA, CASS, CATL, CAVC, CAVK and CAVN.

[10] Each centre was assigned to one of four categories according to geographic location in Canada. East included 6 centres in the Atlantic provinces, Quebec included 12 centres, Ontario included 26 centres and West included 10 centres in the western provinces.

[11] Allocation code assigned at randomization by central office. 1 = letrozole, 2 = placebo.

**Table 4-4. Descriptive analysis and Cox proportional hazards regression of individual participant data from PR7.**

<table>
<tr><td rowspan="3">Component</td><td rowspan="3">Covariate</td><td rowspan="3">Level</td><td colspan="5">Descriptive analysis</td><td colspan="4">Cox proportional hazards regression</td></tr>
<tr><td colspan="2">One or more missing</td><td colspan="2">None missing</td><td rowspan="2">p-value</td><td>Univariate</td><td rowspan="2">p-value</td><td>Multivariate</td><td rowspan="2">p-value</td></tr>
<tr><td>n</td><td>%</td><td>n</td><td>%</td><td>HR (95% CI) [1]</td><td>HR (95% CI) [1]</td></tr>
<tr><td rowspan="15">Participant</td><td rowspan="4">age [2a, 2b]</td><td>Mean (SD)</td><td colspan="2">72.7 (6.7)</td><td colspan="2">74.2 (6.1)</td><td>0.001</td><td>0.98 (0.96, 0.99)</td><td>0.004</td><td>0.97 (0.95, 0.99)</td><td>0.0004</td></tr>
<tr><td>50.0 – 69.9</td><td>107</td><td>43.0</td><td>142</td><td>57.0</td><td>0.043</td><td>reference</td><td rowspan="3">0.120</td><td></td><td rowspan="3"></td></tr>
<tr><td>70.0 – 79.9</td><td>187</td><td>34.8</td><td>350</td><td>65.2</td><td rowspan="2">0.018 (trend)</td><td>0.78 (0.62, 1.00)</td><td></td></tr>
<tr><td>80.0 – 89.9</td><td>44</td><td>32.1</td><td>93</td><td>67.9</td><td>0.79 (0.56, 1.12)</td><td></td></tr>
<tr><td rowspan="2">physical impairment [3]</td><td>0</td><td>283</td><td>37.8</td><td>465</td><td>62.2</td><td rowspan="2">0.113</td><td>reference</td><td rowspan="2">0.530</td><td></td><td rowspan="2"></td></tr>
<tr><td>1</td><td>55</td><td>31.4</td><td>120</td><td>68.6</td><td>0.91 (0.68, 1.22)</td><td></td></tr>
<tr><td rowspan="4">race</td><td>White</td><td>233</td><td>40.1</td><td>348</td><td>59.9</td><td rowspan="4">not calc-ulated</td><td>reference</td><td rowspan="4">not calc-ulated</td><td></td><td rowspan="4"></td></tr>
<tr><td>Other</td><td>4</td><td>33.3</td><td>8</td><td>66.7</td><td></td><td></td></tr>
<tr><td>Unknown</td><td>4</td><td>0.16</td><td>21</td><td>0.84</td><td></td><td></td></tr>
<tr><td>Missing</td><td>97</td><td>31.8</td><td>208</td><td>68.2</td><td></td><td></td></tr>
<tr><td>cognitive function [4]</td><td>Median (IR)</td><td colspan="2">83.3 (16.7)</td><td colspan="2">83.3 (16.7)</td><td>0.410</td><td>1.00 (0.99, 1.01)</td><td>0.923</td><td></td><td></td></tr>
<tr><td>pain [5]</td><td>Median (IR)</td><td colspan="2">0.0 (16.7)</td><td colspan="2">0.0 (16.7)</td><td>0.817</td><td>1.00 (1.00, 1.01)</td><td>0.626</td><td></td><td></td></tr>
<tr><td>fatigue [6]</td><td>Median (IR)</td><td colspan="2">11.1 (22.2)</td><td colspan="2">11.1 (22.2)</td><td>0.657</td><td>1.00 (0.99, 1.01)</td><td>0.854</td><td></td><td></td></tr>
<tr><td>nausea and vomiting [7]</td><td>Median (IR)</td><td colspan="2">0.0 (0.0)</td><td colspan="2">0.0 (0.0)</td><td>0.136</td><td>1.00 (0.98, 1.02)</td><td>0.722</td><td></td><td></td></tr>
<tr><td>global health [8]</td><td>Median (IR)</td><td colspan="2">83.3 (25.0)</td><td colspan="2">83.3 (25.0)</td><td>0.796</td><td>1.00 (0.99, 1.00)</td><td>0.824</td><td></td><td></td></tr>
<tr><td rowspan="10">Centre</td><td rowspan="5">level of activity – patients [9a, 9b]</td><td>Median (IR)</td><td colspan="2">24.0 (37.0)</td><td colspan="2">20.0 (34.0)</td><td>0.140</td><td>1.01 (1.00, 1.01)</td><td>0.005</td><td>1.00 (1.00, 1.01)</td><td>0.121</td></tr>
<tr><td>≤ 9</td><td>78</td><td>33.6</td><td>154</td><td>66.4</td><td>0.603</td><td>reference</td><td rowspan="4">0.156</td><td></td><td rowspan="4"></td></tr>
<tr><td>10 - 19</td><td>71</td><td>35.3</td><td>130</td><td>64.7</td><td rowspan="3">0.189 (trend)</td><td>1.18 (0.85, 1.62)</td><td></td></tr>
<tr><td>20 - 39</td><td>86</td><td>38.4</td><td>138</td><td>61.6</td><td>1.35 (1.00, 1.84)</td><td></td></tr>
<tr><td>≥ 40</td><td>103</td><td>38.7</td><td>163</td><td>61.3</td><td>1.36 (1.01, 1.83)</td><td></td></tr>
<tr><td rowspan="2">post-graduate residency training program [10]</td><td>Yes</td><td>291</td><td>39.6</td><td>444</td><td>60.4</td><td rowspan="2">0.0002</td><td>1.80 (1.32, 2.45)</td><td rowspan="2">0.0002</td><td>1.81 (1.31, 2.51)</td><td rowspan="2">0.0004</td></tr>
<tr><td>No</td><td>47</td><td>25.0</td><td>141</td><td>75.0</td><td>reference</td><td>reference</td></tr>
<tr><td rowspan="3">location [11]</td><td>East</td><td>27</td><td>47.4</td><td>30</td><td>52.6</td><td rowspan="3">0.006</td><td>1.68 (1.07, 2.63)</td><td rowspan="3">0.016</td><td>1.45 (0.92, 2.28)</td><td rowspan="3">0.012</td></tr>
<tr><td>Quebec</td><td>45</td><td>30.4</td><td>103</td><td>69.6</td><td>1.01 (0.69, 1.48)</td><td>0.74 (0.50, 1.10)</td></tr>
<tr><td>Ontario</td><td>202</td><td>40.2</td><td>301</td><td>59.8</td><td>1.41 (1.06, 1.87)</td><td>1.20 (0.89, 1.61)</td></tr>
</table>

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | West | 64 | 29.8 | 151 | 70.2 | | reference | | reference | |
| Study | type of treatment [12] | C | 158 | 34.4 | 301 | 65.6 | 0.168 | 1.03 (0.83, 1.27) | 0.800 | | |
| | | I | 180 | 38.8 | 284 | 61.2 | | reference | | | |

[1] Hazard Ratio (95% Confidence Interval).

Age at randomization in years: [2a] continuous, and categories created based on [2b] decades.

[3] Performance status as measured by ECOG scale.

[4] Transformed scale score for Cognitive Function, calculated as ((1–(COGNITIVE_RS–1)/3)*100) from questions 20, 25. N Miss = 0,1 : 1,0.

[5] Transformed scale score for Pain, calculated as (((PAIN_RS–1)/3)*100) from questions 9, 19. N Miss = 1,1 : 1,0.

[6] Transformed scale score for Fatigue, calculated as (((FATIGUE_RS–1)/3)*100) from questions 10, 12, 18. N Miss = 1,1 : CENSOR = 1,0.

[7] Transformed scale score for Nausea and Vomiting, calculated as (((NAUSVOM_RS–1)/3)*100) from questions 14, 15. N Miss = 1,1 : 1,0.

[8] Transformed scale score for Global Health / Quality of Life, calculated as (((GLOBAL_RS–1)/6)*100) from questions 29, 30. N Miss = 1,1 : 1,0.

Number of patients entered on study: [9a] continuous, and [9b] categories.

[10] Post-graduate residency training program in surgical oncology, gynecologic oncology, medical oncology, radiation oncology or hematology. Each centre was assigned Yes or No. Yes included 19 centres: CAAJ, CABN, CAEJ, CAGH, CAGQ, CAHC, CAHN, CAKK, CAKO, CALC, CALM , CAMN, CAMP, CANL, CARM, CATC, CATW, CAVA and CAVF. No included 11 centres: CAER, CAMU, CANH, CANW, CAPN, CAPT, CASA, CASS, CAVD, CAVG and CAVK.

[11] Each centre was assigned to one of four categories according to geographic location in Canada. East included 4 centres in the Atlantic provinces, Quebec included 4 centres, Ontario included 12 centres and West included 10 centres in the western provinces.

[12] Allocation code assigned at randomization by central office. C = continuous hormonal therapy, I = intermittent hormonal therapy.

**Table 4-5. Descriptive analysis and Cox proportional hazards regression of individual participant data from LY12.**

| Component | Covariate | Level | Descriptive analysis | | | | | Cox proportional hazards regression | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | One or more missing | | None missing | | p-value | Univariate | p-value | Multivariate | p-value |
| | | | n | % | n | % | | HR (95% CI) [1] | | HR (95% CI) [1] | |
| Participant | age [2a, 2b] | Median (IR) | 55.0 (13.4) | | 55.1 (12.2) | | 0.851 | 1.00 (0.99, 1.01) | 0.707 | | |
| | | 20.0 – 49.9 | 85 | 60.7 | 55 | 39.3 | 0.860 | reference | 0.871 | | |
| | | 50.0 – 59.9 | 105 | 57.7 | 77 | 42.3 | | 0.93 (0.70, 1.24) | | | |
| | | 60.0 – 70.0 | 91 | 58.7 | 64 | 41.3 | | 0.94 (0.70, 1.26) | | | |
| | employment status [3] | Work for pay | 49 | 55.7 | 39 | 44.3 | 0.657 | reference | 0.578 | | |
| | | On leave | 83 | 53.9 | 71 | 46.1 | | 0.93 (0.66, 1.33) | | | |
| | | Unemployed | 46 | 61.3 | 29 | 38.7 | | 1.01 (0.68, 1.51) | | | |
| | | Retired | 43 | 60.6 | 28 | 39.4 | | 1.21 (0.81, 1.83) | | | |
| | gender | Female | 115 | 60.8 | 74 | 39.2 | 0.486 | reference | 0.624 | | |
| | | Male | 166 | 57.6 | 122 | 42.4 | | 0.94 (0.74, 1.20) | | | |
| | insurance | Yes | 217 | 56.8 | 165 | 43.2 | 0.738 | 0.86 (0.38, 1.94) | 0.722 | | |
| | | No | 6 | 66.7 | 3 | 33.3 | | reference | | | |
| | stage | I | 21 | 56.8 | 16 | 43.2 | 0.818 | reference | **0.157** | | |
| | | II | 61 | 55.4 | 49 | 44.6 | | 0.91 (0.55, 1.49) | | | |
| | | III | 77 | 61.1 | 49 | 38.9 | | 1.25 (0.77, 2.02) | | | |
| | | IV | 122 | 59.8 | 82 | 40.2 | | 1.26 (0.79, 1.99) | | | |
| | caregiver | No one | 52 | 50.0 | 52 | 50.0 | 0.102 | reference | **0.009** | | |
| | | Spouse | 131 | 59.0 | 91 | 41.0 | | 1.35 (0.98, 1.86) | | | |
| | | Someone else | 38 | 61.3 | 24 | 38.7 | | 1.40 (0.92, 2.13) | | | |
| | | Missing | 60 | 67.4 | 29 | 32.6 | | 1.91 (1.31, 2.76) | | | |
| | physical impairment [4] | 0 | 100 | 56.5 | 77 | 43.5 | 0.369 | reference | **0.013** | | |
| | | 1 | 140 | 59.1 | 97 | 40.9 | | 1.32 (1.02, 1.71) | | | |
| | | 2 | 29 | 60.4 | 19 | 39.6 | | 1.56 (1.03, 2.36) | | | |
| | | 3 | 12 | 80.0 | 3 | 20.0 | | 2.24 (1.23, 4.08) | | | |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | ethnicity | Hispanic/ Latino | 3 | 37.5 | 5 | 62.5 | 0.127 | reference | 0.406 | | |
| | | Non-Hispanic | 247 | 59.7 | 167 | 40.3 | | 1.46 (0.47, 4.56) | | | |
| | | Not reported | 4 | 33.3 | 8 | 66.7 | | 0.85 (0.19, 3.78) | | | |
| | | Unknown | 8 | 47.1 | 9 | 52.9 | | 0.95 (0.25, 3.58) | | | |
| | physical well-being [5] | Median (IR) | 21.0 (9.0) | | 22.0 (10.0) | | 0.496 | 0.98 (0.96, 1.00) | **0.036** | | |
| | soc/family well-being [6] | Median (IR) | 24.0 (5.8) | | 24.0 (5.8) | | 0.913 | 0.98 (0.95, 1.00) | **0.099** | | |
| | emotional well-being [7] | Mean (SD) | 15.3 (3.5) | | 15.7 (3.1) | | 0.195 | 0.97 (0.94, 1.01) | **0.140** | | |
| | functional well-being [8] | Mean (SD) | 15.5 (6.6) | | 16.4 (6.6) | | 0.144 | 0.97 (0.95, 0.99) | **0.001** | 0.97 (0.95, 0.99) | **0.003** |
| | total FACT-G score [9] | Mean (SD) | 73.7 (15.9) | | 75.5 (16.1) | | 0.239 | 0.99 (0.98, 1.00) | **0.004** | | |
| | trial outcome index [10] | Mean (SD) | 35.3 (11.8) | | 36.5 (12.3) | | 0.302 | 0.99 (0.98, 1.00) | **0.004** | | |
| Centre | level of activity – patients [11a, 11b] | Median (IR) | 14.0 (22.0) | | 18.5 (31.0) | | **0.032** | 1.00 (0.99, 1.00) | **0.007** | 0.99 (0.99, 1.00) | **0.010** |
| | | ≤ 9 | 104 | 63.8 | 59 | 36.2 | **0.054** (trend) | reference | **0.150** | | |
| | | 10 - 19 | 62 | 58.5 | 44 | 41.5 | | 0.85 (0.62, 1.17) | | | |
| | | 20 - 39 | 66 | 59.5 | 45 | 40.5 | | 0.91 (0.67, 1.24) | | | |
| | | ≥ 40 | 49 | 50.5 | 48 | 49.5 | | 0.67 (0.48, 0.95) | | | |
| | post-graduate residency training program [12] | Yes | 241 | 59.4 | 165 | 40.6 | 0.633 | 1.04 (0.74, 1.45) | 0.815 | | |
| | | No | 40 | 56.3 | 31 | 43.7 | | reference | | | |
| | location [13] | East | 51 | 77.3 | 15 | 22.7 | **0.001** | 2.33 (1.51, 3.58) | **0.001** | 2.20 (1.41, 3.44) | **0.005** |
| | | Quebec | 26 | 57.8 | 19 | 42.2 | | 1.34 (0.81, 2.23) | | 1.33 (0.80, 2.23) | |
| | | Ontario | 169 | 59.3 | 116 | 40.7 | | 1.42 (0.99, 2.05) | | 1.61 (1.10, 2.36) | |
| | | West | 35 | 43.2 | 46 | 56.8 | | reference | | reference | |
| Study | type of treatment [14] | 1 | 145 | 59.2 | 100 | 40.8 | 0.901 | 0.89 (0.70, 1.12) | 0.327 | | |
| | | 2 | 136 | 58.6 | 96 | 41.4 | | reference | | | |

[1] Hazard Ratio (95% Confidence Interval).

Age at randomization in years: [2a] continuous, and categories created based on [2b] combined decades.

[3] Work status category 'Unemployed' includes homemaker and other.

[4] Performance status as measured by ECOG scale.

[5] Scale score for Physical Well-Being, calculated as PWB = sum of responses for questions 1-7. N Miss = 15,5: 1,0.

[6] Scale score for Social/Family Well-Being, calculated as SWB = sum of responses for questions 8-14. N Miss = 15,5 : 1,0.

[7] Scale score for Emotional Well-Being, calculated as EWB = sum of responses for questions 15-20. N Miss = 15,5 : 1,0.

[8] Scale score for Functional Well-Being, calculated as FWB = sum of responses for questions 21-27. N Miss = 15,5 : 1,0.

[9] Total score for FACT-G, calculated as sum of PWB, SWB, EWB, and FWB. N Miss = 15,5 : 1,0.

[10] Score for Trial Outcome Index, calculated as sum of PWB and FWB. N Miss = 15,5 : 1,0.

Number of patients entered on study: [11a] continuous, and [11b] categories.

[12] Post-graduate residency training program in surgical oncology, gynecologic oncology, medical oncology, radiation oncology or hematology. Each centre was assigned Yes or No. Yes included 19 centres: CAAJ, CABN, CABP, CAGH, CAHA, CAHF, CAHN, CAHO, CAKK, CALC, CALM, CAMN, CAMP, CAMS, CANL, CARM, CATC, CATW and CAVA. No included 7 centres: CAEM, CAER, CAGS, CAMU, CAPT, CASA and CASS.

[13] Each centre was assigned to one of four categories according to geographic location in Canada. East included 5 centres in the Atlantic provinces, Quebec included 6 centres, Ontario included 9 centres and West included 6 centres in the western provinces.

[14] Allocation code assigned at randomization by central office. 1 = salvage treatment with R-GDP (gemcitabine, dexamethasone, cisplatin, rituximab), 2 = salvage treatment with R-DHAP (dexamethasone, cytarabine, cisplatin, rituximab). After two cycles, subjects on each arm proceeded to autologous stem cell transplant.

4.8 <u>References</u>


1. Priestman TJ, Baum M. Evaluation of quality of life in patients receiving treatment for advanced breast cancer. *Lancet* 1976;1:899-901.
2. Sedgwick P. Phases of clinical trials. *BMJ* 2011;343:d6068 doi: 10.1136/bmj.d6068.
3. Stockler MR et al. Patient-reported outcome results from the open-label phase III Aurelia trial evaluating bevacizumab-containing therapy for platinum-resistant ovarian cancer. *J Clin Oncol* 2014;32:1309-1316.
4. Szklo M, Nieto FJ. *Epidemiology beyond the basics, second edition.* Jones and Bartlett Publishers: Sudbury, 2007.
5. Moinpour CM et al. Challenges posed by non-random missing quality of life data in an advanced-stage colorectal cancer clinical trial. Psycho-Oncology 2000;9:340-354.
6. Lipscomb J, Reeve BB, Clauser SB, et al. Patient-reported outcomes assessment in cancer trials: taking stock, moving forward. *J Clin Oncol* 2007;25:5133-5140.
7. National Research Council (2010). *The Prevention and Treatment of Missing Data in Clinical Trials.* Panel on Handling Missing Data in Clinical Trials. Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.
8. Little RL, D'Agostino R, Cohen ML et al. The prevention and treatment of missing data in clinical trials. *N Engl J Med* 2012; 367:1355-1361.
9. Little RJ, Cohen ML, Dickersin K et al. The design and conduct of clinical trials to limit missing data. *Stat Med* 2012;31:3433-3443.

10. Calvert M, Kyte D, Mercieca-Bebber R, et al. Guidelines for inclusion of patient-reported outcomes in clinical trial protocols: the SPIRIT-PRO extension. *JAMA* 2018;319:483-494.
11. Bottomley A, Pe M, Sloan J, et al. Moving forward toward standardizing analysis of quality of life data in randomized cancer clinical trials. *Clinical Trials* 2018 DOI:10.1177/1740774518795637.
12. Calvert M et al. Reporting of patient-reported outcomes in randomized trials: the CONSORT PRO extension. *JAMA* 2013;309:814-822.
13. Palmer MJ, Mercieca-Bebber R, King M, Calvert M, Richardson H, Brundage M. A systematic review and development of a classification framework for factors associated with missing patient-reported outcome data. *Clinical Trials* 2018;15:95-106.
14. Land SR, Ritter MW, Costantino JP, et al. Compliance with patient-reported outcomes in multicenter clinical trials: methodologic and practical approaches. *J Clin Oncol* 2007;25:5113-5120.
15. Osoba D, Zee B. Completion rates in health-related quality-of-life assessment: approach of the National Cancer Institute of Canada Clinical Trials Group. *Stat Med* 1998;17:603-612.
16. Hahn HA, Webster KA, Cella D, Fairclough DL. Missing data in quality of life research in Eastern Cooperative Oncology Group (ECOG) clinical trials: problems and solutions. *Stat Med* 1998;17:547-559.
17. Palmer MJ, Mercieca-Bebber R, King M, Calvert M, Richardson H, Brundage M. What evidence suggests a relationship between patient age and missing PRO data on randomized clinical trials (RCTs)? *Qual Life Res* 2017;26(Supplement 1):138-139.

18. Stang A. Nonresponse research – an underdeveloped field in epidemiology. *Eur J Epidemiol* 2003;18:929-931.

19. Osoba D, Bezjak A, Brundage M, Pater J. Evaluating health-related quality of life in cancer clinical trials: the National Cancer Institute of Canada Clinical Trials Group experience. *Value Health* 2007;10 Suppl 2:S138-45.

20. Sadura A, Pater J, Osoba D, Levine M, Palmer M, Bennet K. Quality-of-life assessment: patient compliance with questionnaire completion. *J Natl Cancer Inst* 1992;84:1023-1026.

21. Osoba D, Dancey J, Zee B, Myles J, Pater J. Health-related quality-of-life studies of the National Cancer Institute of Canada Clinical Trials Group. *J Natl Cancer Inst Monogr* 1996;20:107-111.

22. Bezjak A et al. Quality-of-life outcomes for adjuvant chemotherapy in early-stage non-small-cell lung cancer: results from a randomized trial, JBR.10. *J Clin Oncol* 2008;26:5052-5059.

23. Bezjak A et al. Symptom improvement in lung cancer patients treated with erlotinib: quality of life analysis of the National Cancer Institute of Canada Clinical Trials Group study BR.21. *J Clin Oncol* 2006;24:3831-3837.

24. Au H-J et al. Health-related quality of life in patients with advanced colorectal cancer treated with cetuximab: overall and kras-specific results of the NCIC CTG and AGITG CO.17 trial. *J Clin Oncol* 2009;27:1822-1828.

25. Ringash J et al. Quality of life in patients with k-ras wild-type colorectal cancer. The CO.20 phase 3 randomized trial. *Cancer* 2014;120:181-189.

26. Chow E et al. Dexamethasone in the prophylaxis of radiation-induced pain flare after palliative radiotherapy for bone metastases: a double-blind, randomised placebo-controlled, phase 3 trial. *Lancet Oncol* 2015;16:1463-72.

27. Brotto L et al. Randomized study of sequential cisplatin-topotecan/carboplatin-paclitaxel versus carboplatin-paclitaxel: effects on quality of life. *Support Care Cancer* 2016;24:1241-1249.

28. Fletcher J. What is heterogeneity and is it important? *BMJ* 2007;334:94-6.

29. Whelan TJ et al. Assessment of quality of life in MA.17, a randomized, placebo-controlled trial of letrozole after 5 years of tamoxifen in postmenopausal women. *J Clin Oncol* 2005;23:6931-6940.

30. Crook JM et al. Intermittent androgen suppression for rising PSA level after radiotherapy. *N Engl J Med* 2012;367:895-903.

31. Crump M et al. Randomized comparison of gemcitabine, dexamethasone, and cisplatin versus dexamethasone, cytarabine, and cisplatin chemotherapy before autologous stem-cell transplantation for relapsed and refractory aggressive lymphomas: NCIC CTG LY.12. *J Clin Oncol* 2014;32:3490-3496.

32. McDowell I. *Measuring health: a guide to rating scales and questionnaires.* Oxford University Press, New York, 2006.

33. Fayers PM, Aaronson NK, Bjordal K, Groenvold M, Curran D, Bottomley A. *The EORTC QLQ-C30 Scoring Manual (3rd Edition)*. Published by: European Organisation for Research and Treatment of Cancer, Brussels, 2001.

34. Webster K, Cella D, Yost K. The functional assessment of chronic illness therapy (FACIT) measurement system: properties, applications, and interpretation. *Health Qual Life Outcomes* 2003;1:79.

35. Vittinghoff E, Glidden DV, Shiboski SC, McCulloch CE. *Regression methods in statistics. Linear, logistic, survival, and repeated measures models. Second edition*. Springer, 2012.

36. Kleinbaum DG, Kupper LL, Muller KE, Nizam A. *Applied regression analysis and other multivariable methods, 3rd edition.* Duxbury Press: Boston, 1997.

37. Spruance SL, Reid JE, Grace M, Samore M. Hazard ratio in clinical trials. *Antimicrob Agents Chemother* 2004;48:2787-2792.

38. Hopwood P, Harvey A, Davies J, et al. Survey of the administration of quality of life questionnaires in three multicentre randomised trials in cancer. *Eur J Cancer* 1997;34:49-57.

39. Courvoisier DS, Eid M, Lischetzke T. Compliance to a cell phone-based ecological momentary assessment study: the effect of time and personality characteristics. *Psychol Assess* 2012;24:713-720.

40. Bernhard J, Cella D, Coates AS, et al. Missing quality of life data in cancer clinical trials: serious problems and challenges. *Stat Med* 1998;17:517-532.

41. Ganz PA, Haskell CM, Figlin RA, et al. Estimating the quality of life in a clinical trial of patients with metastatic lung cancer using the Karnofsky performance status and Functional Living Index – Cancer. *Cancer* 1988;61:849-856.

42. Concato J, Peduzzi P, Holford TR, Feinstein AR. Importance of events per independent variable in proportional hazards analysis. I. Background, goals, and general strategy. *J Clin Epidemiol* 1995;48:1495-1501.

43. Peduzzi P, Concato J, Feinstein AR, Holford TR. Importance of events per independent variable in proportional hazards regression analysis. II. Accuracy and precision of regression estimates. *J Clin Epidemiol* 1995;48:1503-1510.

44. Mercieca-Bebber R, Palmer MJ, Brundage M, et al. 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**:e010938.

45. Wright JR, Crooks D, Ellis PM, Mings D, Whelan TJ. Factors that influence the recruitment of patients to phase III studies in Oncology. *Cancer* 2002;95:1584-91

46. Price KC, Barrett BK, Roark JM. Cancer and Leukemia Group B clinical research associates committee. *Clin Cancer Res* 2006;12 (11 Suppl):3642-3644.

# Chapter 5

## What is the lived experience of clinical research associates with missing patient-reported outcome data?

5.1 Abstract

**Purpose:** Clinical Research Associates (CRAs) have a fundamental role in collecting patient-reported outcome (PRO) data from cancer patients in randomized clinical trials (RCTs) and are an untapped source for what may cause missing PRO data. The purpose was to describe the influences of missing PRO data, as revealed through the lived experience of CRAs.

**Methods:** Semi-structured interviews were audio-recorded with 11 CRAs, transcribed verbatim, and analysed according to descriptive phenomenology.

**Results:** All CRAs shared ten themes. Missing PRO data can occur in different patterns and may evoke an emotional reaction in the CRA. PROs for routine clinical care compete with PROs for RCTs. The paper and electronic formats have benefits and drawbacks. Instrument factors and patient factors influence missing PRO data. Assessment of PROs at disease progression is particularly difficult. Deficiencies in centre research infrastructure can lead to missing PRO data. CRAs develop relationships with patients that may help reduce missing PRO data. It is not always possible to adapt the assessment schedule and to provide sufficient time to complete the instrument. There is a need for field guidance, and a motivation and knowledge base to address issues.

**Conclusions:** PROs for routine clinical care compete with PROs for RCTs. Factors inherent to the instruments, patients, centres, staff, and RCTs may influence missing PRO data. Adapting the assessment schedule and providing sufficient time for instrument completion were challenging, particularly at disease progression. Developing field guidance would benefit from involving CRAs and their collective wisdom.

## 5.2 <u>Introduction</u>

Missing patient-reported outcome (PRO) data are problematic. A systematic review of the literature identified several factors reported to be associated with missing PRO data [1]. Front-line staff with experience collecting PRO data are an additional source for information regarding influences of missing PRO data. In Canada, Clinical Research Associates (CRAs) work at a cancer centre or hospital and have a fundamental role in collecting PRO data from cancer patients. Among many tasks, CRAs have responsibility for administering an instrument to patients on phase III randomized clinical trials (RCTs) and processing the data afterward.

A focus group with CRAs in the Canadian Cancer Trials Group (CCTG, a multi-centre cooperative oncology group with history collecting PRO data in RCTs [2]) provided a window into their experience with missing PRO data. CRAs revealed an enthusiasm for the topic, contributed their knowledge of the clinic environment, and demonstrated an in-depth understanding of cancer patients. It became clear that CRAs were an untapped source of information relevant to identifying the causes of missing PRO data.

The objective was to describe the lived experience of CRAs with missing PRO data. To address this, we conducted in-depth semi-structured interviews of individual CRAs with questions framed within descriptive phenomenology.

## 5.3 <u>Methods</u>

The study methods followed the recommendations articulated by Morse *et al* [3-4] to ensure their rigour.

5.3.1 Investigator Responsiveness

As personal assumptions and prior knowledge may influence the research process, MP developed a reflexivity statement [5], reflected throughout the study, and met regularly with a qualitative research expert (TK).

5.3.2 Methodological Coherence

Descriptive phenomenology [6], which originated in the philosophy of Husserl [7], was the inspiration for methodology. Rather than insisting on each participant's pre-reflective experience and searching for the universal essence of the phenomenon (which are characteristics of European phenomenology), the study explored CRAs experience with missing PRO data *per se* and their context of potential influences leading to missing PRO data [8]. This practical approach applied to research focused on the reality of CRAs experience as they engage with the phenomenon through their daily activities [9].

The data collection method was semi-structured interviews [10-11], either held face-to-face or over the telephone. A letter of invitation was sent via e-mail and followed with confirmation of format, date, and time. Each CRA was given the informed consent form, provided with a brief overview of the study, and had opportunity to ask questions before beginning the interview. Participation was voluntary, involving no financial incentives.

The interview guide framed discussion within the CRA's lived experience with missing PRO data, ensuring similarity across interviews. All questions were open-ended and included prompts to clarify intent and to elicit additional information. The CRA had opportunity to debrief at the end with the audio-recorder turned off, then provided background and demographic information on a one-page form. Subsequently, MP prepared a field note to capture salient impressions.

5.3.3 Purposeful Sampling

In qualitative research, judgement is used, rather than sample size calculation, to balance a lack of compelling data from too few participants with the cost of increasing the number of participants to achieve saturation [12]. Maximum variation sampling was used to select information-rich participants [13], to document uniqueness and shared patterns that cut across CRAs. The criteria for constructing a diverse sample included the geographic location of the centre (within the western provinces, Ontario, Quebec, or eastern provinces of Canada), and the presence or absence of a post-graduate residency training program in oncology at the centre. This approach identified CRAs having a range of experience with respect to missing PRO data. At most, 20 CRAs would participate [14]. A pool of information-rich candidates was created according to the sampling criteria, from the CCTG roster of approximately 950 CRAs in Canada, with CRAs having involvement on specific RCTs or membership on Disease Site Committees, and supplemented with CRAs having previous participation in the focus group or by recommendation of a key informant. Candidates were individually invited for interview. Three unsuccessful attempts at contact by e-mail prompted purposeful sampling of another CRA.

5.3.4 Data Management and Analysis

Each interview was audio-recorded and transcribed verbatim by an experienced transcriber. The interviewer (MP) reviewed each transcription with the audio-recording and made edits, assuring the quality of content and preserving anonymity of the CRAs.

Consistent with descriptive phenomenology, the analysis involved bracketing, horizonalization, and synthesis of findings, following modification of the Stevick-Colaizzi-Keen method [15]. Bracketing, a fundamental methodological concept [16], mitigated the influence of MP, who set aside personal assumptions and prior knowledge, to be open to the data. Horizonalization included several steps. 1) Identifying relevant statements in the transcription by considering the significance of each statement with the research question. 2) Listing the relevant statements and evaluating the uniqueness of each. 3) Identifying invariant meaning units by removing the repetitive and overlapping (redundant) statements. 4) Organizing invariant meaning units into clusters, and grouping related clusters into themes. Synthesis of findings involved preparing a textual description of CRAs experience with missing PRO data and a structural description of contexts influencing missing PRO data. Verbatim examples were selected to illustrate key information and to summarize the scope of experience.

5.3.5 Credibility

All CRAs were active, having personal experience with missing PRO data, on RCTs. The analysis followed principles and included direct support to ensure MP was not alone in the process [4]. MP and TK participated in an investigator triangulation exercise [17], then all authors participated in a face-validity exercise, reaching consensus on the final themes and the final clusters of meaning statements. The overall description was thick and rich with data provided by 11 CRAs.

**Appendix E.3** provides the letter of invitation, informed consent form, interview guide, background and demographic information form, an example of field note, and summary of the triangulation and face-validity exercises.

## 5.4 <u>Results</u>

Eleven CRAs had interviews between February 2018 and October 2018. Two CRAs explicitly declined, for reasons of other priority and recent retirement. The conversations were mostly face-to-face and lasted on average 61 minutes. Six of the CRAs worked at centres in Ontario, and of these, four had a post-graduate residency training program (**Table 5-1)**. Most CRAs were more than 30 years of age, had college or higher education, and had more than ten years of experience. Eight CRAs reported receiving specific training for collection of PROs on RCTs (**Table 5-2)**. The 10 themes, shared by all CRAs, and the associated clusters of meaning statements, are presented in **Table 5-3.** The following sections provide highlights.

### 5.4.1 <u>Missing PRO data can occur in different patterns and may evoke an emotional reaction in the CRA</u>

Missing PRO data were an infrequent experience. One CRA said 'over the years, for my studies, it's very not the norm to miss it' and another CRA explained 'our site has been very good throughout the years in collecting them'.

The CRAs had experience with different types of missing PRO data. One CRA recalled a lung cancer patient and described how the entire instrument may be missing.

> Half way through her treatment she said I don't want to complete them. She flat out refused. I'm not sure if she was thrown off from that questionnaire at the beginning.

A CRA described how a questionnaire within the instrument may be missing.

> The longer ones . . . they are probably doing maybe four or five questionnaires, designed to assess different things . . . the sponsor has taken three or four of them, piling them in and having multiple ones that look different. A lot of them too will have a face sheet. If the face sheet isn't able to be taken off . . . that might prompt them to think "oh I am

> done", because the face sheet doesn't look like anything that you do. . . . Often if they're multiple ones stuck on top of each other with face sheets then sometimes they miss some.

Several CRAs experienced a missing page of an instrument because the patient 'didn't flip through it properly' or 'they flipped through two pages at once', which was 'often just an oversight' while the patient was going through the instrument.

All CRAs experienced missing responses to specific questions, recalling patients 'going fast reading', 'simply miss a question', or 'they forgot it'.

The missing PRO data can precipitate negative feelings in the CRAs who take this responsibility seriously. One CRA described coordinating a scheduled assessment with a nurse: 'it's very frustrating . . . CRAs get them back and they're not completed. It goes in a big circle of not getting done'. Another CRA said 'It really feels terrible to forget one of those questionnaires. I've really missed sometimes'. A third CRA said 'We always feel bad when it happens (laughs). It is never intentional'.

#### 5.4.2 <u>PROs for routine clinical care compete with PROs for RCTs</u>

In many provinces, PROs for clinical care are mandated and used as symptom screening tools. All patients are required to complete an assessment of their symptoms before their appointment with the doctor.

> . . . [name of cancer agency deleted] I think mandates it. Patients are told every time they're seeing a doctor they have to do that assessment. If they are coming in for just treatment they don't do it.

Sometimes patients are not sure why they are asked similar questions in an instrument for a RCT after previously completing an instrument for clinical care. Different CRAs had a similar experience. For example, one CRA described a patient response “I already answered this question. I already said if that I was tired”.

Patients can experience overload of what they experience as duplication, and will refuse to complete the instrument for a RCT after previously completing an instrument for clinical care. Two CRAs explained,

> Everybody’s asking them the same questions, in a different context, so it I think at some point, especially after when patients come into the centre frequently, it’s a bit of an overload.
>
> Patient compliance can be an issue. Sometimes they’ll say “I don’t want to do it” . . . that’s what they think “**again**?”

### 5.4.3 Both the paper format and the electronic format have benefits and drawbacks

While the paper format is still used for some trials, the CRAs identified a trend towards collecting PRO data with electronic devices. Each format was experienced as having characteristics that can increase or reduce missing PRO data (**Table 5-4**).

CRAs described their experience collecting PRO data with an electronic tablet (**Table 5-5**). The comments reflect variability in the quality of the technology, but also its availability for patients. Some trials have one device for all patients, or one device for every patient, to use in the centre, while other trials require patients to take the device home and complete assessments on an ongoing basis, or as one CRA indicated ‘for the rest of their life’.

Older patients in particular find technology to be challenging and some will not complete an instrument on a device (**Table 5-6**). The comments reflect diversity of experience regarding a criterion for older patients but consensus that technology may not be appropriate for all patients.

Interestingly, none of the CRAs described any patient having an emotional reaction to the paper format. In contrast, CRAs described some patients being frightened by an electronic device while other patients enjoyed been part of a study with new technology. The electronic format was not the unanimous preference of CRAs (**Table 5-7**), but was thought best for future patients. One CRA said,

> In the next few years, more people are using electronic devices, it will be less of a concern. That is the way of the future.

However, another CRA expressed concerns for these devices in the present.

> Maybe future generations that certainly will be the way to go. Right now not always. That will cause missing data. They can't use the device (laughs).

#### 5.4.4 Missing PRO data are influenced by the quantity, repetition, and wording of questions, and by the format of the instrument

CRAs described how the number, repetition, and appearance of questions within the instrument can lead to missing PRO data. In some instruments, the wording of questions can be confusing or difficult to understand (**Table 5-8**), or address sensitive content (**Table 5-9**).

#### 5.4.5 Missing PRO data are influenced by personal characteristics of the patient, attributes of the cancer, and toxicity from the treatment

The CRAs noted personal attributes they experience as influencing missing PRO data including age and gender. With respect to age, CRAs have different experiences regarding whether age is an influence (**Table 5-10**). With respect to gender, CRAs described how patients may respond to particular questions. Women, for example, are typically more likely to skip certain sections if they feel uncomfortable with them. Men aren't as embarrassed about certain questions and may be more comfortable answering, but will often miss a subsequent page in the instrument by forgetting to turn the current page after answering the questions.

#### 5.4.6 Assessment of PROs at progression of disease is particularly difficult

A typical RCT requires assessment of PROs at specific time points, to measure the health status of patients during treatment and follow-up. On some RCTs, the CRAs are expected to administer the instrument at the time of disease progression, which can be challenging to do for several reasons. For example, being aware of the patient's disease progression may evoke an emotional reaction in the CRA that hinders them from administering the instrument (**Table 5-11**). Also, the CRAs noted the patient's personality they experience can influence whether the instrument is completed at progression (**Table 5-12**).

#### 5.4.7 Deficiencies in research infrastructure at the centre can lead to missing PRO data

CRAs work on many different trials. One CRA, after being asked whether trials were an influence of missing PRO data, described a trial sponsored by pharma that was more strict requiring high compliance, in comparison to other trials. 'We really have to make sure that the patient fills everything out if that's what they want to do, triple checking and making sure everything's done right.' The CRA did not elaborate on the mechanism of funding but conveyed the notion priority given to trials with higher funding can lead to missing PRO data on trials with lower funding: 'Sometimes it's not possible. We have to trust that [pause] it gets done.'

#### 5.4.8 CRAs develop relationships with patients that may help reduce missing PRO data

Building a rapport with the patients was an important step in developing good relationships so that the patients feel comfortable doing what they are asked. One CRA observed patients 'want to help you do your job'. Another CRA described making a patient with metastatic pancreatic cancer comfortable in the clinic.

> She'd been through so much up in [name of city deleted]. She had a feeding tube . . . I went up three times to talk chat with them. I had other things to give them. I just sat and felt comfortable just sitting chatting with them. It was his birthday that day, we talked about that. It was her birthday coming . . . She came with her spouse . . . I just wanted them to feel comfortable in the clinic . . . This study, it's not easy, they're in three weeks out of four for chemo. The last time I saw them I said "I'll be back next week to see you see how you're doing".

Being in clinic helped the CRAs to strengthen their relationships, necessary because they administered questionnaires to the patients. One CRA articulated the reality of their experience.

> You get attached to them. You have to. I don't see how you don't . . . One patient died on adjuvant chemo. I think she had a reaction to that cyclo . . . She had no family. She asked [name of physician deleted] if I'd come up. She was just bloating.

Three CRAs indicated being a resource for the patients was a key responsibility.

> They know that we're a resource at their demand . . . We find out all these answers for them . . . Scheduling issues causes a lot of anxiety so it's just nice that we're there. I still have a patient, he's been off the trial treatment for 5 months. I still get emails . . . things get ironed out . . . It's not part of my job. I don't mind going above and beyond.

> Being there for the patients, they know you're their coordinator. You're the go to. That's a big part of it too. You have to be there.

> . . . we are the telephone triage, we take all the calls from our patients that are not doing well. If they are on immunotherapy, we're always open to the phone calls.

Clearly, CRA responses suggest empathy for the patients. One CRA recalled a patient with bladder cancer having difficulty completing an instrument and said 'You just think oh the poor person'. Another CRA described how treating a single metastasis with radiation can often delay progression, but for one patient 'it came back [pause] uhm much earlier than expected. It was kind of sad.' This empathy highlights the CRAs sensitivity to patients when collecting PRO data.

### 5.4.9 It is not always possible to adapt the assessment schedule and to provide sufficient time to complete the instrument

CRAs described many challenges with administering the instrument according to the assessment schedule of the protocol. Among these challenges is providing sufficient time for the patient to complete the instrument in a clinic room, which may be difficult because of interruptions. The patient is put in a clinic room and given the instrument to complete, but may set the instrument aside when the doctor arrives and not complete it. Interrupting the patient seemed to be more frequent in some types of centres, where many people were involved. As described by one CRA,

> We're in a teaching hospital. The patients are put in a room. There's a nurse that has to go and do an assessment. There's often a medical student or a resident that goes in, then the doctor. You have very little time to get done what you need done.

### 5.4.10 There is a need for field guidance, and a motivation and knowledge base to address issues

The CRAs offered suggestions regarding what can be done to minimize missing PRO data. Several CRAs recommended improvements to the instrument. Including the time frame such as 'in the past seven days' in each question rather than providing it at the beginning of a group of questions would remind patients to answer each question. Asking only essential questions, rather than asking several questions about the same thing in different ways, would 'minimize the stress on patients'. Creating a cohesive instrument with everything well-spaced would reduce missing PRO data. Many suggestions related to how structures and processes could be put in place. These include providing training, having time and space to complete the instrument, and allowing flexibility with the schedule of assessment so that the CRA has more opportunity to collect PRO data from the patient. Organizing the documents in a plastic sleeve at the front of the chart was noted as a practical way to keep the documents together so that they aren't missed.

## 5.5 Discussion

We obtained information on influences of missing PRO data from interviews with CRAs. Missing PRO data are an infrequent experience that occurs in different patterns and may evoke negative feelings in the CRAs. The infrequency of the phenomenon and emotional reaction upon its occurrence reflect the competence and dedication of CRAs in collecting PRO data from cancer patients on RCTs. Collectively, the CRAs care about data quality.

Collecting PRO data involves the patient being given an instrument at baseline, indicating their answers to the questions, and returning it, repeating this action at subsequent time points during treatment and follow-up. While the process appears straightforward, paradoxically, there are inherent complexities. It is apparent from the CRAs descriptions that differences – both within and across the instruments, patients, centres, staff, and RCTs – contribute to making the collection of PRO data challenging. Although a multiplicity of factors may potentially influence missing PRO data, the findings of this study identify a finite number of themes, with each describing separate influences.

PROs for routine clinical care, like the Edmonton Symptom Assessment System [18] and the Expanded Prostate Cancer Index Composite for Clinical Practice [19], can help health care practitioners improve the clinical care of cancer patients who, as mandated by provincial health authorities, are required to complete a symptom screening tool *before* an assessment for an RCT. This finding raises awareness that, in many provinces of Canada, there is an expectation of patients that is external to and independent of their participation in a given RCT, which may lead to missing PRO data on the RCT through patient uncertainty or overload.

The CRAs did not describe a predominant influence of missing PRO data, but instead described several important influences inherent to the instruments (quantity, repetition, wording of questions, and format), patients (characteristics, attributes of the cancer, and toxicity from the treatment), centres (deficiencies in research infrastructure), staff (developing relationships with patients), and RCTs (data capture in electronic versus paper format). Also, adapting the assessment schedule and providing sufficient time for the patient to complete the instrument were challenging for the CRAs, particularly at disease progression. There is a need for field guidance, and a motivation and knowledge base among CRAs to develop such guidance.

These descriptions add to our previous knowledge [1]. The results contribute new evidence that PROs for routine clinical care compete with PROs for RCTs. The results contribute additional qualitative evidence to the existing literature for several factors involving patient characteristics, instruments, research infrastructure, and data capture format.

The main strength of our study is the incorporation of recommended strategies to improve rigor. Additional steps were taken to ensure credibility. We consider each theme, and the associated clusters of meaning statements, to have importance. The high-quality, detailed descriptions derive significance emerging from sampling heterogeneity [13]. The CRAs, having experience on CCTG RCTs, were purposefully sampled from university-based or community centres in four different geographic locations of Canada, and share every theme. We are confident each theme accurately reflects the associated clusters of meaning statements, and every cluster is relevant to missing PRO data. Saturation, the extent that data are comprehensive and provide a range of perspectives, was achieved when no new themes emerged after 11 interviews [20].

There are limitations to this study. CRAs may have withheld some information regarding influences of missing PRO data. One CRA said after the interview, with the audio-recorder turned off, that they would not admit making an error. The effect of the audio-recorder on content is not known. Interviews were planned but were not conducted with CRAs at centres in the East and in Quebec having a post-graduate residency training program, because saturation was achieved. The results may be limited by their absence but the effect is not known.

CRAs have a fundamental role in collecting high-quality data [21]. Surprisingly, qualitative inquiry of research staff has been used uncommonly to identify influences of missing PRO data. In our systematic review of the literature [1], one study from the United States [22] used a focus group to identify issues with retention. Subsequently, investigators from the United Kingdom [23] and Australia [24] used interviews of research staff to understand current practices for identifying missing PRO data and procedures for following up. Our study, through interviews with CRAs in Canada, demonstrates CRAs are insightful, and their experience contributes new and additional evidence for the influences leading to missing PRO data.

### 5.6 Conclusion

PROs for routine clinical care compete with PROs for RCTs. Factors inherent to the instruments, patients, centres, staff, and RCTs may influence missing PRO data. Adapting the assessment schedule and providing sufficient time for instrument completion were challenging, particularly at disease progression. Developing field guidance would benefit from involving CRAs and their collective wisdom.

### 5.7 Acknowledgements

The authors thank: 1) the Clinical Research Associates who generously volunteered their time and participated in the interviews, 2) Janice Flegg and Tracie Hanna who were very helpful in identifying candidate CRAs, and 3) Adam Lodzinski who expertly moderated the focus group in the preliminary qualitative work.

**Table 5-1. Characteristics of centres and interviews.**

| Clinical Research Associate | Centre | | Interview | |
|---|---|---|---|---|
| | Geographic Location | Presence of Post-Graduate Residency Training Program | Type | Length (minutes) |
| A | Ontario | Yes | Face-to-face | 91 |
| B | Ontario | Yes | Face-to-face | 77 |
| C | Ontario | Yes | Face-to-face | 63 |
| D | Ontario | Yes | Face-to-face | 48 |
| E | West | No | Face-to-face | 52 |
| F | East | No | Face-to-face | 70 |
| G | Ontario | No | Face-to-face | 36 |
| H | Quebec | No | Face-to-face | 55 |
| I | West | Yes | Telephone | 65 |
| J | West | Yes | Telephone | 64 |
| K | Ontario | No | Telephone | 55 |

**Table 5-2. Characteristics of participants.**

| | | |
|---|---|---|
| Worked as a Clinical Research Associate at another centre | | |
| | Yes | 2 |
| | No | 9 |
| Worked as a Clinical Research Associate for an organization other than Canadian Cancer Trials Group [1] | | |
| | Another cooperative group | 9 |
| | Industry | 10 |
| | Within centre | 7 |
| | Other | 1 |
| Age in years | | |
| | < 21 | 1 |
| | 21–30 | 1 |
| | 31–40 | 4 |
| | 41–50 | 0 |
| | 51–60 | 5 |
| Highest level of education | | |
| | Secondary (high) school diploma or certificate | 1 |
| | Registered Apprenticeship or other trades certificate or diploma | 1 |
| | College, CEGEP [2], or other non-university certificate or diploma | 4 |
| | University bachelor's degree | 3 |
| | University master's degree | 2 |
| Self-identification [1] | | |
| | Certified Clinical Research Coordinator [CCRC® (ACRP [3])] | 1 |
| | Certified Clinical Research Professional [CCRP® (SOCRA [4])] | 6 |
| | Clinical Research Associate | 7 |
| | Clinical Research Nurse | 1 |
| | Registered Nurse | 1 |
| | Clinical Trials Data Coordinator | 1 |
| Years of experience with collection of patient-reported outcomes on randomized clinical trials | | |
| | 1–5 | 2 |
| | 6–10 | 2 |
| | 11–15 | 4 |
| | 16–20 | 1 |
| | 21–25 | 0 |
| | > 25 | 2 |
| Received specific training for collection of patient-reported outcomes on randomized clinical trials | | |
| | Yes, from a site initiation visit | 4 |
| | Yes, for an electronic system or device | 2 |
| | Yes, from a course, seminar or numerous webinars | 1 |
| | Yes, from case report form guidelines for entering data | 1 |
| | No, relied on the protocol with advice from colleagues | 2 |
| | No, relied on self-education from journals and on-line literature | 1 |

[1] Total exceeds 11 as more than one answer provided, [2] Collège d'enseignement général et professionnel,
[3] Association of Clinical Research Professionals, [4] The Society of Clinical Research Associates

**Table 5-3. Themes (bold) and the associated clusters of meaning statements (bullets) identified in analyses of data from Clinical Research Associates (n = 11).**

| |
|---|
| **Missing PRO data can occur in different patterns and may evoke an emotional reaction in the CRA.**<br>• Missing PRO data are an infrequent experience.<br>• The entire instrument, a questionnaire within the instrument, a page of an instrument, or responses to specific questions, may be missing.<br>• Missing PRO data can precipitate negative feelings in the CRA. |
| **PROs for routine clinical care compete with PROs for RCTs.**<br>• In many provinces, PROs for clinical care are mandated and used as symptom screening tools.<br>• Sometimes patients are not sure why they are asked similar questions in an instrument for a RCT after previously completing an instrument for clinical care.<br>• Sometimes patients experience overload and will refuse to complete the instrument for a RCT after previously completing an instrument for clinical care. |
| **Both the paper format and the electronic format have benefits and drawbacks.**<br>• Each format has characteristics that can increase or reduce missing PRO data.<br>• CRAs have responsibility for entry of data collected by paper and could make a data entry error.<br>• The quality and availability of the technology depends on the vendor and may not be apparent when a centre decides to join the RCT.<br>• Older patients in particular find technology to be challenging and some will not complete an instrument on a device.<br>• The electronic format is not appropriate for some scenarios, such as for follow-up when the patient stops protocol treatment and is receiving palliative care.<br>• The electronic format can evoke an emotional reaction in the patient, is not always the unanimous preference of CRAs, but is thought to be best for future patients.<br>• The electronic format may not offer an improvement in compliance over the paper format. |

**Table 5-3. (continued)**

| |
|---|
| **Missing PRO data are influenced by the quantity, repetition, and wording of questions, and by the format of the instrument.**<br>• The length of the instrument can lead to missing PRO data.<br>• Repetitive questions in an instrument are frustrating for patients.<br>• The wording of questions can be confusing or difficult to understand.<br>• The content addressed by some questions can be sensitive to some patients.<br>• A smaller font size and busy page layout contribute to missing PRO data.<br>• Some questions relating to treatment are difficult for the patient to answer at baseline when treatment has not been received. |
| **Missing PRO data are influenced by personal characteristics of the patient, attributes of the cancer, and toxicity from the treatment.**<br>• CRAs have different experiences regarding whether age is an influence.<br>• Women are typically more likely to skip certain sections whereas men will often miss a page.<br>• Retired people are often more willing to complete an instrument.<br>• Some patients have difficulty reading and require assistance to complete the instrument. A caregiver can help, by reading the questions to the patient and recording the answers.<br>• Some patients may require use of an interpreter or translator but if they have to leave, the patients can’t complete the instrument.<br>• Finding out the assigned treatment was standard of care may make the patient less likely to complete the instrument.<br>• Patients with a diagnosis of brain cancer have more difficulty completing the instrument.<br>• Patients on adjuvant therapy may be more likely to complete the instrument but miss some of the sections.<br>• Some patients are fatigued and don’t want to complete an instrument.<br>• Some patients develop a mental fog from chemotherapy which can compromise their ability to complete an instrument. |
| **Assessment of PROs at progression of disease is particularly difficult.**<br>• The patient’s deteriorating health may make them unable to attend the clinic.<br>• The CRA may suspect that the patient is progressing but the doctor has sole authority to inform the patient of disease progression.<br>• The CRA has responsibility to perform trial requirements and may administer the instrument with bad news pending.<br>• Field guidance from the sponsor may encourage the CRA to collect PRO data at progression, but the CRA may not do it (with support of the doctor).<br>• The CRA may not be aware the patient has progressed or is not able to administer the instrument before the patient is informed of progression.<br>• Being aware of the patient’s disease progression may evoke an emotional reaction in the CRA that hinders them from administering the instrument.<br>• Being informed of disease progression by the doctor may evoke an emotional reaction in the patient that hinders them from completing the instrument.<br>• The patient’s personality influences whether the instrument is completed at progression. |

**Table 5-3. (continued)**

| |
|---|
| **Deficiencies in research infrastructure at the centre can lead to missing PRO data.**<br>• Inadequate space in the clinic to meet with the patient, lack of organization in preparing the chart for the patient's next visit, and priority given to RCTs with higher funding can lead to missing PRO data.<br>• Not having sufficient CRAs for back-up can lead to missing PRO data.<br>• Forgetting to administer the instrument to the patient, or not checking whether the patient completed the instrument, can lead to missing PRO data. |
| **CRAs develop relationships with patients that may help reduce missing PRO data.**<br>• CRAs develop a rapport with the patients to help them feel comfortable in the clinic.<br>• The CRAs are a resource for the patients.<br>• The CRAs have empathy for the patients.<br>• Frequent visits (such as, during treatment) reinforce familiarity. |
| **It is not always possible to adapt the assessment schedule and to provide sufficient time to complete the instrument.**<br>• The amount of time spent with the patient is variable but is less after randomization.<br>• The patient has insufficient time to complete the instrument in the clinic because of interruptions.<br>• The CRA is busy and is not able to administer the instrument.<br>• The necessity for a test or procedure on the day of assessment can make it difficult for the CRA to administer the instrument before the patient is seen by the doctor or receives treatment.<br>• The patient has had a test or procedure and, anticipating receipt of bad news, does not want to complete the instrument.<br>• Rigid adherence to an assessment schedule can lead to missing PRO data when treatment is deferred but the patient is still expected to come in to clinic to complete the instrument.<br>• A delay in treatment, because the patient is not feeling well or blood counts are too low, can necessitate repeating the assessment but the patient may be unwilling to comply.<br>• It's time consuming for the patient to complete an instrument and some patients will leave the clinic because they are working or have other appointments.<br>• The patient may complete the instrument on a day that was beyond the allowable window for the expected day of assessment. |
| **There is a need for field guidance, and a motivation and knowledge base to address issues.**<br>• Trials are getting much more complex.<br>• It seems like there is more missing PRO data now.<br>• CRAs have enthusiasm for cancer research and are concerned about missing PRO data.<br>• CRAs wonder about the quality of the PRO data that are collected.<br>• The expectation for CRA review of PRO data when collected in a diary may be different from when PROs are collected in an instrument.<br>• The contribution of a monitor to reduce missing PRO data at the centre is not consistent across RCTs.<br>• The CRAs have collective wisdom and can be a resource regarding what can be done to minimize missing PRO data. |

**Table 5-4. Each format has characteristics that can increase or reduce missing PRO data.**

| Characteristics that can increase missing PRO data | |
|---|---|
| Paper | Electronic |
| Sometimes patients don't complete all the questions | Need to charge |
| Can provide two answers for the same question | Have to load them |
| Not able to provide comments to sponsor | Had problems downloading |
| Having to spend time photocopying | Are not reliable |
| Photocopying error | Don't have a backup |
| Preserving anonymity | Need wireless connection |
| Sending it off to the data people | Assessment takes more time |
| Can go missing | Finicky to touch |
| Having to keep a copy | Having to keep a copy |
| | Hard for patients |
| | There are different procedures for Canadian patients |
| | It is possible for a patient to provide data for one of multiple questionnaires on the tablet |
| | Have to use disinfectant |
| | Need to have a code |
| | Missing can affect the whole study |
| | Flow of administration is different and more cumbersome |
| **Characteristics that can reduce missing PRO data** | |
| Paper | Electronic |
| Availability | Can't miss any questions |
| Don't need to charge | Cannot provide two answers for the same question |
| Can move information on the cover sheet | May prompt assessment |
| Can record comments | Don't have to print documents |
| Can put several on patient's chart | Less paper in charts |
| | Don't have to enter data |
| | Data automatically go to the sponsor |
| | Getting e-mails from the sponsor |
| | The system can flag when health utilities has been missed |
| | Keeping track of patients in a web-based system |
| | Have to give it to the patient physically |
| | CRA is not supposed to see the answers |

**Table 5-5. CRAs describe their experience collecting PRO data with an electronic tablet.**

| CRA | Quote |
|---|---|
| A | The bane of our existence right now are these devices. I find that a big hindrance right now is these electronic version of that. That’s a big mistake, common mistake I would say now. They’re nothing but headaches the electronic devices I find. They’re frustrating. |
| B | The electronic, prior to going onto the trial, I have to register myself an account, whoever else is my backup for the trial, generate accounts. I log in with my account. We screen a patient. I add that patient to the e-PRO, randomized onto the trial. Prior to every visit I log in, activate that visit for the patient, I turn it over to them. |
| D | We’ve had some studies using electronic tablets or handhold devices. They’ve pretty much been disasters (laughs). |
| E | On one study that we do, that’s advanced breast, that’s a drug company one . . . it’s on a laptop. |
| F | The research part of it the quality of life to administer either on a PRO tablet on an e-PRO tablet, six or seven studies. We’ve had them for five six years different studies with uhm tablets. It seems uhm [pause] less personal. The electronic device is more the problem. |
| G | umhum! yup! yup<br>[Response to And were all the patients on that study expected to complete an instrument with a tablet?] |
| H | We have a tablet in the hospital for some studies. It’s basically like a paper questionnaire but on a tablet. It’s fairly easy to complete. |
| I | We can use the iPad here in [name of city deleted]. We do have a system where certain people get are given uhm an iPad to report . . . on the radiation side . . . a pilot study, several main questionnaires . . . for the GU patients, some head-and-neck questionnaires. It looks like a nice way to uhm collect that right now. |
| J | Newer studies are, yes. We have two of them. Breast study, an oral uhm drug, do uhm the quality of life questionnaire on a **tablet**. They do it uhm [pause] not at every [pause] visit but at every third visit [pause] uhm for the breast study. Prostate study, they complete it using the tablet, a little bit easier. It’s depending on the type of tablet used. It’s also uhm important to note that there are [pause] I think better ones (laughs). |

**Table 5-6. Older patients in particular find technology to be challenging and some will not complete an instrument on a device.**

| CRA | Quote |
|---|---|
| A | We have prostate patients that come in that are in their eighties that are trying to figure out these tablets. Some of them will navigate it but they've never dealt with the piece full . . . We're asking them to take these tablets home and do these diaries and these PROs on these tablets. People that have never been exposed to an iPad or a tablet or those things that it really is a foreign type thing to have to use a device like that . . . You've got to touch, give an extra press [sound of finger hitting table several times] They jam at it or their fingers seem to be too big for the little button. I don't like it. I saw them. I don't even know how they would do the little tiny phone cell phone like pad, get frustrated with the touch screen inside of it, as do I. |
| B | He's elderly, he wouldn't have a cell phone, he's just old school. He has to tap the button. For whatever reason, I don't know if he's using his fingernail, a lot of them use their fingernail [sound of finger hitting table several times] it's like "no no no you've got to use your finger". It's just not very sensitive. He gets kind of frustrated because it's not registering his answers . . . It's just like a tablet, but they give you a stylus pen, that works okay. I tried giving him one of this stylus pen, that didn't seem to really help him any, he just kind of struggles through it . . . Some of the older, it's a little bit overwhelming for them. |
| D | Older patients can't grasp it. You've got to look at your demographics. It's not so good for older ovarian cancer ladies, prostate men. |
| F | Some of them have never used an iPad because they're in the older age. 60 and lower take the tablet, they can handle that, they've used iPads and iPhones and stuff. The younger ones are more comfortable. The tablets are very simple. There's that nice big blue arrow they know they have to touch it. They touch every little tab. They understand. The older ones are not as comfortable. |
| G | An older lady, in her late seventies, a hem(atology) patient, CLL <Chronic Lymphocytic Leukemia> trial. She just hated that (laughs). She didn't want to do it. She just didn't like it. She sometimes either flat out refused or would skip through. It was technology that she wasn't comfortable with. |
| I | I've found a difference when uhm when I was doing the SPROUT sys(tem) the SPROUT system for SC24. There was an option to go online and even a scan option. The older patients didn't really follow that, they did not like that, it's too advanced for them. They didn't even know what the the scan bar code what that means. If we do convert to a system that [pause] allows [pause] patients to either go online and login and type in an an address and type in PINs and things like that they won't [pause] there may be a [pause] uhm [pause] you know a difference in age. I feel the younger generation are more comfortable doing. The older generation would probably prefer the paper version. That had the most issues with [pause] the computer system was I think fifty-five and above. |
| J | Our elderly patients that are seventy eighty years old are not used to the electronic [pause] devices. They may you know "give me a newspaper and a pencil" (laughs). There might be some intimidation with how to complete the quality of life electronically. They do complete it. There's just a lot of anxiety there for the elderly when it comes to electronic devices. It is quite a uhm concern for the [pause] the patients . . .The concern is that it's some cancer patients and the age. They are more prone to the paper and less to any electronic devices. |

**Table 5-7. The electronic format was not the unanimous preference of CRAs.**

| CRA | Quote |
|---|---|
| B | I don't have any problems with [the electronic format] myself. I find it very helpful. I do! I like it. It is good! That's my preference anyway. |
| D | I'm not a fan of the electronic PROs right now. I like paper. It works. Unless you don't have a pen (laughs). |
| H | There's still paper. There's this company actually who brought back to papers. Some sponsor are going back to paper. I think for now I prefer paper because with a device we often have issue, especially when we try to connect. The best would be to have a uhm something that we can connect in the computer [pause] to update the device so it's ready for the visit. Right now we always have to rely on the Wi-Fi or GSM <Global System for Mobile Communications> to connect [pause] and right now I think this is the weakest. |
| I | The reason why I like the iPad is because I have to program it ahead of time, and then I give it to the patient physically, and then they fill it out, and then I can check afterwards that it's been sent off, and then I'm done. |
| J | I enjoy the paper. |

**Table 5-8. The wording of questions can be confusing or difficult to understand.**

| CRA | Quote |
|---|---|
| A | They're wanting to answer them but they just, they get frustrated with the types of questions that are being asked. I think it's the way they're formed . . . "What is your best [pause] health today?" throws people when they've been told they've got an incurable cancer. I think that almost is like a slap in the face to them, **whereas** other people would read that question and have you know record their (answer). Some people read that sentence or will read a sentence and get hung up on the way it's posed. That's what makes them not answer it, or not want to answer it, or have a hard time, or ask me for how they should answer it. |
| B | I always hate this one too. It's like a scale "How would you rate your health today?" 100% is best imaginable and zero is worse than imaginable. People say "Well how do I know what worse than imaginable is? I guess I'd be dead!" (laughs) |
| E | Because they'll go "how am I supposed to answer that like today I feel", no no it's not today it's in the last seven days or two weeks. |
| H | Language can be a barrier even if the patient always been speaking English or French. Sometimes the language used is not what they currently use at home. They're not sure what it means or how to answer to it . . . That could be an issue. Once in a while we have patients who really don't understand some of the words, "belching", that was one of the words. There was another one in French but now it doesn't come to my mind (laughs). The patient was speaking French but then read the sentence. She couldn't remember what was the uhm the meaning of that word. |
| I | There's a section that talks about pain. One of my patients always fills out "I have pain". Because it's such a general question about pain he answers to what he believes is you know general pain. He was saying "well it's not related to this study and I've always had this pain". He always writes us a note on the side "it's from an ankle injury years ago". This happens every time . . . He's wondering "am I supposed to relate this pain specifically to the area of treatment?" or "am I supposed to be just answering as in general which is the ankle pain I've had since uhm since uhm you know for twenty years because I had a I was playing some sort of sport and we were injured?" |
| J | Something to be said for the type of quality of life questionnaires. Have to be the simple ones [pause] uhm nothing too elaborate or or difficult for them to understand. |

**Table 5-9. The content addressed by some questions can be sensitive to some patients.**

| CRA | Quote |
|---|---|
| A | There's also certain parts of the questionnaire that are commonly left out, specifically breast cancer uhm patients whether, about their sexuality. We have some patients who just will skip that section altogether. They'll have done that for the entire time they were on a trial. They just will have no interest in sharing that information with us. |
| B | You have mastectomy or lumpectomy, it changes your appearance, your body image, how you look at yourself. A lot of those questions are tapping into that, asking you "do you feel as feminine as?", "how do you feel attractive?", asking them about their sex life. I just think those questions are just terrible. A lot of patients that I would administer them they just hated doing them for that reason. One lady she's like "what kind of pervert wants to know this information?" (laughs) It's very intrusive . . . I had a patient who, she absolutely refused to answer any of those questions. She'd answer all the other ones . . . it was just this one section when it got into body appearance, how many times you're having intercourse, with or without a partner . . . I wouldn't feel comfortable answering those questions either. That would be a reason why you'd be missing that particular data was patients just don't feel comfortable. |
| C | I suppose an example would be males, when it comes to impotence. Some of the questionnaires are asking those questions, their sexual behaviours. |
| D | Some questions can be sensitive and personal, some people aren't comfortable divulging that, especially prostrate patients, their sexual function, sometimes they're embarrassed by it, they don't want anybody to know that they're struggling. One lady, an ovarian study, always refused to answer one question. She'd say to me "I don't know how to answer this. I don't know what they're asking". The question was "does your treatment make you feel less of a woman?" |
| E | There's no particular questions they miss unless it's a question that sexuality don't like to answer that. Don't need to, can leave that blank if they don't want to answer it . . . Men don't seem to care at all about those questions. I find men more open about that than women when they ask the certain questions about their sex life or whatever. |
| F | They will think twice if they want to talk about it if it's cancer of the prostrate, "do you feel fulfilled sexually" . . . He was married. He says "so uhm if I don't have to answer that question I won't". He didn't answer. |
| H | . . . if they're satisfied sexually and things like that. Some patients do not want to answer these questions. |
| I | There's some questions that of course they don't like answering . . . |

**Table 5-10. CRAs have different experiences regarding whether age is an influence.**

| CRA | Quote |
|---|---|
| C | We find that younger patients, if they're on an oral medication, are a little less compliant than older people (laughs). |
| D | A young cervix cancer patient . . .  she was okay answering the very personal questions . . . whether that has to do with the age difference. Prostrate patients are generally older. A young woman, I don't know if that mattered, it could. |
| E | I didn't find a trend, why people miss questions or miss signing, unless they're really elderly . . . now I would say probably 80 and over. |
| F | My patients are over 60 65. A lot of them are diabetic. They have that blurred vision. There's a lot of different factors. |
| H | The seventies are new fifty because [pause] the population [pause] is a bit uhm is better fit. Even at seventy years old we have patient you would never think are seventy [pause], they look very fit, they look in shape. It really depends on the condition of the person. Uhm, right now, no – compare my older and younger patients [pause] it's about the same. |
| I | When it comes to the paper version that we've been doing it doesn't really matter. |
| J | Population is elderly patients. The average is approximately like seventy years old. Anything younger than that is actually a young patient for us. (laughs) Working with many eighty year old patients, sometimes is difficult to complete the questionnaire for them, struggling to just [pause] not necessarily understand the question but physically just doing the questionnaire. |

**Table 5-11. Being aware of the patient's disease progression may evoke an emotional reaction in the CRA that hinders them from administering the instrument.**

| CRA | Quote |
|---|---|
| A | I sometimes will even gauge going in there. If I went in and she was weepy I wouldn't even bring my tablet . . . |
| B | The physician goes in and says "unfortunately your cancer is progressed and we're going to take you off of this treatment", talks about other treatment options, or there is no treatment options. Obviously, it's very emotional. I'm to come in and ask them to do a discontinuation e-PRO at that point. Sometimes they handle that information very well. I've had patients too upset. I feel like a bit of a jerk, too. "Hey I'm really sorry yeah and you're off the study and da da duh but the study would really like me to have you do this questionnaire how do you feel about that?" They're like "no". |
| C | . . . they progressed, they're not feeling well, they may not want to complete the QofLs. If you look at a lot of the studies, once they do progress, QofLs aren't needed any more. Perhaps that's why, because would I want to chase a patient for that? That would be the last thing I would probably want to chase them for. At that point that's not important to them. It is to the study, but not to the individual. |
| D | . . . umhum [spoken very quietly] yeah I'm sure I have [long pause of three seconds] it is difficult. "I I understand this is a a hard time, you're able to refuse any study procedures you're not comfortable with, but would you complete this today or" |
| E | Oh, I get very emotional with them, too. It bothers me when they're upset. |
| F | It doesn't. [Response to How does that affect you?] I mustn't felt very good because I asked him if it's okay [pause] if we could do it together. I must not have been satisfied that he didn't do it at all because I wanted really to know how he felt. |
| G | We don't even [pause] uhm [pause] typically we wouldn't bother them with that [spoken very slowly] [spoken very quietly] at that point yeah [spoken very quietly]. |
| I | It's very [pause] very sad. It's [pause] not a great [pause] great position to be in. It makes me s(ad) yeah it makes me kind of [pause] you know sad as well. I approached him with it. I tell them like "I really understand if you don't want to do this" so [pause] "I can I can always [pause] see you the next time" or something like that [spoken very quietly] I think they did end up doing the next one. |

**Table 5-12. The patient's personality influences whether the instrument is completed at progression.**

| CRA | Quote |
|---|---|
| B | I think it's just personality. I just feel like some people have a different handle on their cancer, their diagnosis, how they accept it. It's a difference in personality I think . . . I recently had another patient, exact same design as that trial but a different type of lung cancer . . . progression . . . found to be on the immunotherapy, on the clinical trial drug . . . He handled it very very well . . . He was obviously upset at the news (but) he didn't baulk at doing the discontinuation questionnaire at all. |
| C | It would depend probably on the patient themselves. Some of them would really want to do it. I think that all has to do with personality and how they live their lives. If they're committed to something they're committed to something . . . I think it's all personality dependent. I really do. |
| D | Is it your upbringing, background, general beliefs? I think it all has to do with it. |
| E | Depends on the shape of the person . . . some patients are a lot more stressed out than other patients, just dealing with [long pause of three seconds] the whole experience with cancer. |
| F | I'll give you a a solid experience that I've had. He fell under standard arm, didn't fall on the treatment arm which was immunotherapy. We offered him the quality of questionnaire on cycle one day one. He developed severe toxicities from the regular chemo: low platelets, absolute neutrophil count went down, he had **no** energy. After our second or third attempt to try to make sure that his uhm labs would resume so he could continue he ended up having a global deterioration and felt worse and worse. I went up to the department to see the patient. I asked him. I'll leave this with you so you can complete it when you have a chance. I went back upstairs. It's not that he didn't want to he didn't have the energy, was quite weak. I said "you know if you want I can ask you the questions and you'll just answer the question and give me the answers". He was quite advanced. He was trying his last chance at something. He completed it, made me repeat questions sometimes. He just didn't want to have to [pause] go ahead and do it himself. I didn't mind to have to uhm to read it to him. |
| H | We can also offer if they want to meet with the doctor to discuss continuation or not before we do anything. The subject we went to see the doctor said "listen they just got a bad news you know, they they were called for this appointment [pause] uhm and were we were not made aware of it, so they were seen before us". He came back and completed it and continued on the study. |
| I | There's certain people for sure who are more likely to complete them. I think it depends on the patient. Most people [pause] are okay with doing the questionnaires. |

## 5.8 References


1. Palmer MJ, Mercieca-Bebber R, King M, Calvert M, Richardson H, Brundage M. A systematic review and development of a classification framework for factors associated with missing patient-reported outcome data. *Clinical Trials* 2018;15:95-106.
2. Osoba D, Bezjak A, Brundage M, Pater J. Evaluating health-related quality of life in cancer clinical trials: the National Cancer Institute of Canada Clinical Trials Group experience. *Value Health* 2007;10 Suppl 2:S138-45.
3. Morse JM, Barrett M, Mayan M, Olson K, Spiers J. Verification strategies for establishing reliability and validity in qualitative research. *Int J Qual Methods* 2002;1:13-22.
4. Morse JM. Critical analysis of strategies for determining rigor in qualitative inquiry. *Qual Health Res* 2015;25:1212-1222.
5. Saumure K, Given LM. Rigor of qualitative research. *The SAGE encyclopedia of qualitative research methods.* SAGE Publications: Thousand Oaks, 2008.
6. Creswell JW, Hanson WE, Plano Clark VL, Morales A. Qualitative research designs: selection and implementation. *The Counseling Psychologist* 2007;35:236-264.
7. Husserl E. *Ideas: General introduction to pure phenomenology* (translator, Gibson WRB). Humanities Press: New York, 1931(Original work published 1913).
8. Caelli K. The changing face of phenomenological research: traditional and American phenomenology in nursing. *Qual Health Res* 2000;10:366-377.
9. Dowling M, Cooney A. Research approaches related to phenomenology: negotiating a complex landscape. *Nurse Res* 2012;20:21-27.
10. Brinkmann S and Kvale S. *InterViews. Learning the craft of qualitative research interviewing, third edition.* SAGE Publications: Thousand Oaks, 2015.

11. McIntosh MJ, Morse JM. Situating and constructing diversity in semi-structured interviews. *Glob Qual Nurs Res* 2015 doi:10.1177/2333393615597674.
12. Morse JM. Analytic strategies and sample size. *Qual Health Res* 2015;25:1317-1318.
13. Patton MQ. *Qualitative research & evaluation methods, 3rd edition.* SAGE Publications: Thousand Oaks, 2002.
14. Morse JM. All data are not equal. *Qual Health Res* 2015;25:1169-1170.
15. Moustakas, C. *Phenomenological research methods.* Thousand Oaks: Sage Publications, 1994.
16. Gearing RE. Bracketing. *The SAGE encyclopedia of qualitative research methods*. SAGE Publications: Thousand Oaks, 2008.
17. Morris PL. Triangulation. *The SAGE encyclopedia of communication research methods*. SAGE Publications: Thousand Oaks, 2018.
18. Bruera E, Kuehn N, Miller MJ, Selmser P, Macmillan K. The Edmonton System Assessment System (ESAS): A simple method for the assessment of palliative care patients. *J Palliat Care* 1991;7:6-9.
19. Brundage MD, Barbera L, McCallum F, Howell DM. A pilot evaluation of the expanded prostate cancer index composite for clinical practice (EPIC-CP) tool in Ontario. *Qual Life Res* 2019;28:771-782.
20. Morse JM. "Data were saturated . . ." *Qual Health Res* 2015;25:587-588.
21. Rico-Villademoros F, Hernando T, Sanz J-L, et al. The role of the clinical research coordinator – data manager – in oncology clinical trials. *BMC Med Res Methodol* 2004;4:6.
22. Sherman DW, McSherry CB, Parkas V, et al. Recruitment and retention in a longitudinal palliative care study. *Appl Nurs Res* 2005;18:167-177.

23. Kyte D, Ives J, Draper H, Keeley T, Calvert M. Inconsistencies in quality of life data collection in clinical trials: a potential source of bias? Interviews with research nurses and trialists. *PLOS One* 2013;8:e76625.

24. Mercieca-Bebber R, Calvert M, Kyte M, Stockler M, King MT. The administration of patient-reported outcome questionnaires in cancer trials: interviews with trial coordinators regarding their roles, experiences, challenges and training. *Contemp Clin Trials Commun* 2018;9:23-32.

# Chapter 6

# General Discussion

## 6.1 Overview

The goal of the thesis was to use quantitative and qualitative approaches to identify the causes of missing PRO data. The dissertation consists of four parts. Part 1 identified factors reported to be associated with missing PRO data. Part 2 examined some of the factors in secondary analyses of quantitative data. Part 3 identified influences of missing PRO data from primary collection and analysis of qualitative data. Part 4 consolidated results from Part 3 with results from Part 2.

## 6.2 Summary of the Main Findings

The first manuscript is *"A systematic review and development of a classification framework for factors associated with missing patient-reported outcome data"*. Two databases were systematically searched for factors reported to have association with missing PRO data. Forty-six unique categories of factors were identified. These factors were organized into a framework having five main components: Instrument (e.g. length), Participant (e.g. sociodemographic characteristics), Centre (e.g. research infrastructure), Staff (e.g. error in administering the instrument), and Study (e.g. type of treatment). The Classification Framework is an inventory of factors reported in the literature to be associated with missing PRO data. It may assist researchers to improve the design of new RCTs and to implement procedures to reduce missing PRO data.

The second manuscript is *"Factors predicting missing instruments in three cancer randomized clinical trials"*. Cox regression was used to analyze individual participant data from three RCTs: MA17 (breast cancer), PR7 (prostate cancer), and LY12 (non-Hodgkin's lymphoma). The outcome was time from randomization to the first missing instrument. Variables for 15 factors were used as covariates, selected based on availability and previously-reported putative associations with missing PRO data. Statistically significant independent associations with outcome were found for: baseline age (PR7, HR = 0.97); level of well-being (LY12, HR = 0.97); centre level of activity (LY12, HR = 0.99); presence of post-graduate residency training program (MA17, HR = 1.66; PR7, HR = 1.81); and centre geographic location (PR7, HR = 1.20; LY12, HR = 1.61, for Ontario vs. West). Many factors reported in the literature to have association with missing PRO data were not significantly associated in these three RCTs with time to the first missing instrument: employment status, insurance, stage, race/ethnicity, gender, caregiver, cognitive status, physical impairment, toxicity, type of treatment.

The third manuscript is *"What is the lived experience of clinical research associates with missing patient-reported outcome data?"* In-depth semi-structured interviews were conducted with a purposeful sample of CRAs. Ten themes were identified in analysis based on descriptive phenomenology. The CRAs described their experience with different patterns of missing PRO data. PROs for routine clinical care compete with PROs for RCTs. Several important influences of missing PRO data were inherent to the instruments (quantity, repetition, wording of questions, and format), patients (characteristics, attributes of the cancer, and toxicity from the treatment), centres (deficiencies in research infrastructure), staff (developing relationships with patients), and RCTs (data capture in electronic versus paper format). Also, adapting the assessment schedule and providing sufficient time for the patient to complete the instrument were challenging for the CRAs, particularly at disease progression. There is a need for field guidance, and a motivation and knowledge base among CRAs to develop such guidance.

## 6.3 Meeting the Thesis Objectives

### 6.3.1 Part 1

The objectives for Part 1 were to perform a systematic review of the literature and to create a Classification Framework for factors reported to be associated with missing PRO data. The PhD candidate (MJP) achieved these objectives in the first manuscript. MJP independently screened 6027 records identified from a systematic search of the MEDLINE and CINAHL databases using search terms. An independent screening of the same search results was done by a second reviewer in Australia. Consensus was used to identify a final set of papers. One hundred articles from 70 journals [Appendix E.1] met the eligibility criteria reflecting: relevance to PROs, missing data in prospective studies (including RCTs), and examination of predictors or causes of missing PRO data. From these papers, MJP extracted 663 excerpts of text that described factors associated with missing PRO data. Through an iterative process of consensus with his supervisors, he reduced the number of excerpts to a list of unique factors, and organized the list in a Classification Framework. The extent to which each factor predicts missing instruments was largely unknown, in part because few factors were derived from regression models.

The search criteria for the systematic review were unrestricted with respect to disease. In all, 57% of the 100 eligible papers had focus on cancer, 23% did not specify disease, and the remaining 20% reported for a variety of health conditions. In order to justify proceeding to Part 2 of the thesis, MJP first examined the evidence for an association between one factor – patient age – and missing PRO data on 23 studies identified in the systematic review [Appendix B]. There was variation in: context, analytical approaches, and in the conceptualization of age and of missing PRO data. A decision was made to focus on a series of RCTs in cancer for Part 2, as this approach would provide robustness of study design.

### 6.3.2 Part 2

The primary objective for Part 2 was to determine which characteristics of the instruments, participants, centres, and trials predict time to the first missing instrument after randomization. MJP achieved this objective in the second manuscript. Quantitative regression analyses examined the association of 11 Participant factors, three Centre factors, and one Study factor with time to the first missing instrument in three RCTs. It was not possible to examine Instrument factors due to collinearity with the RCTs. The results contribute new evidence that subjects from centres with a post-graduate residency training program had shorter time to the first missing instrument on MA17 and on PR7. The results contribute additional quantitative evidence to the existing literature for four factors: baseline age, level of well-being, level of activity, and geographic location. The results do not provide evidence of a significant association for ten factors: employment status, insurance, stage, race/ethnicity, gender, caregiver, cognitive status, physical impairment, toxicity, type of treatment.

### 6.3.3 Part 3

The objective for Part 3 was to describe the lived experience of CRAs with missing PRO data. MJP achieved this objective in the third manuscript. The CRAs descriptions revealed that differences – both within and across the instruments, patients, centres, staff, and RCTs – may influence missing PRO data. The results contribute new evidence that PROs for routine clinical care compete with PROs for RCTs. The results contribute additional qualitative evidence to the existing literature for several factors involving patient characteristics, instruments, research infrastructure, and data capture format.

### 6.3.4 Part 4

The primary objective for Part 4 was to consolidate the qualitative and quantitative results. Section 6.4 describes how MJP applied methodological triangulation to consolidate the qualitative results with the quantitative results. Section 6.5 highlights the key findings.

## 6.4 Application of Methodological Triangulation

The purpose for mixing methods is to integrate information from different sources through convergent validation [1]. Consolidation was achieved through methodological triangulation [2], which is important for two reasons: to maximize understanding from the available evidence of missing PRO data, and to enhance the rigour of the mixed-methods design [3]. For every factor, the qualitative result was compared with the quantitative result. Each factor was then assigned to one of three categories: convergent (if results were consistent), divergent (if results were inconsistent), or distinct (if the factor only appeared in results for one method).

## 6.5 Findings from Consolidation of the Qualitative and Quantitative Results

The results for post-graduate residency training program and age were convergent (Table 6.1). The presence of a post-graduate residency training program may lead to a shorter time to the missing instrument as the patient has insufficient time to complete the instrument in clinic because of interruptions, particularly by trainees and students involved in the circle of care. Subjects with older age may have longer time to first missing instrument on some but not all RCTs. For example, CRA F described many 65-year old subjects with blurred vision, and CRA J described many 80-year old subjects with physical limitations, having difficulty completing the instrument (Table 5-10). Here, the qualitative results explain the quantitative results.

Nine factors were divergent (Table 6-1). The qualitative results provide insight on one factor that was significant (level of activity), and eight factors that were not significant (employment status, stage, gender, caregiver, cognitive status, physical impairment, toxicity, and type of treatment). Here, the qualitative results provide evidence on how each factor may lead to missing PRO data that was not apparent from the quantitative results.

The remaining factors were distinct. Of these, four factors only appeared in the quantitative results, and 49 factors only appeared in the qualitative results.

The quantitative results identified two factors that were significant (level of well-being and location) and two factors that were not significant (race/ethnicity and insurance). The last two factors may not have been significant because of limited variability in these covariates for subjects from Canada, since the vast majority were Caucasian and had health insurance.

The qualitative results contributed breadth and depth to a greater understanding of factors influencing missing PRO data. This point is illustrated with two examples. The finding that 'PROs for routine clinical care compete with PROs for RCTs' highlights the importance of new factors external to the RCTs being influences of missing PRO data within RCTs. The finding that 'Assessment of PROs at disease progression is difficult' highlights the importance of appreciating CRAs and cancer patients are humans not just sources of data who may experience an emotional reaction that hinders them from administering or completing the instrument.

The results are intriguing and add to our understanding of missing PRO data but are not definitive. The thesis committee recommended MJP not proceed with revising the Classification

Framework (as per secondary objective in Part 4) based on these results because it would be premature to do so. Additional qualitative research would help increase understanding. Involving a panel of experts in the review of content would provide robustness to inform a revision.

6.6 Strengths

Each manuscript includes strengths that are relevant to each objective. This section describes strengths of the entire dissertation.

A pragmatic approach was used to understand a common but complex problem from different sources of information. Recommended strategies were applied in each Part to enhance rigour. The quantitative component informed the design of the qualitative component. CRAs were purposefully sampled based on the geographic location and the presence or absence of a post-graduate residency training program in oncology at the centre. In addition, the qualitative component included CRAs with experience working on RCTs for the quantitative component, thereby providing linkage between the two sources of data in the mixed-methods design.

The quantitative results added value to what was learned from the qualitative results. The quantitative analyses examined the independent association of specific factors with time to the first missing instrument in three RCTs. The semi-structured interviews with CRAs did not examine *a priori* the influence of specific factors on missing PRO data. Rather, CRAs were encouraged to describe their personal experience with important influences.

Cox proportional hazards regression obtained precise HRs and narrow 95% CI. The HR is familiar to clinicians. The HR provides information on the magnitude and the direction. The 95%

CI indicates statistical significance at $p < 0.05$ if the upper tail is below one or the lower tail is above one. Knowledge of these properties can guide comparisons. For example, consider the results for presence of a post-graduate residency training program on MA17 (HR = 1.66, 95% CI 1.06 – 2.60) and for the level of well-being on LY12 (HR = 0.97, 95% CI 0.95 – 0.99). Both are statistically significant, but a stronger association with time to the first missing instrument was observed for presence of a post-graduate residency training program than with the level of well-being. The directions of association are different: the HR above one represents shorter time to the first missing instrument, while the HR below one represents longer time.

This dissertation demonstrates the benefit of the mixed-methods design. In particular, CRAs have collective expertise and qualitative inquiry of their lived experience revealed important influences of missing PRO data that were not apparent from quantitative analyses of RCT data.

## 6.7 Limitations

This section describes the effect of bias and confounding on the internal validity of the dissertation, effect modification, and the external validity of the dissertation.

### 6.7.1 Publication Bias

Publication bias may have been present in Part 1, but was not present in Part 2 or in Part 3 because the sources of data did not involve publications. The search criteria restricted the systematic review to English-language records and these may be susceptible to publication bias. However, prior to 1998, in general each paper contributed one factor, nine papers from the international workshop on missing PRO data contributed 27 factors in 1998, and subsequently, the yield diminished with no new factors identified since 2012 (Figure 3-2). So, although non-

English records may contain factors reported to be associated with missing PRO data, it is very unlikely that these would contribute new information to the unique factors that were obtained from the English records and were included in the Classification Framework. Publication bias is considered to have minimal effect on the internal validity.

6.7.2 Selection Bias

Selection bias is not a concern for Part 1 because two reviewers independently screened all records identified through database searching against *a priori* eligibility criteria. Selection bias may have been present in Part 2 but is not a concern. The 23 RCTs that were eligible for selection were stratified according to treatment intent (Figure 4-1) as palliative (n = 13 RCTs, including treatment for an incurable metastatic cancer) or curative (n = 10 RCTs, including adjuvant treatment for a potentially curative cancer). The initial plan was to select all RCTs having curative treatment but, after discussion with the thesis committee, it was decided to select RCTs with variety in context and large sample size: first from the breast disease site (MA17), then from the genito-urinary disease site (PR7), and lastly from the hematologic disease site (LY12). The selection was not influenced by the extent of missing PRO data or the characteristics of participants and centres, because this information was only apparent after the data had been provided. Selection bias may have been present in Part 3 but is not a concern. Maximum variation sampling was used to select CRAs on diversity of geographic location (within the western provinces, Ontario, Quebec, or eastern provinces of Canada) and of post-graduate residency training program in oncology (present or absent). This approach identified CRAs having a range of experience with respect to missing PRO data for interview, thereby reducing vulnerability to selection bias [4].

6.7.3 Information Bias

Information bias may have been present in Part 1, in Part 2, and in Part 3. Misclassification of one or more of the 46 categories of factors within 5 components for the Classification Framework (Part 1) was unlikely because an iterative process of consensus involving MJP and his supervisors guided placement of content. Misclassification of one or more of the 60 clusters within 10 themes (Part 3) was unlikely because an iterative process of consensus involving MJP, his supervisors, and the qualitative expert on the thesis committee (TK) guided placement of content (Appendix E.3, Face-validity Exercise). Misclassification of one or more of the 54 MA17 centres, 30 PR7 centres, or 26 LY12 centres for post-graduate residency training program (Part 2) is a potential threat to internal validity. However, most LY12 centres were also on PR7 and were also on MA17. Non-differential misclassification on MA17 may have biased the HR towards a null value but because coding was consistent the magnitude of bias should be similar on each RCT, so that comparisons of HRs between RCTs were unbiased.

6.7.4 Confounding

Confounding was a potential concern for Part 2 only. The aim was to evaluate the independent association between factors identified in Part 1 and missing PRO data in multivariate regression analyses. However, there was a limited selection of variables available from the three databases because the RCTs were not specifically designed to evaluate missing PRO data. Consequently, only 15 (of 46) factors identified in Part 1 were examined in Part 2, and final multivariate models were limited to available covariates. In general, the univariate HRs did not change substantially (i.e. > 10%) after adjustment for other covariates in the multivariate models. The results in Part 2 were susceptible to unmeasured confounding from other factors (e.g., procedural knowledge of staff, for factor age) which may have distorted the HRs. The extent and direction is not known.

6.7.5 Effect Modification

The Classification Framework (Part 1) provides information on factors reported in the literature to have association with missing PRO data, but does not indicate whether there is modification of effect between factors with missing PRO data. There was *a priori* intention to examine potential effect modification involving age, gender, and type of cancer in Part 2. Collinearity of type of cancer with the three RCTs, and collinearity of gender with two of the three RCTs, precluded examination of effect modification in MA17 (all subjects were female) and in PR7 (all subjects were male). In LY12, univariate Cox regression revealed non-significant effect modification between age and gender with time to the first missing instrument ($p = 0.416$). In Part 3, some CRAs described effect modification involving age, gender, and type of cancer. For example, CRA D said: "Older patients can't grasp it. You've got to look at your demographics. It's not so good for older ovarian cancer ladies, prostate men" (Table 5-6). The dissertation provides some insights about potential effect modification but the paucity of data is recognized as a limitation.

6.7.6 External Validity

The results of the dissertation should be interpreted with an understanding that the Classification Framework was based on international data reported in the literature, the quantitative analyses were based on data provided by CCTG for subjects in Canada with curative breast cancer, prostate cancer, or non-Hodgkin's lymphoma, and the qualitative analysis was based on data provided by CRAs having experience working at CCTG centres for subjects in Canada with different cancers. The results may not be generalizable to RCTs: (a) in patients with other types of cancer, (b) conducted outside of Canada, and/or (c) in settings not associated with a cooperative oncology group. The external validity would improve with a survey of CRAs in

Canada (and by extension, United Kingdom and Australia) regarding their level of endorsement. Surveying oncologists would also be informative.

Although 20% of the 100 eligible papers from the systematic review reported for patients with a variety of non-cancer health conditions (refer to 3.4 Results for list), quantitative analyses did not examine factors from the Classification Framework in RCTs of these patients, nor did qualitative inquiry extend to CRAs involved in their care. Whether the results are generalizable beyond cancer is not known. The results may be generalizable to other chronic diseases, if the trajectory and clinical setting are similar to cancer (e.g., assessment of PROs at progression), but caution should be exercised in the absence of evidence.

### 6.8 Contributions of this Research

This dissertation contributes to international efforts providing guidance for improving the collection, analysis, reporting, and presentation of PROs. Collectively, the SPIRIT-PRO [5, Appendix G], SISAQOL [6], CONSORT-PRO [7], and guidance on data presentation formats [8], are a robust set of documents that can help researchers with writing the protocol, analysing the PRO data, reporting the PRO results, and graphically displaying the PRO findings in clinical practice. But, despite careful study design and analysis plans, missing PRO data can weaken the internal validity of the RCT. A symposium at ISOQOL highlighted how different approaches could be used to maximise the impact of PRO results from RCTs [9]. Missing PRO data can be minimized by understanding independent factors that influence missing PRO data and applying complementary strategies to reduce negative consequences [10, Appendix H; 11, Appendix F].

## 6.9 <u>Conclusions and Future Directions</u>

This dissertation provides evidence to facilitate understanding of which factors influence missing PRO data, so steps can be taken to improve the quality of PROs by minimizing missing data.

Original data collected from cancer patients who participated on an RCT would provide a perspective on missing PRO data. CRAs, in a sense, were a proxy who described patient-specific issues across their experience with a wide variety of patients, but, of course, the perspective of patients is best revealed from patients themselves. Interviews with patients might identify new influences, for example, their experience with anxiety, frustration or burn-out during the course of the RCT, but one would not anticipate substantially different results from what was obtained in interviews with CRAs. The current evidence would benefit from future research with patients whose contribution could supplement the results and conclusions from this dissertation.

Conditional on obtaining suitable funding, future work should be directed towards: developing field guidance, and conducting additional qualitative inquiry with a methodology other than descriptive phenomenology. These are briefly described in the following paragraphs.

The qualitative component identified a need for field guidance, and a motivation and knowledge base among CRAs to develop such guidance. The CRAs have collective wisdom and it would make sense to involve them in developing field guidance to include which factors influence missing PRO data and their suggestions regarding what can be done to minimize missing PRO data. The format should contain an educational component and strategies for dealing with common scenarios (e.g., older patients, PROs for routine clinical care compete with PROs for RCTs).

Grounded theory would be the appropriate choice for a new study investigating how different factors act to create missing PRO data. The theory of planned behaviour [12, 13] is useful for understanding the dynamics of human behaviours in relation to intended behaviours. According to the theory demographic characteristics, personality traits, and the environment influence the individual's behavioural, normative, and control beliefs. These determine attitudes, subjective norms, and perceptions of behavioural control, which feed into the individual's intention to engage in the behaviour. The theory illustrates the paradoxical complexity of missing PRO data that, on the surface, is straightforward. Its application to interviews of patients would identify how different factors influence an individual's decision to complete a PRO instrument.

Ethnography would be the appropriate choice for another study investigating how the sponsor may influence missing PRO data. For example, the culture of personnel within the CCTG Statistical and Operations Office, having involvement with the collection of PRO data on specific RCTs, may contribute to reducing missing PRO data through educational initiatives and monitoring of data flow. The effect of these efforts is not known. Interviews with personnel could supplement data from observing the work environment, and data from reviewing standard operating procedures, protocols, and other documents.

Table 6-1. Application of methodological triangulation.

| Factor | Result from the quantitative analysis | Result from the qualitative analysis | Comparison of results between the two methods | Assignment to category |
|---|---|---|---|---|
| post-graduate residency training program | In the final multivariate models for MA17 and for PR7, subjects from centres with a post-graduate residency training program had an increase in the relative risk (HR = 1.66, 95% CI 1.06 – 2.60; HR = 1.81, 95% CI 1.31 – 2.51). In regression modelling for subjects on LY12, significant associations were not observed with missing instruments. | The patient has insufficient time to complete the instrument in the clinic because of interruptions. | The results are consistent. | Convergent |
| age | In the final multivariate model for PR7, subjects with a one-year increase in age had a three percent decrease in the relative risk (HR = 0.97, 95% CI 0.95 – 0.99). In regression modelling for subjects on MA17 and for subjects on LY12, significant associations were not observed with missing instruments. | CRAs have different experiences regarding whether age is an influence. | The results are consistent. | Convergent |
| level of activity | In the final multivariate model for LY12, subjects from centres with a one-unit increase in accrual had a one percent decrease in the relative risk (HR = 0.99, 95% CI = 0.99, 1.00). | The CRA is busy and is not able to administer the instrument. | The results are inconsistent. | Divergent |
| employment status | LY12 not significant | Retired people are often more willing to complete an instrument. | The results are inconsistent. | Divergent |

| | | | | |
|---|---|---|---|---|
| stage | MA17, LY12<br>not significant | Patients on adjuvant therapy may be more likely to complete the instrument but miss some of the sections. | The results are inconsistent. | Divergent |
| gender | LY12 not significant | Women are typically more likely to skip certain sections whereas men will often miss a page. | The results are inconsistent. | Divergent |
| caregiver | LY12 not significant | Some patients have difficulty reading and require assistance to complete the instrument. A caregiver can help, by reading the questions to the patient and recording the answers. | The results are inconsistent. | Divergent |
| cognitive status | PR7 not significant | Patients with a diagnosis of brain cancer have more difficulty completing the instrument. | The results are inconsistent. | Divergent |
| physical impairment | MA17, PR7, LY12<br>not significant | The patient’s deteriorating health may make them unable to attend the clinic. | The results are inconsistent. | Divergent |
| toxicity | MA17, PR7<br>not significant | Some patients are fatigued and don’t want to complete an instrument. | The results are inconsistent. | Divergent |
| type of treatment | MA17, PR7, LY12<br>not significant | Some patients develop a mental fog from chemotherapy which can compromise their ability to complete an instrument. | The results are inconsistent. | Divergent |

## 6.10 References

[1] Fielding NG. Triangulation and mixed methods designs: data integration with new research technologies. *J Mix Methods Res* 2012;6:124-136.

[2] Morris PM. Triangulation. *The SAGE Encyclopedia of Communication Research Methods.* SAGE Publications: Thousand Oaks, 2018.

[3] Heale R, Forbes D. Understanding triangulation in research. *Evid Based Nurs* 2013 doi:10.1136/eb-2013-101494.

[4] Collier D, Mahoney J. Insights and pitfalls: selection bias in qualitative research. *World Politics* 1996;49:56-91.

[5] Calvert M, Kyte D, Mercieca-Bebber R, et al. Guidelines for inclusion of patient-reported outcomes in clinical trial protocols: the SPIRIT-PRO extension. *JAMA* 2018;319:483-494.

[6] Bottomley A, Pe M, Sloan J et al. Moving forward toward standardizing analysis of quality of life data in randomized clinical trials. *Clinical Trials* 2018;15:624-630.

[7] Calvert M et al. Reporting of patient-reported outcomes in randomized trials: the CONSORT PRO extension. *JAMA* 2013;309:814-822.

[8] Snyder C, Smith K, Holzner B, Rivera YM, Bantug E, Brundage M. Making a picture worth a thousand numbers: recommendations for graphically displaying patient-reported outcomes data. *Qual Life Res* 2019;28:345-356.

[9] Calvert M et al. Symposium #1: Minimising research waste and maximising the impact of patient reported outcome trial results. *Qual Life Res* 2018;27(Suppl 1):S3.

[10] Mercieca-Bebber R, Palmer MJ, Brundage M, Calvert M, Stockler MR, King MT. Missing patient-reported outcome data: understanding associated factors and strategies to reduce its instance and impact through careful design, quality assurance and reporting. *Qual Life Res* 2018;27(Suppl 1):S4.

[11] 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:e010938.

[12] Ajzen I. The theory of planned behaviour: reactions and reflections. *Psychol Health* 2011;26(9):1113-1127.

[13] Conner M, Sparks P. Theory of planned behaviour and health behaviour. Chapter 5 in, *Predicting Health Behaviour: Research and Practice with Social Cognition Models, second edition.* Edited by Conner M and Norman P. Open University Press: Berkshire England, 2005.

# Appendix A

## Research Ethics Board Approvals

## HSREB Approval for Focus Group on May 2, 2015

QUEEN'S UNIVERSITY HEALTH SCIENCES & AFFILIATED TEACHING HOSPITALS RESEARCH ETHICS BOARD-DELEGATED REVIEW
April 10, 2015

Mr. Michael Palmer
Department of Public Health Sciences
Queen's University

Dear Mr. Palmer
**Study Title: EPID-505-15 Improving the quality of patient reported outcomes by minimizing missing data**
**File # 6015243**
**Co-Investigators: Dr. H. Richardson, Dr. M. Brundage**

I am writing to acknowledge receipt of your recent ethics submission. We have examined the protocol, interview guide (2015Mar30) and revised information/consent form for your project (as stated above) and consider it to be ethically acceptable. This approval is valid for one year from the date of the Chair's signature below. This approval will be reported to the Research Ethics Board. Please attend carefully to the following listing of ethics requirements you must fulfill over the course of your study:

**Reporting of Amendments**: If there are any changes to your study (e.g. consent, protocol, study procedures, etc.), you must submit an amendment to the Research Ethics Board for approval. Please use event form: HSREB Multi-Use Amendment/Full Board Renewal Form associated with your post review file # **6015243** in your Researcher Portal (https://eservices.queensu.ca/romeo_researcher/)

**Reporting of Serious Adverse Events**: Any unexpected serious adverse event occurring locally must be reported within 2 working days or earlier if required by the study sponsor. All other serious adverse events must be reported within 15 days after becoming aware of the information. Serious Adverse Event forms are located with your post-review file **6015243** in your Researcher Portal (https://eservices.queensu.ca/romeo_researcher/)

**Reporting of Complaints**: Any complaints made by participants or persons acting on behalf of participants must be reported to the Research Ethics Board within 7 days of becoming aware of the complaint. Note: All documents supplied to participants must have the contact information for the Research Ethics Board.

**Annual Renewal**: Prior to the expiration of your approval (which is one year from the date of the Chair's signature below), you will be reminded to submit your renewal form along with any new changes or amendments you wish to make to your study. If there have been no major changes to your protocol, your approval may be renewed for another year.

Yours sincerely,

Albert F. Clark

Chair, Health Sciences Research Ethics Board
April 10, 2015

**Investigators please note that if your trial is registered by the sponsor, you must take responsibility to ensure that the registration information is accurate and complete**

**QUEEN'S UNIVERSITY HEALTH SCIENCES & AFFILIATED TEACHING HOSPITALS RESEARCH ETHICS BOARD**

The membership of this Research Ethics Board complies with the membership requirements for Research Ethics Boards and operates in compliance with the Tri-Council Policy Statement; Part C Division 5 of the Food and Drug Regulations, OHRP, and U.S DHHS Code of Federal Regulations Title 45, Part 46 and carries out its functions in a manner consistent with Good Clinical Practices.

**Federalwide Assurance Number: #FWA00004184, #IRB00001173**

**Current 2015 membership of the Queen's University Health Sciences & Affiliated Teaching Hospitals Research Ethics Board:**

**Dr. A.F. Clark,** Emeritus Professor, Department of Biomedical and Molecular Sciences, Queen's University (Chair)

**Dr. H. Abdollah,** Professor, Department of Medicine, Queen's University

**Dr. R. Brison,** Professor, Department of Emergency Medicine, Queen's University

**Dr. D. Groll,** Assistant Professor, Department of Psychiatry, Queen's University

**Ms. J. Hudacin,** Community Member

**Mr. D. McNaughton**, Community Member

**Ms. S. Rohland,** Privacy Officer, ICES-Queen's Health Services Research Facility, Research Associate, Division of Cancer Care and Epidemiology, Queen's Cancer Research Institute

**Dr. M. Sawhney**, Assistant Professor, School of Nursing, Queen's University

**Ms. J. Sheldrick**, Community Member

**Dr. A. Singh**, Professor, Department of Psychiatry, Queen's University

**Dr. J. Thorpe,** Clinical Research Associate, Department of Anesthesiology and Perioperative Medicine, Queen's University

**Ms. K. Weisbaum**, LL.B. and Adjunct Instructor, Department of Family Medicine (Bioethics)

**Dr. J. Whiteley**, Community Member

## HSREB Renewal of Ethics Clearance

**QUEEN'S UNIVERSITY HEALTH SCIENCES & AFFILIATED TEACHING HOSPITALS RESEARCH ETHICS BOARD (HSREB)**

**HSREB Renewal of Ethics Clearance**

April 04, 2016

Mr. Michael Palmer
Department of Public Health Sciences
Cancer Research Institute

**ROMEO/TRAQ #: 6015243**
**Department Code: EPID-505-15**
**Study Title: Improving the quality of patient reported outcomes by minimizing missing data**
**Review Type: Delegated**
**Date Ethics Clearance Effective: April 10, 2016**
**Ethics Clearance Expiry Date: April 10, 2017**

Dear Mr. Palmer,

The Queen's University Health Sciences & Affiliated Teaching Hospitals Research Ethics Board (HSREB) has reviewed the application. This study, including all currently approved documentation has been granted ethical clearance until the expiry date noted above.

Prior to the expiration of your ethics clearance, you will be reminded to submit your renewal report through ROMEO. Any lapses in ethical clearance will be documented below.

Yours sincerely,

Albert F. Clark.

Chair, Health Sciences Research Ethics Board

*The HSREB operates in compliance with, and is constituted in accordance with, the requirements of the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (TCPS 2); the International Conference on Harmonisation Good Clinical Practice Consolidated Guideline (ICH GCP); Part C, Division 5 of the Food and Drug Regulations; Part 4 of the Natural Health Products Regulations; Part 3 of the Medical Devices Regulations, Canadian General Standards Board, and the provisions of the Ontario Personal Health Information Protection Act (PHIPA 2004) and its applicable regulations. The HSREB is qualified through the CTO REB Qualification Program and is registered with the U.S. Department of Health and Human Services (DHHS) Office for Human Research Protection (OHRP). Federalwide Assurance Number: FWA#:00004184, IRB#:00001173*

*HSREB members involved in the research project do not participate in the review, discussion, or decision.*

## HSREB Renewal of Ethics Clearance

**QUEEN'S UNIVERSITY HEALTH SCIENCES & AFFILIATED TEACHING HOSPITALS RESEARCH ETHICS BOARD (HSREB)**

**HSREB Renewal of Ethics Clearance**

March 13, 2017

Mr. Michael Palmer
Department of Public Health Sciences
Queen's University

**ROMEO/TRAQ #: 6015243**
**Department Code: EPID-505-15**
**Study Title: Improving the quality of patient reported outcomes by minimizing missing data**
**Review Type: Delegated**
**Date Ethics Clearance Effective: April 10, 2017**
**Ethics Clearance Expiry Date: April 10, 2018**

Dear Mr. Palmer,

The Queen's University Health Sciences & Affiliated Teaching Hospitals Research Ethics Board (HSREB) has reviewed the application. This study, including all currently approved documentation has been granted ethical clearance until the expiry date noted above.

Prior to the expiration of your ethics clearance, you will be reminded to submit your renewal report through ROMEO. Any lapses in ethical clearance will be documented below.

Yours sincerely,

Albert F. Clark.

Chair, Health Sciences Research Ethics Board

*The HSREB operates in compliance with, and is constituted in accordance with, the requirements of the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (TCPS 2); the International Conference on Harmonisation Good Clinical Practice Consolidated Guideline (ICH GCP); Part C, Division 5 of the Food and Drug Regulations; Part 4 of the Natural Health Products Regulations; Part 3 of the Medical Devices Regulations, Canadian General Standards Board, and the provisions of the Ontario Personal Health Information Protection Act (PHIPA 2004) and its applicable regulations. The HSREB is qualified through the CTO REB Qualification Program and is registered with the U.S. Department of Health and Human Services (DHHS) Office for Human Research Protection (OHRP). Federalwide Assurance Number: FWA#:00004184, IRB#:00001173*

*HSREB members involved in the research project do not participate in the review, discussion, or decision.*

**HSREB Amendment Acknowledgement / Ethics Clearance**

QUEEN'S UNIVERSITY HEALTH SCIENCES & AFFILIATED TEACHING HOSPITALS RESEARCH ETHICS BOARD (HSREB)

**HSREB Amendment Acknowledgment/Ethics Clearance**

March 28, 2017

Mr. Michael Palmer
Department of Public Health Sciences
Queen's University

**ROMEO/TRAQ #: 6015243**
**Department Code: EPID-505-15**
**Study Title: Improving the quality of patient reported outcomes by minimizing missing data**
**Review Type: Delegated**
**Date Ethics Clearance Issued: March 28, 2017**

Dear Mr. Palmer,

The Queen's University Health Sciences & Affiliated Teaching Hospitals Research Ethics Board (HSREB) has reviewed the amendment application and granted ethics approval/acknowledgement for the documents listed below.

- Interview Guide for Clinical Research Associates – 2017Mar02
- Informed Consent Form – 2017Mar02
- Informed Consent Form – 2017Mar20

Yours sincerely,

Albert F. Clark
Chair, Health Sciences Research Ethics Board

The HSREB operates in compliance with, and is constituted in accordance with, the requirements of the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (TCPS 2); the International Conference on Harmonisation Good Clinical Practice Consolidated Guideline (ICH GCP); Part C, Division 5 of the Food and Drug Regulations; Part 4 of the Natural Health Products Regulations; Part 3 of the Medical Devices Regulations, Canadian General Standards Board, and the provisions of the Ontario Personal Health Information Protection Act (PHIPA 2004) and its applicable regulations. The HSREB is qualified through the CTO REB Qualification Program and is registered with the U.S. Department of Health and Human Services (DHHS) Office for Human Research Protection (OHRP). Federalwide Assurance Number: FWA#:00004184, IRB#:00001173

*HSREB members involved in the research project do not participate in the review, discussion, or decision.*

## HSREB Renewal of Ethics Clearance

QUEEN'S UNIVERSITY HEALTH SCIENCES & AFFILIATED TEACHING HOSPITALS
RESEARCH ETHICS BOARD (HSREB)

**HSREB Renewal of Ethics Clearance**

April 03, 2018

Mr. Michael Palmer
Department of Public Health Sciences
Queen's University

**ROMEO/TRAQ #: 6015243**
**Department Code: EPID-505-15**
**Study Title: Improving the quality of patient reported outcomes by minimizing missing data**
**Review Type: Delegated**
**Date Ethics Clearance Effective: April 10, 2018**
**Ethics Clearance Expiry Date: April 10, 2019**

Dear Mr. Palmer,

The Queen's University Health Sciences & Affiliated Teaching Hospitals Research Ethics Board (HSREB) has reviewed the application. This study, including all currently approved documentation has been granted ethical clearance until the expiry date noted above.

Prior to the expiration of your ethics clearance, you will be reminded to submit your renewal report through ROMEO. Any lapses in ethical clearance will be documented below.

Yours sincerely,

Albert J. Clark

Chair, Health Sciences Research Ethics Board

*The HSREB operates in compliance with, and is constituted in accordance with, the requirements of the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (TCPS 2); the International Conference on Harmonisation Good Clinical Practice Consolidated Guideline (ICH GCP); Part C, Division 5 of the Food and Drug Regulations; Part 4 of the Natural Health Products Regulations; Part 3 of the Medical Devices Regulations, Canadian General Standards Board, and the provisions of the Ontario Personal Health Information Protection Act (PHIPA 2004) and its applicable regulations. The HSREB is qualified through the CTO REB Qualification Program and is registered with the U.S. Department of Health and Human Services (DHHS) Office for Human Research Protection (OHRP). Federalwide Assurance Number: FWA#:00004184, IRB#:00001173*

*HSREB members involved in the research project do not participate in the review, discussion, or decision.*

## HSREB Renewal of Ethics Clearance

QUEEN'S UNIVERSITY HEALTH SCIENCES & AFFILIATED TEACHING HOSPITALS RESEARCH ETHICS BOARD (HSREB)

**HSREB Renewal of Ethics Clearance**

March 28, 2019

Mr. Michael Palmer
Department of Public Health Sciences
Queen's University

**TRAQ #: 6015243**
**Department Code: EPID-505-15**
**Study Title: "Improving the quality of patient reported outcomes by minimizing missing data"**
**Review Type: Delegated**
**Date Ethics Clearance Effective: March 28, 2019**
**Ethics Clearance Expiry Date: March 28, 2020**

Dear Mr. Palmer:

The Queen's University Health Sciences & Affiliated Teaching Hospitals Research Ethics Board (HSREB) has reviewed the application. This study, including all currently approved documentation has been granted ethical clearance until the expiry date noted above.

Prior to the expiration of your ethics clearance, you will be reminded to submit your renewal report through TRAQ. Any lapses in ethical clearance will be documented below.

Lapses in Ethics Clearance: N/A

Yours sincerely,

Albert F. Clark

**Albert F. Clark, PhD**
**Chair, Queen's University Health Sciences and Affiliated Teaching Hospitals Ethics Board**

*The HSREB operates in compliance with, and is constituted in accordance with, he requirements of the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (TCPS 2); the international Conference on Harmonisation Good Clinical Practice Consolidated Guideline (ICH GCP); Part C, Division 5 of the Food and Drug Regulations; Part 4 of the Natural Health Product Regulations; Part 3 of the Medical Devices Regulations, and the provisions of the Ontario Personal Health Information Protection Act (PHIPA 2004) and its applicable regulations. The HSREB is qualified through the CTO REB Qualification Program and is registered with the U.S. Department of Health and Human Services (DHHS) Office for Human Research Protection (OHRP). Federalwide Assurance Number: FWA#: 00004184, IRB#: 00001173*

*HSREB members involved in the research project do not participate in the review, discussion or decision.*

## Appendix B

## What Evidence Suggests a Relationship Between Patient Age and Missing PRO Data on Randomized Clinical Trials (RCTs)?



The poster, presented at International Society for Quality of Life Research in Philadelphia, is provided on the following pages.

## Introduction

Missing PRO data can compromise the implementation of results from RCTs in clinical practice. The development of strategies to reduce missing PRO data could be guided by evidence confirming a relationship of missing data with patient characteristics.

## Aim

To examine the evidence for an association between patient age and missing PRO data on studies identified in a systematic review.

## Method

We searched MEDLINE and CINAHL to identify articles reporting on missing PRO data. Two reviewers independently screened records against eligibility criteria and identified articles that examined factors thought to cause missing PRO data in prospective health care studies. We then identified articles that examined the relationship between age and missing PRO data. We described which conditions were studied, and in the relevant RCTs, what findings were evident for associations between patient age and missing PRO data, at baseline, and during or after treatment.

## Results

One hundred articles were identified in the initial set. Twenty-three articles examined the relationship between patient age and missing PRO data **(Figure 1)** in different health conditions **(Figure 2).** Twelve articles presented a statistical analysis that included age and missing PRO data.

**Summary**

Eight of the 12 articles that reported a statistical analysis observed an association between older age and a decrease in missing PRO data (2 RCTs and 6 cohort studies). Four of the 12 articles observed an association between older age and an increase in missing PRO data (1 RCT and 3 cohort studies).

The articles with RCTs were of particular interest because RCTs can provide the best level of evidence. Three articles presented results of logistic regression for RCTs in cancer. Two articles reported odds ratios (ORs). One article just reported the significance level. The magnitude of the ORs were of clinical interest. In comparison with patients aged > 70 years: OR = 0.57 for ages < 50, OR = 0.89 for ages 50 – 60, and OR = 1.01 for ages 60 – 70. At baseline, the odds of missing PRO data decreased by 2% for each one year increase in age (OR = 0.98), and decreased by 16% for each ten year increase in age. During treatment, the odds of missing PRO data increased by 2% for each one year increase in age (OR = 1.02).

Heterogeneity precluded meta-analysis. Specifically, there was variation in: context, conceptualization of age and missing PRO data, and analytical approaches.

**Conclusions**

Evidence is sparse regarding a replicable association between patient age and missing PRO data on RCTs. Further research is required to confirm the direction and strength of associations at baseline, and during or after treatment, in different contexts.

**Acknowledgements**

MJP was supported by Queen's University Terry Fox Research Institute Transdisciplinary Training Program in Cancer Research, Queen's Graduate Award and Public Health Sciences Graduate Growth Funding. Partial support was provided by the Division of Cancer Care and Epidemiology in the Cancer Research Institute at Queen's University. RMB was supported by Sydney Catalyst, courtesy of the Cancer Institute New South Wales. MK was supported by the Australian Government through Cancer Australia.

**Presentation Details, Further Information**

Poster Number: 3054; October 21, 2017; Saturday Poster Session II

Please contact Michael J. Palmer by email: 81mjp6@queensu.ca

**References**

1. Bernhard et al. *Stat Med* 1998;17:517-532.
2. Bernhard et al. *Stat Med* 1998;17:587-601.
3. De Graaf et al. *Compr Psychiatry* 2013;54:1131-1139.
4. Di Maio et al. *Health Qual Life Outcomes* 2003;1:44.
5. Greenhalgh. *Qual Life Res* 2005;14:1363-1374.
6. Hahn et al. *Stat Med* 1998;17:547-559.
7. Hopwood. *Semin Oncol* 1996;23 (5, Suppl 10):60-64.
8. Kaasa et al. *Stat Med* 1998;17:623-632.
9. Kopp et al. *J R Soc Med* 2003;96:442-448.
10. Lamers et al. *Compr Psychiatry* 2012;53:63-70.
11. Land et al. *J Clin Oncol* 2007;25:5113-5120.
12. Marcus. *Am J Public Health* 1983;73:72-77.
13. Matthews et al. *BMC Public Health* 2006:6:185.
14. Mein et al. *BMC Med Res Methodol* 2012;12:164.
15. Menn et al. *Health Qual Life Outcomes* 2010;8:39.
16. Moynihan. *Eur J Cancer* 1998;34:9-11.
17. Olschewski. *Stat Med* 1998;17:571-575.
18. Osoba et al. *Stat Med* 1998;17:603-612.
19. Pinchon et al. *Am J Hematol* 2009;84:671-7.
20. Reimer. *Drugs Aging* 2010;27:791-800.
21. Schamber et al. *J Arthroplasty* 2013;28:1449-1453.
22. Solberg et al. *Acta Orthopaedica* 2011;82:56-63.
23. Yancik et al. *J Psychosoc Oncol* 1989;7:59-74.

**Figure 1. Flow Diagram.**

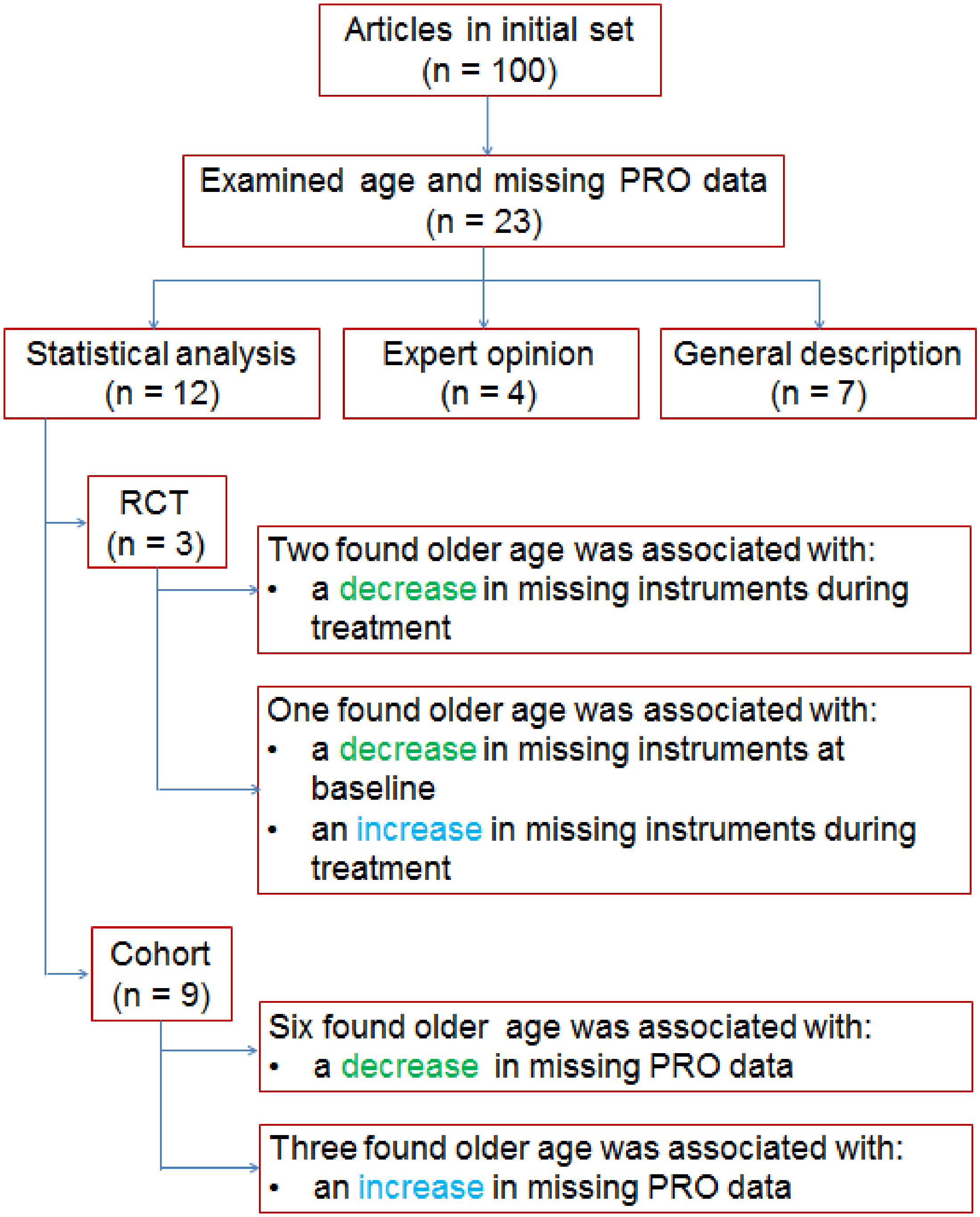
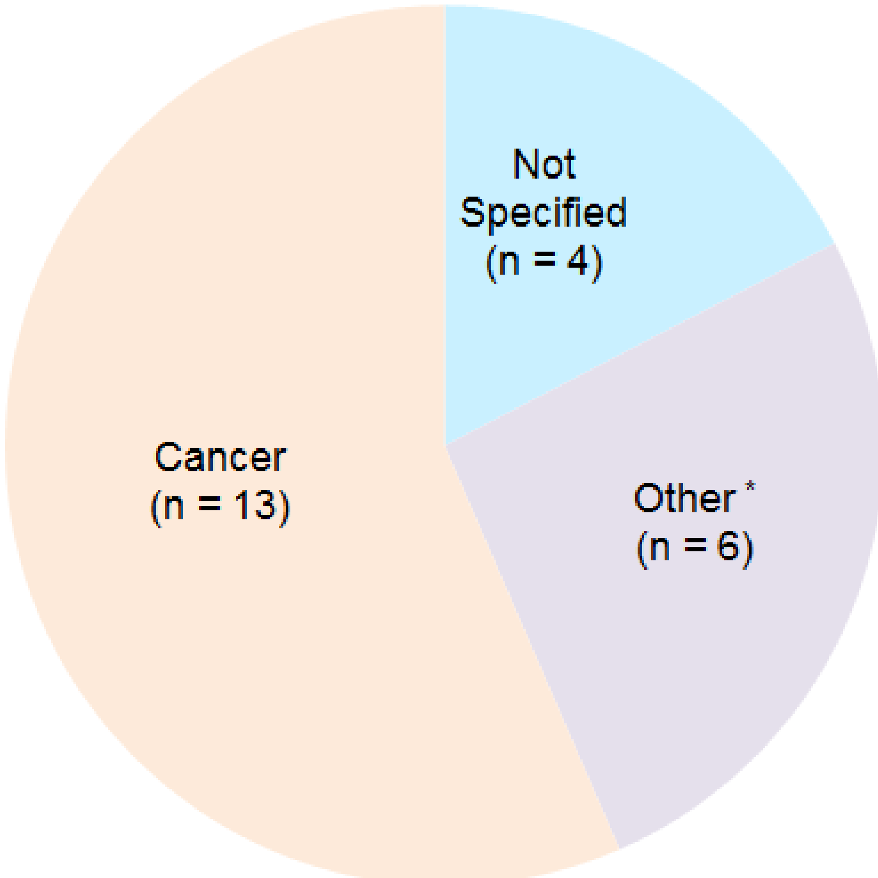


**Figure 2. Scope of Disease.**




* anxiety, depression, multiple sclerosis, total joint arthroplasty, lumbar spine disorders, chronic obstructive pulmonary disease

# What Evidence Suggests a Relationship Between Patient Age and Missing PRO Data on Randomized Clinical Trials (RCTs)?

Michael J. Palmer[1,2], Rebecca Mercieca-Bebber[3,4], Madeleine King[3,4], Melanie Calvert[5,6], Harriet Richardson[1,2], Michael Brundage[1,2]

[1]Department of Public Health Sciences, [2]Cancer Care and Epidemiology, Cancer Research Institute, Queen's University, Kingston, Canada; [3]Central Clinical School, Sydney Medical School, [4]Quality of Life Office, School of Psychology, Lifehouse C39Z, University of Sydney NSW 2006, Australia; [5]Centre for Patient Reported Outcomes Research, [6]Institute of Applied Health Research, University of Birmingham, Birmingham, United Kingdom

## Introduction

Missing PRO data can compromise the implementation of results from RCTs in clinical practice.

The development of strategies to reduce missing PRO data could be guided by evidence confirming a relationship of missing data with patient characteristics.

## Aim

To examine the evidence for an association between patient age and missing PRO data on studies identified in a systematic review.

## Method

As part of a previous review, we searched MEDLINE and CINAHL to identify articles reporting on missing PRO data. Two reviewers independently screened 6027 records against eligibility criteria and identified articles that examined factors thought to cause missing PRO data in prospective health care studies.

For this review, we screened 100 eligible articles and identified articles that examined a relationship between age and missing PRO data. We described which conditions were studied, and in the relevant RCTs, any evidence of an association between patient age and missing PRO data

- at baseline
- during or after treatment

## *Articles


| Statistical analysis |
|---|
| Land et al. *J Clin Oncol* 2007;25:5113-5120. |
| Osoba et al. *Stat Med* 1998;17:603-612. |
| Hahn et al. *Stat Med* 1998;17:547-559. |
| de Graaf et al. *Compr Psychiatry* 2013;54:1131-1139. |
| Greenhalgh. *Qual Life Res* 2005;14:1363-1374. |
| Lamers et al. *Compr Psychiatry* 2012;53:63-70. |
| Marcus. *Am J Public Health* 1983;73:72-77. |
| Matthews et al. *BMC Public Health* 2006;6:185. |
| Solberg et al. *Acta Orthopaedica* 2011;82:56-63. |
| Kopp et al. *J R Soc Med* 2003;96:442-448. |
| Mein et al. *BMC Med Res Methodol* 2012;12:164. |
| Schamber et al. *J Arthroplasty* 2013;28:1449-1453. |
| **Expert opinion** |
| Bernhard et al. Stat Med 1998;17:517-532. |
| Bernhard et al. Stat Med 1998;17:587-601. |
| Hopwood. Semin Oncol 1996;23 (5, Suppl 10):60-64. |
| Moynihan. Eur J Cancer 1998;34:9-11. |
| **General description** |
| Di Maio et al. Health Qual Life Outcomes 2003;1:44. |
| Kaasa et al. Stat Med 1998;17:623-632. |
| Mann et al. Health Qual Life Outcomes 2010;8:39. |
| Olschewski. Stat Med 1998;17:571-575. |
| Pinchon et al. Am J Hematol 2009;84:671-7. |
| Reimer. Drugs Aging 2010;27:791-800. |
| Yancik et al. J Psychosoc Oncol 1989;7:59-74. |


## Results

One hundred articles were identified in the initial set. Twenty-three articles* examined the relationship between patient age and missing PRO data (**Figure 1**) in different health conditions (**Figure 2**). Twelve articles presented a statistical analysis that included age and missing PRO data. The findings are described in the following table.

| Design | Frequency | Older age was associated with: | Note |
|---|---|---|---|
| RCT | 2 | a decrease in missing instruments during treatment | A |
| | 1 | a decrease in missing instruments at baseline, and an increase in missing instruments during treatment | B |
| Cohort | 6 | a decrease in missing PRO data | C |
| | 3 | an increase in missing PRO data | D |

| | |
|---|---|
| A | One article reported p-value ($p < 0.0002$) for > 60 versus ≤ 60 years. The other article provided odds ratios: OR = 1.75 for < 50, OR = 1.12 for 50–60, OR = 0.99 for 60–70 versus > 70 years. |
| B | At baseline, the odds of missing PRO data decreased by 2% for each one year increase in age (OR = 0.98), and decreased by 16% for each ten year increase in age. During treatment, the odds of missing PRO data increased by 2% for each one year increase in age (OR = 1.02). |
| C | Two articles reported odds of missing PRO data decreased by 2% for each one year increase in age. |
| D | One article reported odds of missing PRO data increased by 8% for each one year increase in age. |

Most of the 11 articles that provided expert opinion or a general description reported an association between older age and an increase in missing PRO data.

## Summary

Eight of the 12 articles that reported a statistical analysis observed an association between older age and a decrease in missing PRO data (2 RCTs and 6 cohort studies). One RCT observed an association between older age and, a decrease in missing PRO data at baseline but an increase in missing PRO data during treatment. Three cohort studies observed an association between older age and an increase in missing PRO data.

The three articles with RCTs were of particular interest because RCTs can provide the best level of evidence. The magnitude of the ORs were of clinical interest. Heterogeneity precluded meta-analysis. Specifically, there was variation in: context, conceptualization of age and missing PRO data, and analytical approaches.

## Conclusions

Evidence is sparse regarding a replicable association between age and missing PRO data on RCTs. Further research is required to confirm the direction and strength of associations at baseline, and during or after treatment, in different contexts.

**Figure 1. Flow Diagram.**

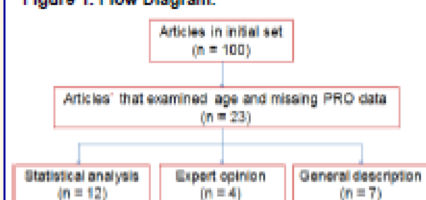
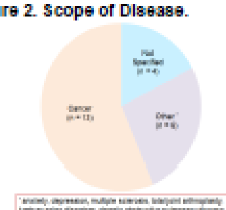



**Figure 2. Scope of Disease.**



## Acknowledgements

MJP was supported by Queen's University Terry Fox Research Institute Transdisciplinary Training Program in Cancer Research, Queen's Graduate Award and Health Sciences Graduate Growth Funding. Partial support was provided by the Division of Cancer Care and Epidemiology in the Cancer Research Institute at Queen's University. RMB was supported by Sydney Catalyst, courtesy of the Cancer Institute New South Wales. MK was supported by the Australian Government through Cancer Australia.

## Presentation Details

Poster Number: 3054
October 21, 2017
Saturday Poster Session II

## Further Information

Please contact Michael J. Palmer
by email:
81mjp6@queensu.ca

## Appendix C

## Are they "sufficiently comparable"? A cautionary example of analysis after pooling individual participant data from three randomized clinical trials.

Michael J. Palmer, Harriet Richardson, Dongsheng Tu, Michael Brundage

Additional quantitative analyses were performed on pooled individual participant data in Part 2. The results are formatted as a commentary for submission to *Translational Behavioral Medicine*.

## Introduction

Patient-reported outcomes describe, from the patient's perspective, symptoms associated with disease, toxicity arising from treatment, the ability to function and to perform activities of daily living, and different dimensions of quality of life. A phase III randomized clinical trial is the gold standard for the evaluation of treatment efficacy and often includes the assessment of patient-reported outcomes. Unfortunately, subjects do not always complete the instrument when expected during treatment or follow-up. Missing data on patient-reported outcomes can seriously threaten the validity of the randomized clinical trial and can undermine its scientific credibility [1], but why these data go missing is largely unknown.

Our research group has interest in identifying the putative causes of missing patient-reported reported outcome data in randomized clinical trials. Although missing data commonly occur, occurrence can be infrequent, hence our desire to pool individual participant data from several randomized clinical trials, to increase the number of events from a larger sample size. Our approach to analysis was novel. Unlike most researchers, we did not aim to estimate the effect of one, designated variable on outcome. Rather, our primary focus was in exploring which of many candidate variables were associated with the first missing instrument after randomization.

Bangdiwala et al [2] provide guidance regarding statistical methodologies to pool across multiple intervention studies. The authors point out on page 229 that whether the studies are "sufficiently comparable" is a key consideration before pooling and accordingly, the concept can provide justification for proceeding with pooling. In this commentary, we describe our analysis and then, after reflecting on our decision to pool individual participant data, provide several caveats for future research worth sharing from this experience.

**Methods**

The study is described in detail elsewhere [3] but briefly summarized here. Three randomized clinical trials were selected from a list of candidates identified according to *a priori* eligibility criteria: MA17 [4], PR7 [5], and LY12 [6]. Individual participant data for subjects from Canada were requested and obtained from the sponsor. Cox proportional hazards regression was used to determine which characteristics of the subjects and the centres predict the time to the first missing instrument after randomization.

Separate analyses were performed for each of the three randomized clinical trials and reported elsewhere [3]. Before pooling, the estimate obtained for post-graduate residency training program from each randomized clinical trial was combined in an estimate weighted by the inverse of variance. A Forest plot was created to facilitate visual judgement of heterogeneity, and Cochran's Q and Higgins $I^2$ tested statistical heterogeneity [7, 8]. To account for clustering of patients within the randomized clinical trials [9], a new variable (RCT = 1 for MA17, RCT = 2 for PR7, RCT = 3 for LY12) was included as a random effect in every model, after pooling individual participant data from the three randomized clinical trials [10]. The pooled analysis provided an opportunity to examine effect modification with increased power [11]. Rather than evaluate all possible statistical interactions, the focus was on plausible effect modification involving age and gender, but not type of cancer because of collinearity with the specific randomized clinical trial. The proportional hazards assumption was assessed with: Kaplan-Meier estimates of the survival function [12], log-minus-log survival plots [13], standardized score processes and supremum tests [14].

Many covariates were only included in quantitative analyses of individual participant data from separate randomized clinical trials. Five covariates described characteristics of the subjects: employment status, race / ethnicity, insurance, stage, and caregiver. Scale scores from patient-reported outcome instruments completed by subjects at baseline described their cognitive status (cognitive function), toxicity (pain, fatigue, and nausea and vomiting) and level of well-being (general health, global health / quality of life, physical well-being, social / family well-being, emotional well-being, functional well-being, total FACT-G score, and trial outcome index). One covariate described the study treatment that was assigned at randomization.

Six covariates were common across the randomized clinical trials and were included in the pooled analyses. Three of the six covariates described characteristics of the subjects: age, gender, and performance status as measured by ECOG scale. Three covariates described characteristics of the centres: level of activity, presence of a post-graduate residency training program, and geographic location in Canada. Descriptive analysis and univariate Cox regression analysis were performed for each of the six covariates, and then a multivariate model was developed from those covariates having $p < 0.20$ on univariate Cox regression.

## Results

### Assessment of Heterogeneity

**Figure C-1** presents the Forest plot. The trial-specific hazard ratios were non-homogeneous for post-graduate residency training program, with the lower tails of the MA17 and PR7 confidence intervals above unity and including the combined estimate, and the lower tail of the LY12 confidence interval below unity and the upper tail not including the combined estimate.

**Appendix C-I** presents the calculations for the statistical assessment of heterogeneity. Cochran's Q was 6.66 ($p > 0.05$). Higgins $I^2$ was 70%.

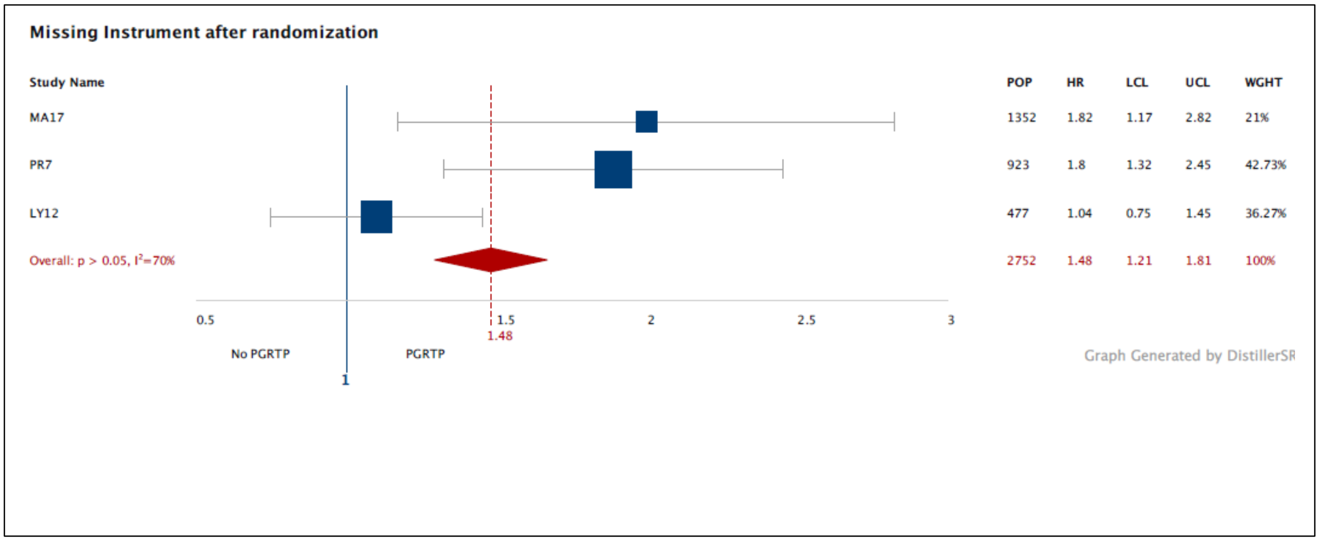



Figure C-1 was created from on-line resource, DistillerSR Forest Plot Generator from Evidence Partners https://www.evidencepartners.com/resources/forest-plot-generator/#forest_plot_6_graph_edit_settings

Descriptive Analysis and Cox Proportional Hazards Regression Analysis

In analysis of pooled individual participant data from three randomized clinical trials, 27% of 2752 subjects had a missing instrument. **Table C-1** presents descriptive analysis and Cox proportional hazards regression analysis.

In descriptive analysis, significant differences between those with one or more missing instruments and those with none missing were observed for the age, gender, and performance status of the subjects, and for the presence of a post-graduate residency training program and geographic location of the centres.

In univariate Cox regression, significant associations with time to the first missing instrument were observed for age and performance status of the subjects, and for presence of a post-graduate residency training program and geographic location of the centres. These four covariates remained in the final multivariate model. Subjects with a one-year increase in age had a one percent decrease in the relative risk of missing instrument. Subjects with a performance status of 2 (HR = 1.52) or 3 (HR = 2.53) had an increase in the relative risk, compared to subjects with a performance status of 0. Subjects at centres with a post-graduate residency training program had a 1.47-fold increase in the relative risk, compared to subjects at centres without a post-graduate residency training program. Subjects from centres in eastern Canada (HR = 1.97), Ontario (HR = 1.46) or Quebec (HR = 1.20) had an increase in the relative risk, compared to subjects from centres in western Canada. The 95% confidence interval for Quebec bounded unity.

Assessment of Effect Modificiation

In Cox regression, there was non-significance for effect modification between the age and gender of subjects with time to the first missing instrument (p = 0.416).

Assessment of Proportional Hazards Assumption

**Appendix C-II** presents assessment of the proportional hazards assumption with Kaplan-Meier estimates of the survival function, with log-minus-log survival plots, and with standardized score processes and supremum tests. Proportional hazards were observed for presence of a post-graduate residency training program and geographic location of the centres, but perhaps not for age and one level of performance status of the subjects.

**Discussion**

According to the homogeneity assumption, the randomized clinical trials estimate the same value for the measure of effect, and any differences are due to random error. Cochran's Q considers how close (or how far) the summary measure of effect is in relation to the overall effect. The value will approach 0 if all studies are close to the overall effect. In our study, Cochran's Q was 6.66. In most studies, the statistical test has little power. In our study, individual participant data from three randomized clinical trials were included in the pooled analysis, so the statistical test, with two degrees of freedom, had low power. Therefore, the non-significant p-value ($p > 0.05$, for Cochran's Q) provides no justification for an assumption of homogeneity. A higher cut-off (for example, $p > 0.20$) could be used to address the problem of low power, but a more thoughtful approach would consider whether contextual expectations and the measure of effect are likely to vary [15].

Heterogeneity is expected and likely to arise from different sources, such as the eligibility criteria for subjects, type of treatment, and length of follow up, so testing for heterogeneity, although a common practice, is not as important as the extent to which the heterogeneity may affect the conclusions [7]. Higgins created $I^2$ to quantify the effect of heterogeneity. The calculation for $I^2$ relates Cochran's Q to the number of randomized clinical trials that are included in the pooled analysis. For a given value of Q, the value for $I^2$ will decrease as the number of randomized clinical trials that are included in the pooled analysis increases. A value of 0% for $I^2$ indicates no observed heterogeneity, while a larger value would indicate heterogeneity. In our study, across the three randomized clinical trials, 70% of total variation in presence of a post-graduate residency training program is due to heterogeneity of the centres rather than chance. This may be considered more than moderate (50%) and less than high (75%) heterogeneity [7]. The presence of post-graduate residency training program was chosen for an assessment of heterogeneity to demonstrate the procedure, but we did not use the results to determine if pooling should proceed, because we did not consider the covariate – or any covariate – to be the primary exposure of interest. We recognize that five other covariates may be a source of potential heterogeneity. Of these five covariates, one was binary (gender), two were continuous (age of the subjects, level of activity of the centres), and two were categorical (performance status of the subjects, geographic location of the centres). We did not feel it was necessary to assess their heterogeneity before proceeding with pooling. Our intention was to explore whether any covariate predicts the time to first missing instrument. After pooling the individual participant data, we performed similar analyses, ensuring a random effect for trial was in every model to account for clustering of subjects.

Pooling provided an opportunity to learn from model development with information across the randomized clinical trials. Although the randomized clinical trials were not designed to measure missing patient-reported outcome data, the assessment of patient-reported outcomes was a secondary objective in each trial. Several additional important requirements were consistent for the collection of data. Collectively, these suggested the three randomized clinical trials were sufficiently comparable [2] and provided justification for proceeding with pooling. One sponsor had oversight for trial conduct and applied similar quality assurance. All treatment had curative intention. All centres were in Canada where collection of patient-reported outcome data was a mandatory requirement. Staff at each centre administered the instruments to eligible subjects in paper format, and sent the instruments with case report forms to the sponsor for processing and data entry according to standard operating procedures. All subjects were at least 18 years of age, had sufficient mental capacity to provide written informed consent, and were literate in the available languages of the instruments. Finally, inclusion in our study required each subject demonstrate an ability and willingness to complete the instrument before randomization, complying with standard eligibility criterion of the sponsor.

Cox proportional hazards regression of PR7 revealed the relative risk for age decreased by three percent for every one-year increase in age in the final multivariate model (Hazard Ratio = 0.97; 95% Confidence Interval = 0.95, 0.99; p-value = 0.0004) [3]. Cox regression after pooling three randomized clinical trials revealed an attenuation of relative risk for age in a different final multivariate model (HR = 0.99; 95% CI = 0.99, 1.00; p-value = 0.198).

Intriguingly, the performance status of subjects was a significant predictor of outcome only in the multivariate model after pooling but not in multivariate models for the individual randomized clinical trials. Conversely, level of activity of the centres was a significant predictor of outcome in the multivariate models for two of the three randomized clinical trials but not after pooling. Osoba [16] and Hahn [17] independently investigated the association of subject-related factors and institution with missing instruments. Of note, both authors ignored clinical heterogeneity by pooling individual participant data from convenience samples of randomized clinical trials that provided treatment intended for cure with randomized clinical trials that provided treatment intended for palliation. In contrast to our study, the two authors used logistic regression for the analysis and conceptualized the outcome in a different manner. Osoba reported a diagnosis of breast or ovarian cancer, and low centre accrual were independent predictors of missing instruments during treatment, but subjects with age >60 years, distant metastases and a performance status of two had fewer missing instruments during treatment [16]. Hahn reported at three-four months, missing instruments increased with age and performance status, and were greater in subjects with a diagnosis of lung compared to breast cancer [17]. Our result for age was consistent with Osoba [16] but in contrast to Hahn [17]. Our result for performance status was consistent with Hahn [17] but in contrast to Osoba [16]. Guidance from prior research is not always helpful and comparison with prior research may be difficult to interpret.

The study design of MA17 and PR7 precluded separate examination of effect modification because the gender of subjects on these two randomized clinical trials was exclusively female or exclusively male, respectively. In contrast, the study design of LY12 did include a mix of women and men but in univariate Cox regression, there was no effect modification between the age and gender of subjects with time to the first missing instrument ($p = 0.527$). One benefit of the pooled analysis was the opportunity to examine effect modification with increased power from inclusion of all 2752 subjects from the three randomized clinical trials. After pooling, there was non-significance for effect modification between the age and gender of subjects with time to the first missing instrument ($p = 0.416$). The data provide insufficient evidence that the null hypothesis of no difference between women and men in the time to the first missing instrument after randomization, given age, is false. Accordingly, researchers who design randomized clinical trials need not have concern that older men may have a greater chance of missing instrument in comparison to older women.

The key assumption of Cox regression is proportional hazards, which means that the ratio of hazard rates is approximately constant at every time point during the study [18]. We assessed the assumption through visual inspection of Kaplan-Meier curves and log-minus-log survival plots, and with statistical tests. Proportional hazards were observed for presence of a post-graduate residency training program at the centres and for geographic location of the centres. We believe possible departures in proportional hazards for age of subjects and for performance status of subjects may be explainable by the large sample size [19].

Reflecting on our experience with pooling individual participant data from three randomized clinical trials revealed three caveats:

1) The analysis of pooled data led to two spurious results. Interpretation of the final multivariate model after pooling, in the absence of results from analyses of the separate randomized clinical trials, would lead one to mistakenly conclude that the performance status of subjects was a significant predictor, and that the level of activity of the centres was not a significant predictor, of the time to first missing instrument after randomization.

2) Important covariates were not available for analysis of pooled data. Of particular note, scale scores from patient-reported outcome instruments that were completed by subjects at baseline were only included in the analyses of the separate randomized clinical trials. One example was the functional well-being scale of the FACT-General instrument on LY12, which had a significant association with outcome.

3) The schedules of assessment were not the same and in fact were quite different. The Kapan-Meier estimate of the survival function is essentially LY12 for the first few months, then MA17 and PR7 until five years, then PR7 for subsequent years. The observation period for LY12 ends before the first assessment for MA17. A few assessment time-points were only common to MA17 and PR7 and these did not include LY12.

Rather than deciding whether the studies are “sufficiently comparable” [2] before pooling, perhaps the question should be are they sufficiently different? In our experience, the results from the pooled analysis were misleading, and do not contribute to our understanding of the putative causes of missing patient-reported reported outcome data in randomized clinical trials. We recommend analysis and reporting of separate randomized clinical trials.

**Table C-1. Descriptive analysis and Cox proportional hazards regression of pooled individual participant data from three randomized clinical trials.**

| Component | Characteristic | Level | Descriptive analysis | | | | | Cox proportional hazards regression | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | One or more missing | | None missing | | p-value | Univariate | p-value | Multivariate | p-value |
| | | | n | % | n | % | | HR (95% CI) [1] | | HR (95% CI) [1] | |
| Participant | age [2a, 2b] | mean (SD) | 63.8 (12.4) | | 65.8 (10.3) | | **0.0001** | 0.99 (0.98, 1.00) | **0.168** | 0.99 (0.99, 1.00) | **0.198** |
| | | 20.0 – 49.9 | 95 | 43.8 | 122 | 56.2 | **<.0001** | reference | 0.824 | | |
| | | 50.0 – 59.9 | 164 | 25.3 | 485 | 74.7 | | 1.02 (0.79, 1.32) | | | |
| | | 60.0 – 69.9 | 212 | 25.5 | 619 | 74.5 | | 1.01 (0.77, 1.33) | | | |
| | | 70.0 – 79.9 | 222 | 25.6 | 647 | 74.4 | | 0.89 (0.64, 1.23) | | | |
| | | 80.0 – 89.9 | 49 | 26.3 | 137 | 73.7 | | 0.89 (0.59, 1.35) | | | |
| | gender | Female | 238 | 15.4 | 1303 | 84.6 | **<.0001** | reference | 0.550 | | |
| | | Male | 504 | 41.6 | 707 | 58.4 | | 0.93 (0.73, 1.18) | | | |
| | physical impairment [3] | 0 | 491 | 23.3 | 1614 | 76.7 | **<.0001** | reference | **0.012** | reference | **0.006** |
| | | 1 | 209 | 36.4 | 366 | 63.6 | | 1.11 (0.93, 1.33) | | 1.07 (0.90, 1.28) | |
| | | 2 | 30 | 52.6 | 27 | 47.4 | | 1.51 (1.02, 2.24) | | 1.52 (1.03, 2.26) | |
| | | 3 | 12 | 80.0 | 3 | 20.0 | | 2.30 (1.28, 4.14) | | 2.53 (1.40, 4.56) | |
| Centre | level of activity – patients [4a, 4b] | Median (IR) | 17.5 (28.0) | | 18.0 (30.0) | | 0.448 | 1.00 (1.00, 1.00) | 0.459 | | |
| | | ≤ 9 | 238 | 28.3 | 604 | 71.7 | 0.716 | reference | 0.876 | | |
| | | 10 - 19 | 163 | 26.2 | 459 | 73.8 | | 0.98 (0.80, 1.19) | | | |
| | | 20 - 39 | 174 | 27.2 | 467 | 72.8 | | 1.03 (0.85, 1.25) | | | |
| | | ≥ 40 | 167 | 25.8 | 480 | 74.2 | | 0.95 (0.77, 1.16) | | | |
| | research infrastructure [5] | Yes | 630 | 30.3 | 1448 | 69.7 | **<.0001** | 1.50 (1.23, 1.84) | **<.0001** | 1.47 (1.19, 1.81) | **0.0003** |
| | | No | 112 | 16.6 | 562 | 83.4 | | reference | | reference | |
| | location [6] | East | 85 | 35.6 | 154 | 64.4 | **<.0001** | 2.05 (1.54, 2.72) | **<.0001** | 1.97 (1.48, 2.62) | **<.0001** |
| | | Quebec | 98 | 23.9 | 312 | 76.1 | | 1.32 (1.00, 1.72) | | 1.20 (0.91, 1.59) | |
| | | Ontario | 445 | 28.8 | 1101 | 71.2 | | 1.53 (1.24, 1.88) | | 1.46 (1.19, 1.80) | |
| | | West | 114 | 20.5 | 443 | 79.5 | | reference | | reference | |

[1] Hazard Ratio (95% Confidence Interval).

Age at randomization in years: [2a] continuous, and categories based on [2b] decades.

[3] Performance status as measured by ECOG scale.

Number of patients entered on study: [4a] continuous, and [4b] categories.

[5] Post-graduate residency training program in surgical oncology, gynecologic oncology, medical oncology, radiation oncology or hematology. Each centre was assigned Yes or No. Yes included 30 centres: CAAJ, CABN, CABP, CAEJ, CAGH, CAGQ, CAHA, CAHC, CAHF, CAHL, CAHM, CAHN, CAHO, CAKK, CAKO, CALC, CALM, CAMG, CAMM, CAMN, CAMP, CAMS, CAMX, CANL, CARM, CATC, CATW, CAVA, CAVF and CAVV. No included 31 centres: CACC, CAEM, CAER, CAGL, CAGS, CAGV, CAJV, CAKP, CALO, CALS, CALY, CAME, CAMH, CAMJ, CAMR, CAMT, CAMU, CAMV, CANH, CANW, CAPN, CAPS, CAPT, CASA, CASS, CATL, CAVC, CAVD, CAVG, CAVK and CAVN.

[6] Each centre was assigned to one of four categories according to geographic location in Canada. East included 7 centres in the Atlantic provinces, Quebec included 13 centres, Ontario included 27 centres and West included 14 centres in the western provinces.

**Appendix C-I. Calculations for Assessment of Heterogeneity.**

In the following table:

RCT = randomized clinical trial
HR = hazard ratio for presence of post-graduate residency training program
95% CI = 95% confidence interval for hazard ratio
b = parameter estimate for presence of post-graduate residency training program
SE = standard error for parameter estimate
Ƃ = weighted average for parameter estimate

| RCT | HR | 95% CI | b | SE | $w = 1 / SE^2$ | wb | (b – Ƃ) | $(b – Ƃ)^2$ | $w (b – Ƃ)^2$ |
|---|---|---|---|---|---|---|---|---|---|
| MA17 | 1.82 | 1.17, 2.82 | 0.59820 | 0.22434 | 19.869 | 11.8856 | +0.20635 | 0.04258 | 0.846 |
| PR7 | 1.80 | 1.32, 2.45 | 0.58899 | 0.15728 | 40.425 | 23.8099 | +0.19714 | 0.03886 | 1.571 |
| LY12 | 1.04 | 0.74, 1.45 | 0.04000 | 0.17074 | 34.303 | 1.3721 | -0.35185 | 0.12379 | 4.246 |
| ∑ | | | | | 94.597 | 37.0676 | | | $\chi_h^2 = 6.66$ |

<u>Calculation of weighted average for parameter estimate</u>

Ƃ = ∑ w b / ∑ w = 37.0676 / 94.597 = 0.39185

<u>Calculation of parameter estimate and 95% confidence interval</u>

$e^{0.39185} = 1.48$, estimated SE of Ƃ = S = $1 / (\sum w)^{1/2} = 1 / (94.597)^{1/2} = 1 / 9.726 = 0.1028$
95% CI for 1.48 = $(e^{0.39185 - 1.96*0.1028}, e^{0.39185 + 1.96*0.1028}) = (e^{0.19036}, e^{0.59334}) = (1.21, 1.81)$

<u>Calculation of Cochran's Q and statistical test</u>

$\chi_h^2 = \sum w (b - Ƃ)^2$, where Ƃ = ∑ w b / ∑ w; $\chi^2$ distribution and df = [number of RCT – 1]

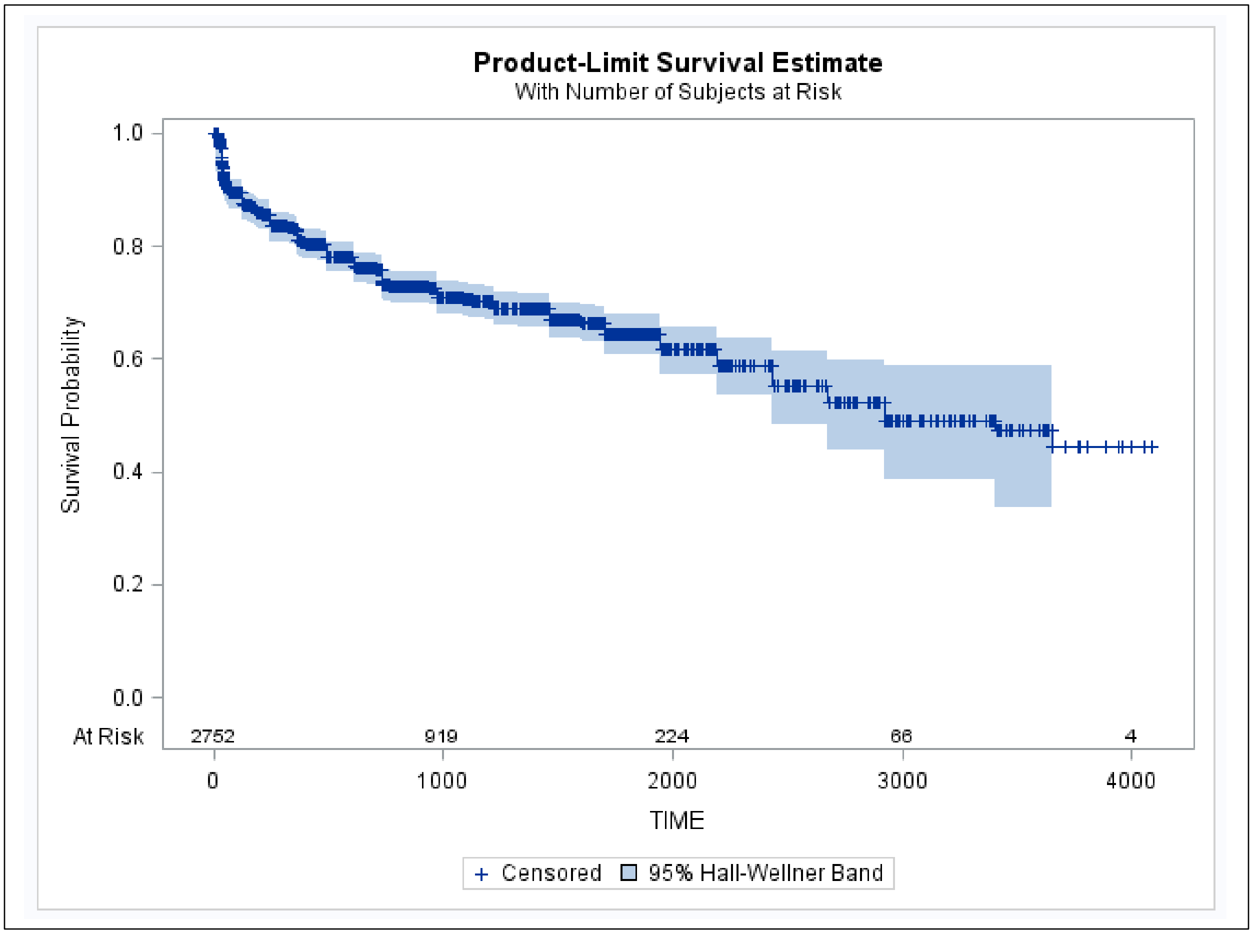
$\chi_h^2 = \sum w (b - Ƃ)^2 = 6.66$ with df = 3 – 1 = 2

For a two-sided test at alpha = 0.05 for the $\chi^2$ distribution, with 2 degrees of freedom: lower-tail critical value is 0.0506, upper-tail critical value is 7.38 (Rossner 2011, Fundamentals of Biostatistics, 7th Edition, page 824, Table 6, Percentage points of the Chi-square distribution). The null hypothesis for the statistical test of heterogeneity is all RCT evaluate the same value for effect. 0.0506 < Cochran's Q < 7.38. The test statistic is not less than the lower-tail critical value, nor is the test statistic greater than the upper-tail critical value. Fail to reject the null hypothesis. The non-significant p-value (p > 0.05) cannot be used as evidence for homogeneity.

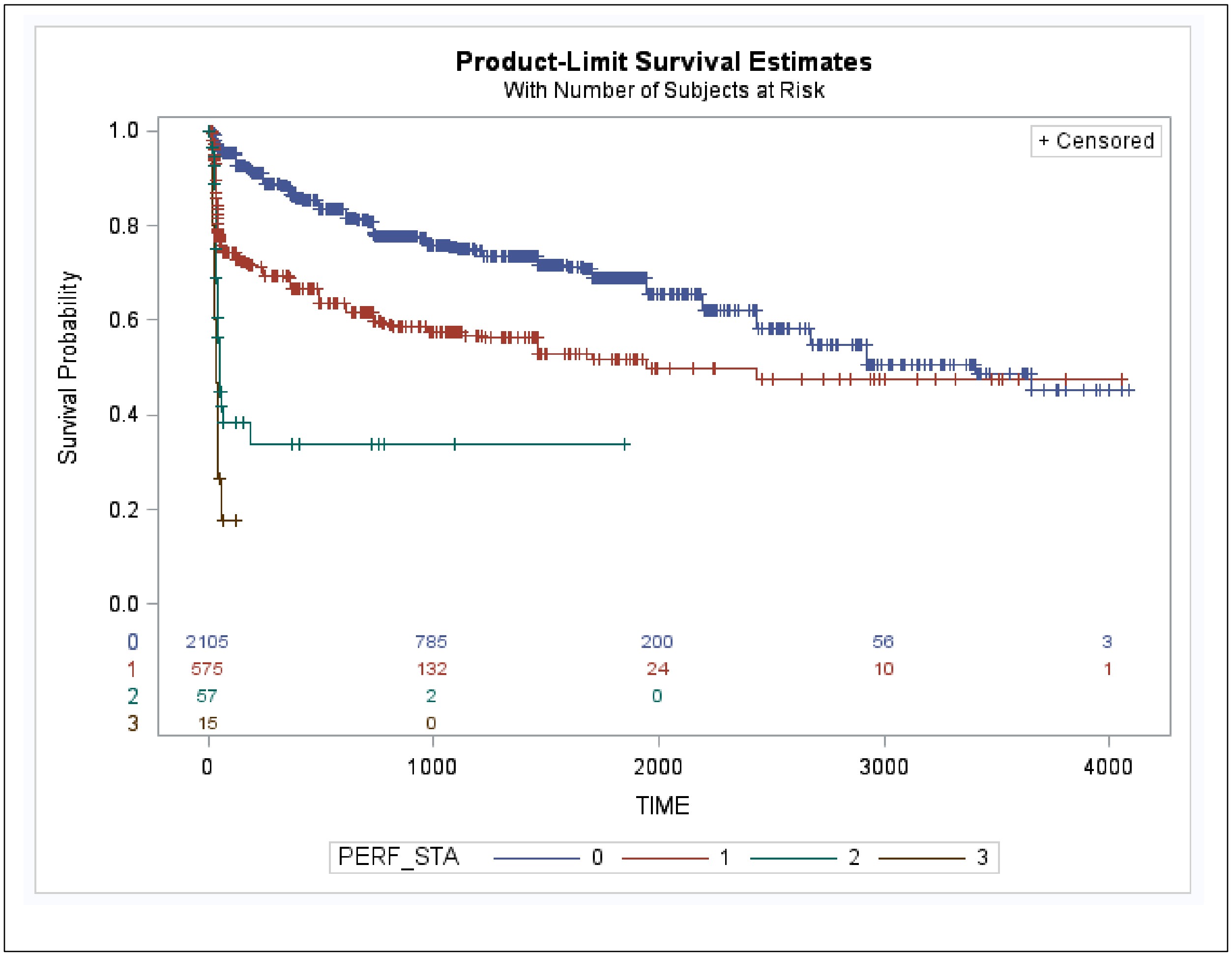
<u>Calculation of Higgins $I^2$ test</u>

$I^2$ = [(Cochran's Q – df) / Cochran's Q ] x 100 = [(6.66 – 2) / 6.66] x 100 = 69.97.

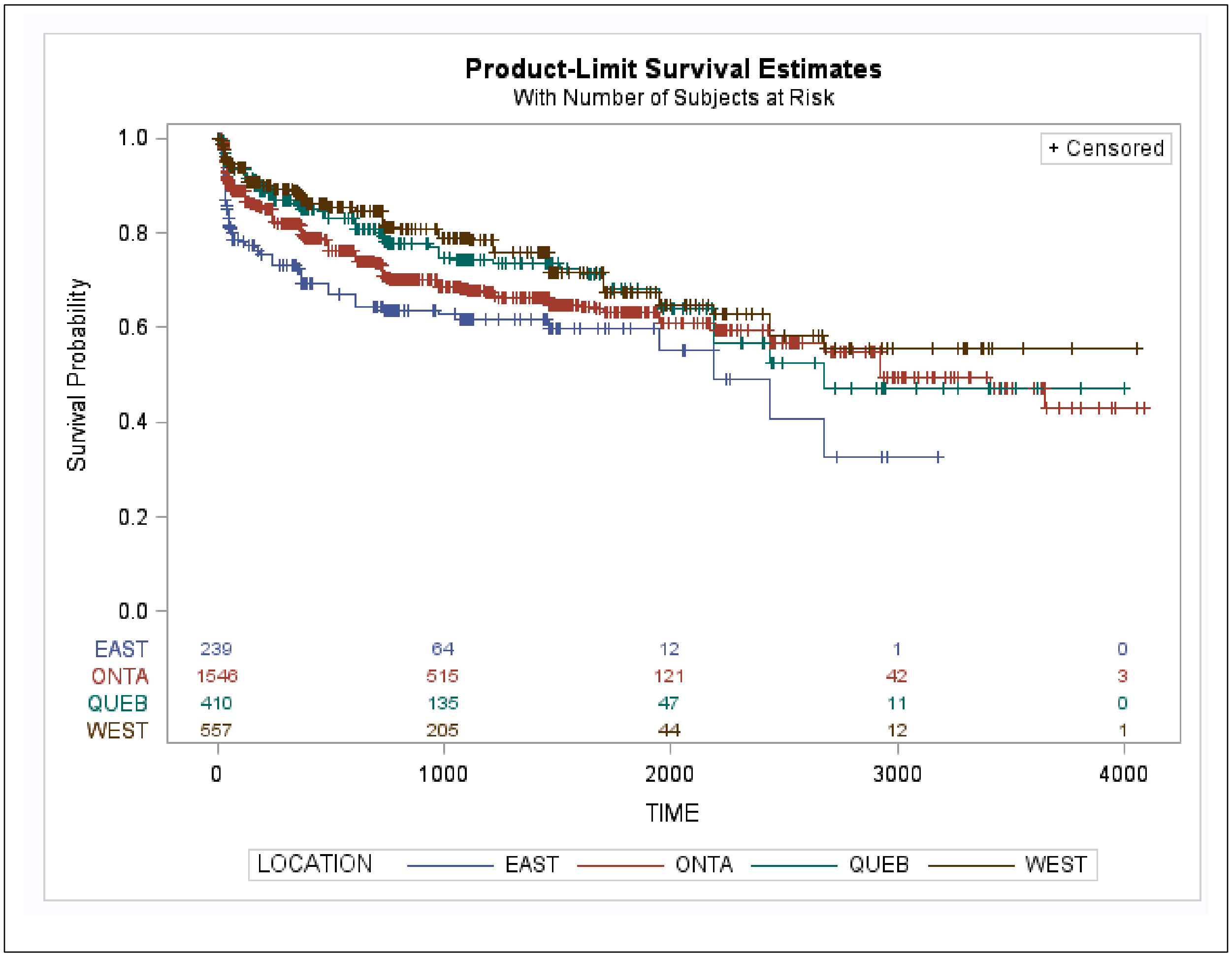
**Appendix C-II.** **Assessment of the proportional hazards assumption with Kaplan-Meier estimates of the survival function, with log-minus-log survival plots, and with standardized score processes and supremum tests.**



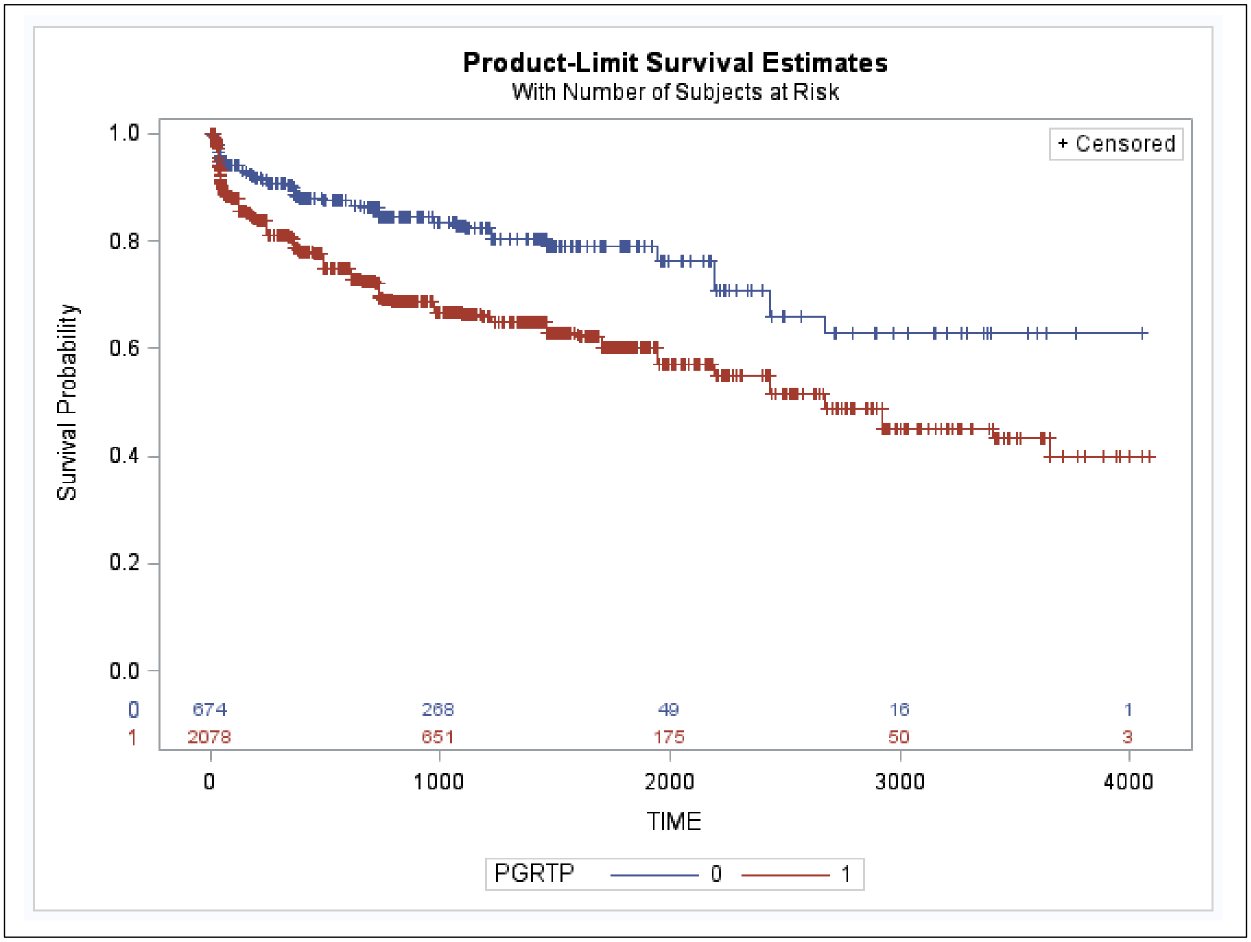

Graph is Kaplan-Meier estimate of survival function. Step function drops when a subject had an event. Graph remains flat between failure times. Vertical ticks indicate censored observations. The survival function drops most steeply at the start. This suggests the hazard rate is highest immediately after randomization. 25% of sample had an event within 0 to 730 days, and 25% of sample had an event within 731 to 2922 days. 50% of sample had an event within 2922 days (= 8 years) of randomization. The survival function does not reach 0 because the observation with longest follow-up was censored. The blue area above and below the survival curve is the Hall-Wellner confidence band. 95% of all confidence intervals are within this band.




Graph is Kaplan-Meier estimate of survival function, stratified by ECOG performance status. The median time to first missing instrument was 3409 days for performance status = 0, 2435 days for performance status = 1, 46 days for performance status = 2, and 37 days for performance status = 3. The Log-rank test was used to assess the equality of survival function over strata. The calculation sums differences between the observed and expected number of failures for each stratum at each time point. The same proportion of subjects is expected to have an event in each interval if the survival function is the same in each stratum. If the proportions differ over time, then the statistic will be large and the null hypothesis of no difference will be rejected. Log-rank weights differences equally at all time points. Log-rank: Chi-square = 378.8, 3 degrees of freedom, $p < 0.0001$. The time from randomization to first missing instrument after randomization was significantly shorter for subjects with a lower performance status.




Graph is Kaplan-Meier estimate of survival function, stratified by location of centre. The median time to first missing instrument was 2191 days for subjects at centres in the Atlantic provinces (EAST), 2678 days at centres in Quebec (QUEB), and 2922 days at centres in Ontario (ONTA). The median time to the first missing instrument was not observed for subjects at centres in the western provinces (WEST). Log-rank: Chi-square = 32.3, 3 degrees of freedom, $p < 0.0001$. The time from randomization to first missing instrument after randomization was significantly shorter for subjects from a centre in the Atlantic provinces, Quebec or Ontario.



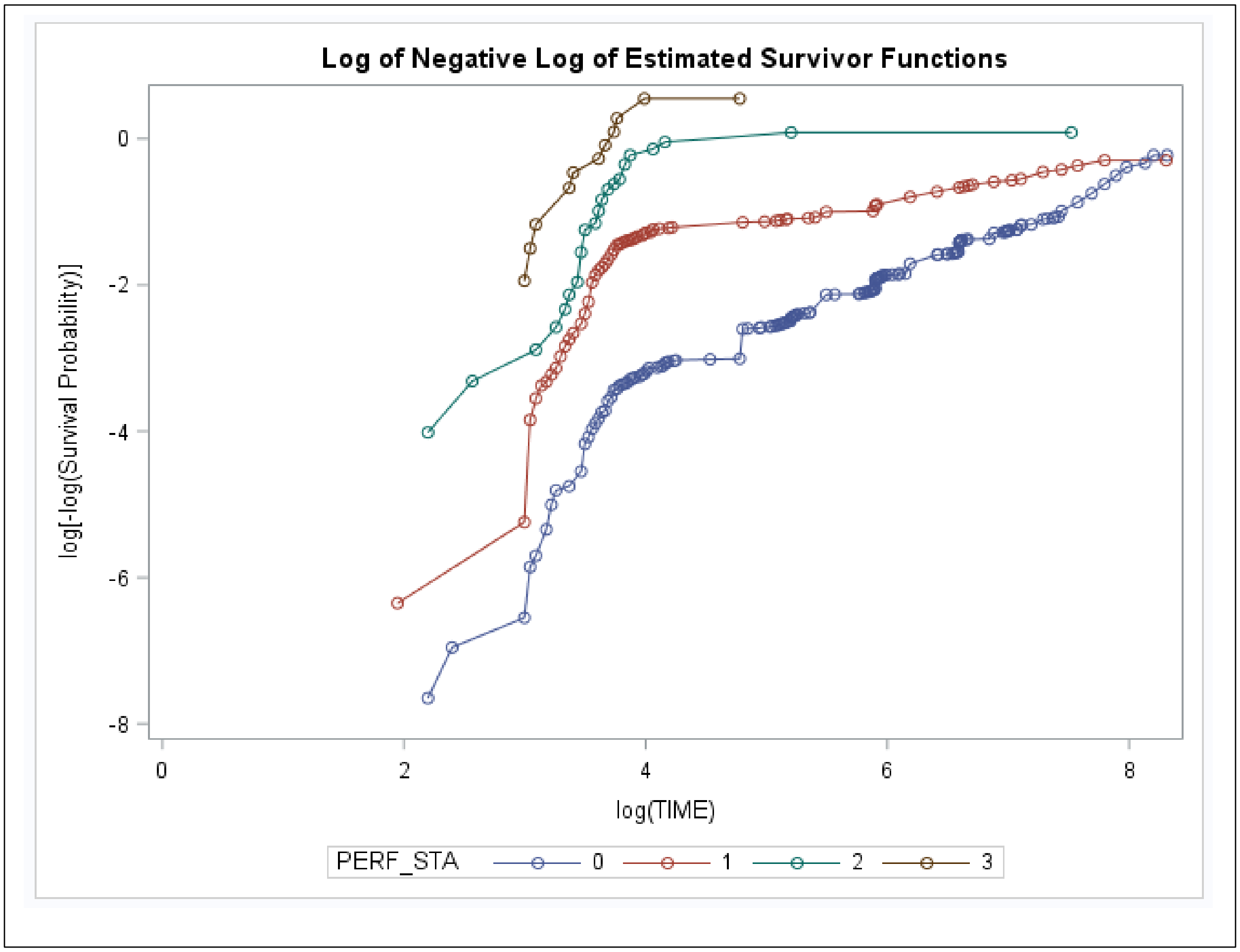

Graph is Kaplan-Meier estimate of survival function, stratified by post-graduate residency training program (PGRTP). The median time to the first missing instrument was 2678 days for subjects at centres with a post-graduate residency training program (PGRTP = 1). The median time to the first missing instrument was not observed for subjects at centres without a post-graduate residency training program (PGRTP = 0). Log-rank: Chi-square = 48.9, 1 degree of freedom, $p < 0.0001$. The time from randomization to first missing instrument after randomization was significantly shorter for subjects at a centre with a post-graduate residency training program.

**Assess the proportional hazards assumption with Kaplan-Meier estimates of survival function**

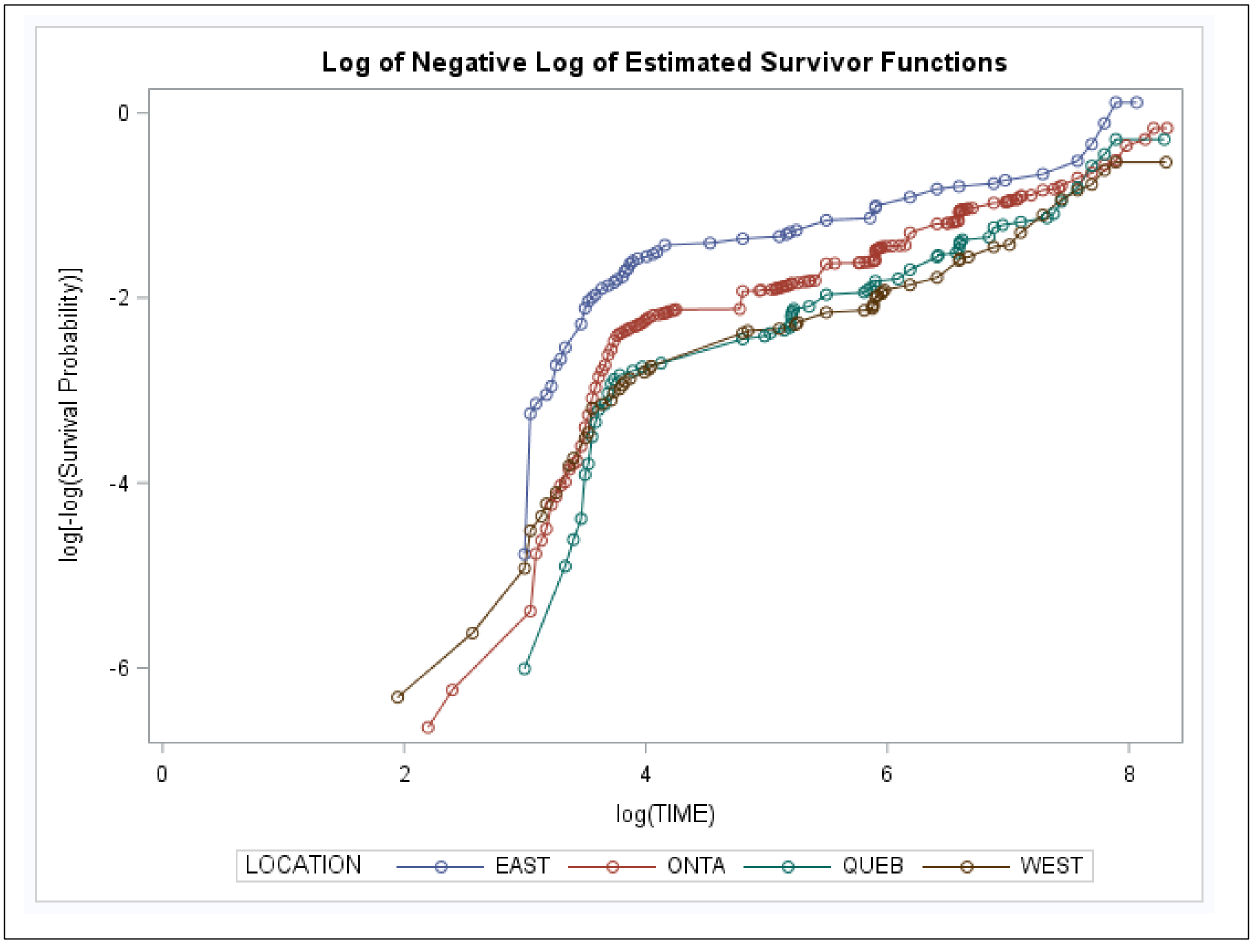
Cox regression assumes that the hazard rate in the comparison group is a constant proportion of the hazard rate in the reference group, at any time point. Accordingly, the hazard ratio does not vary with time. With reference to the previous three figures, performance status and post-graduate residency training program appear to adhere to the proportional hazards assumption. However, with location, the graphs appear to be parallel before 1500 days but afterwards, the graphs for WEST and QUEB appear to cross at about 1500 days, and the graphs for WEST and ONTA appear to cross at about 2700 days, suggesting violation of the assumption.

**Assess the proportional hazards assumption with log-minus-log survival plots**

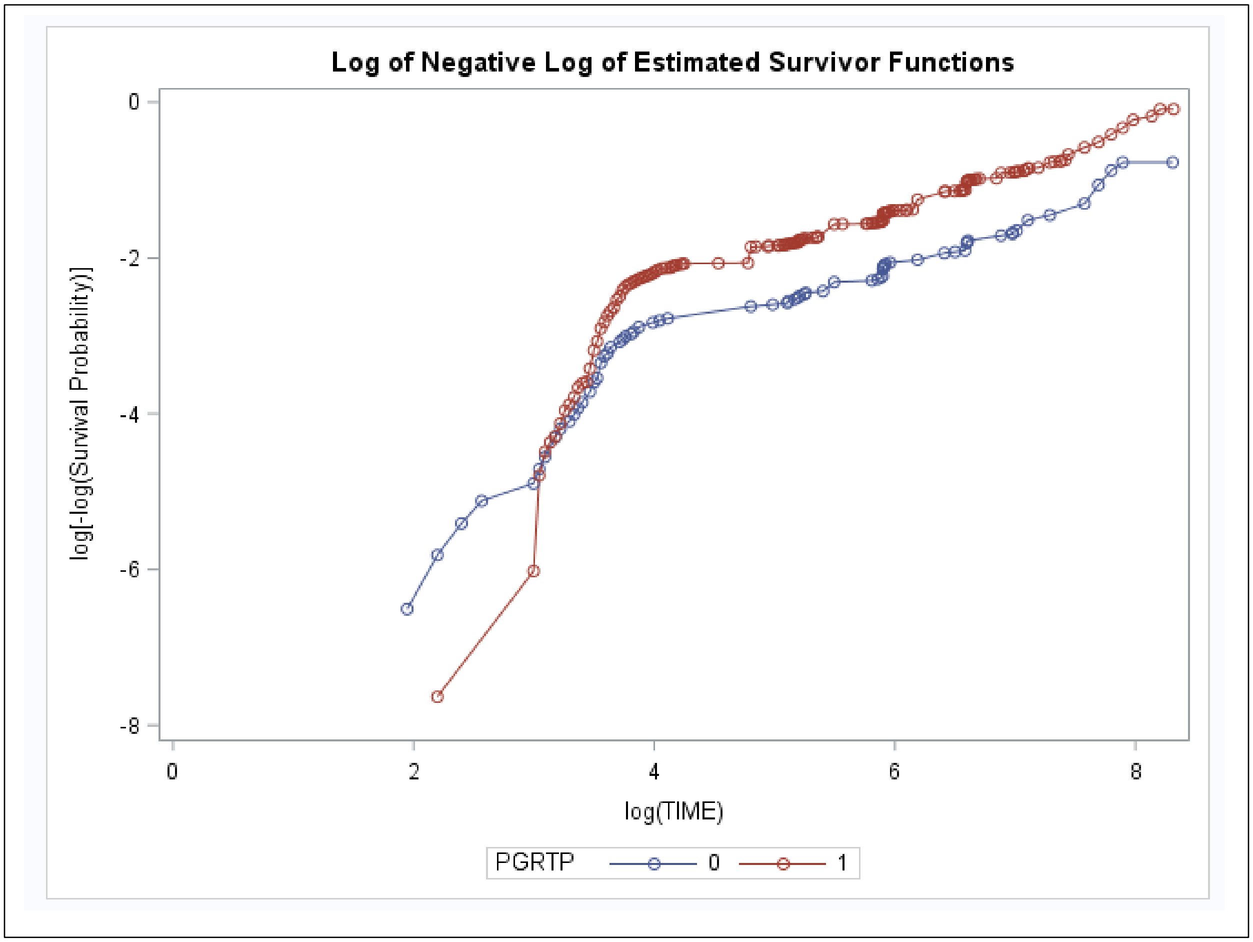
Log-minus-log survival plots are provided for performance status, centre location, and post-graduate residency training program on the next three. Proportionality assumption for AGE was not assessed by log-minus-log survival plot, because a continuous representation was used in model. Parallel lines indicate hazard rates are proportional across strata.




Graph is log-minus-log survival plot for performance status.

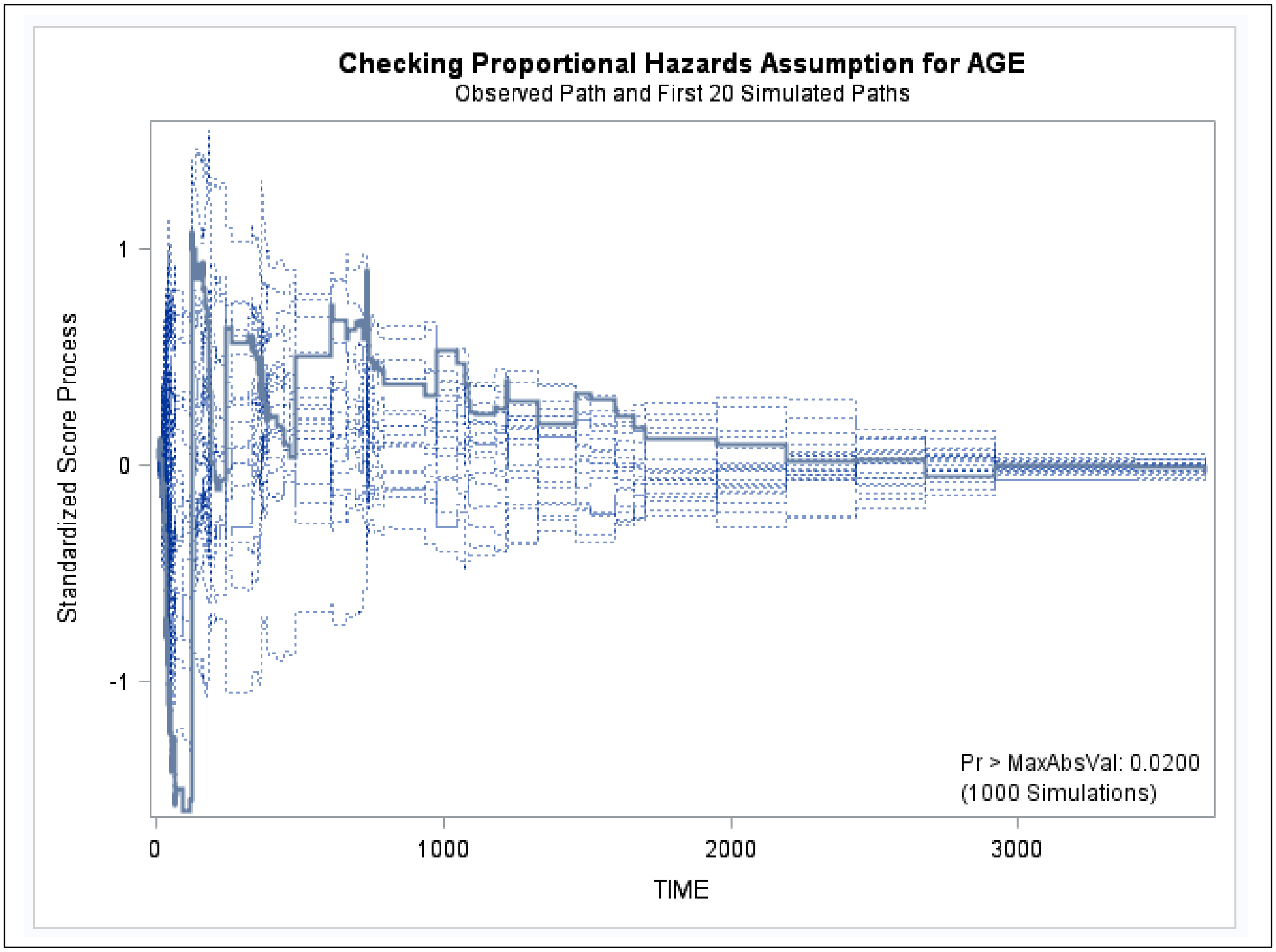



Graph is log-minus-log survival plot for centre location.

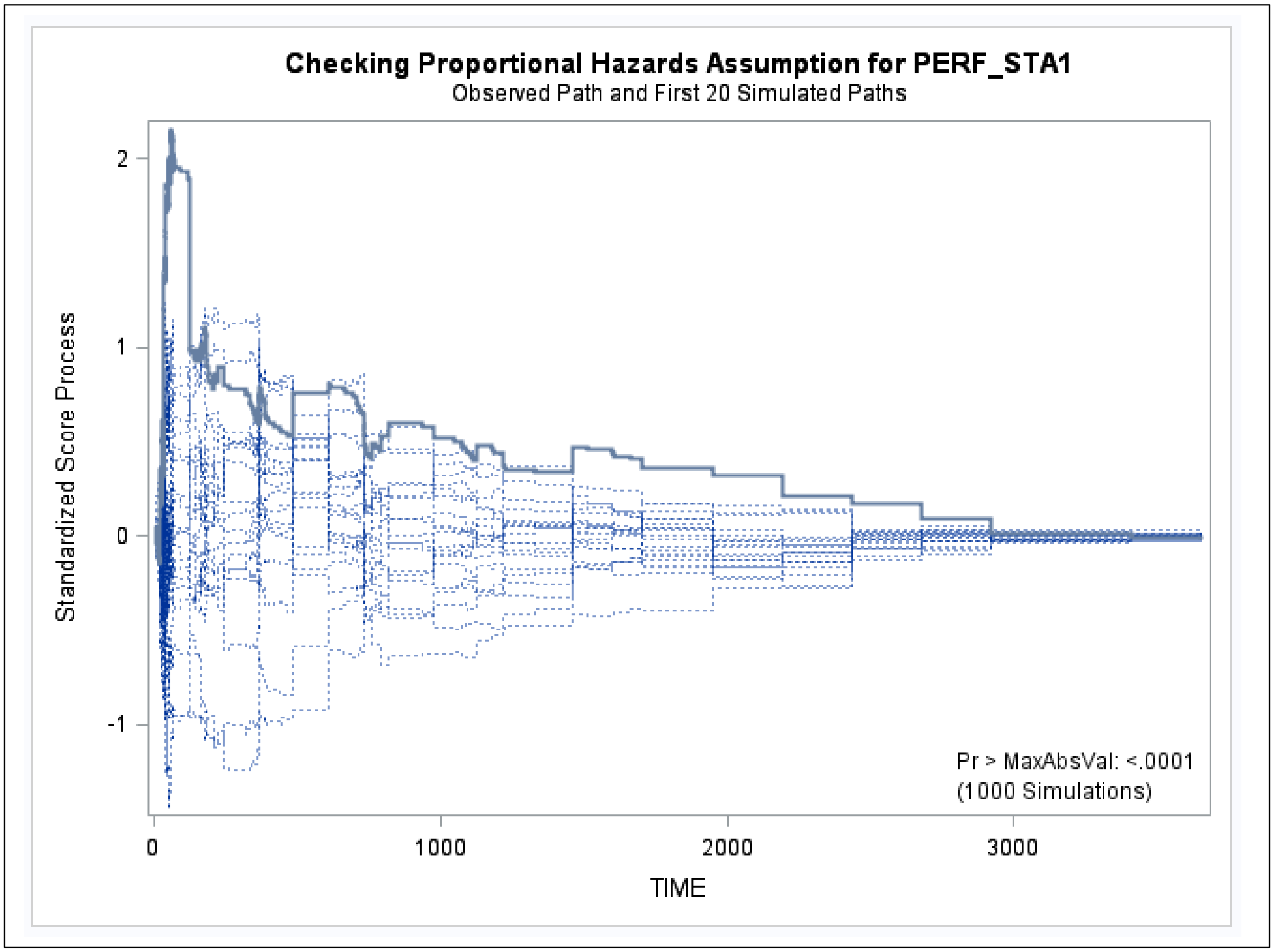



Graph is log-minus-log survival plot for post-graduate residency training program.

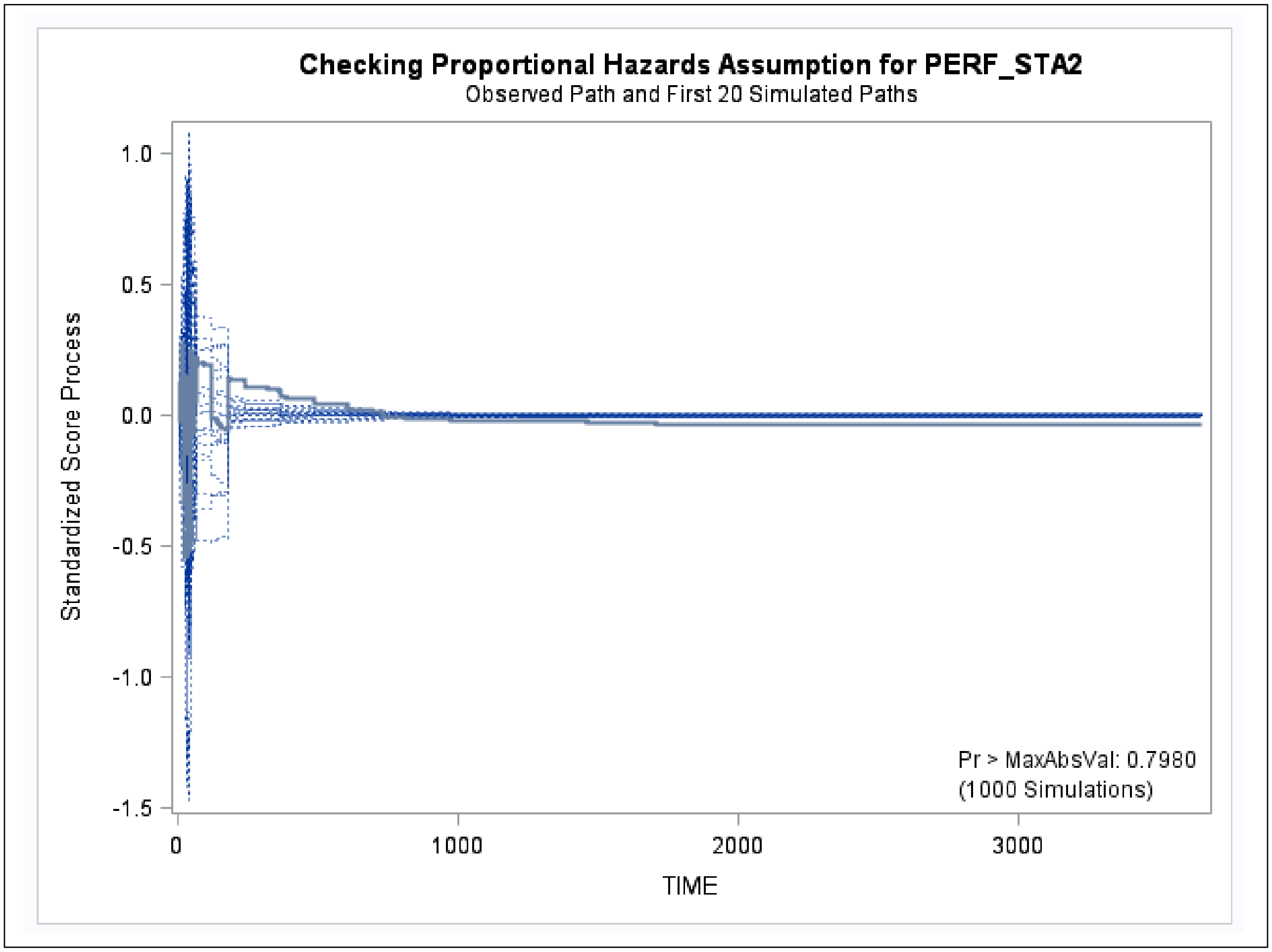
**Assess the proportional hazards assumption with standardized score processes and supremum test**



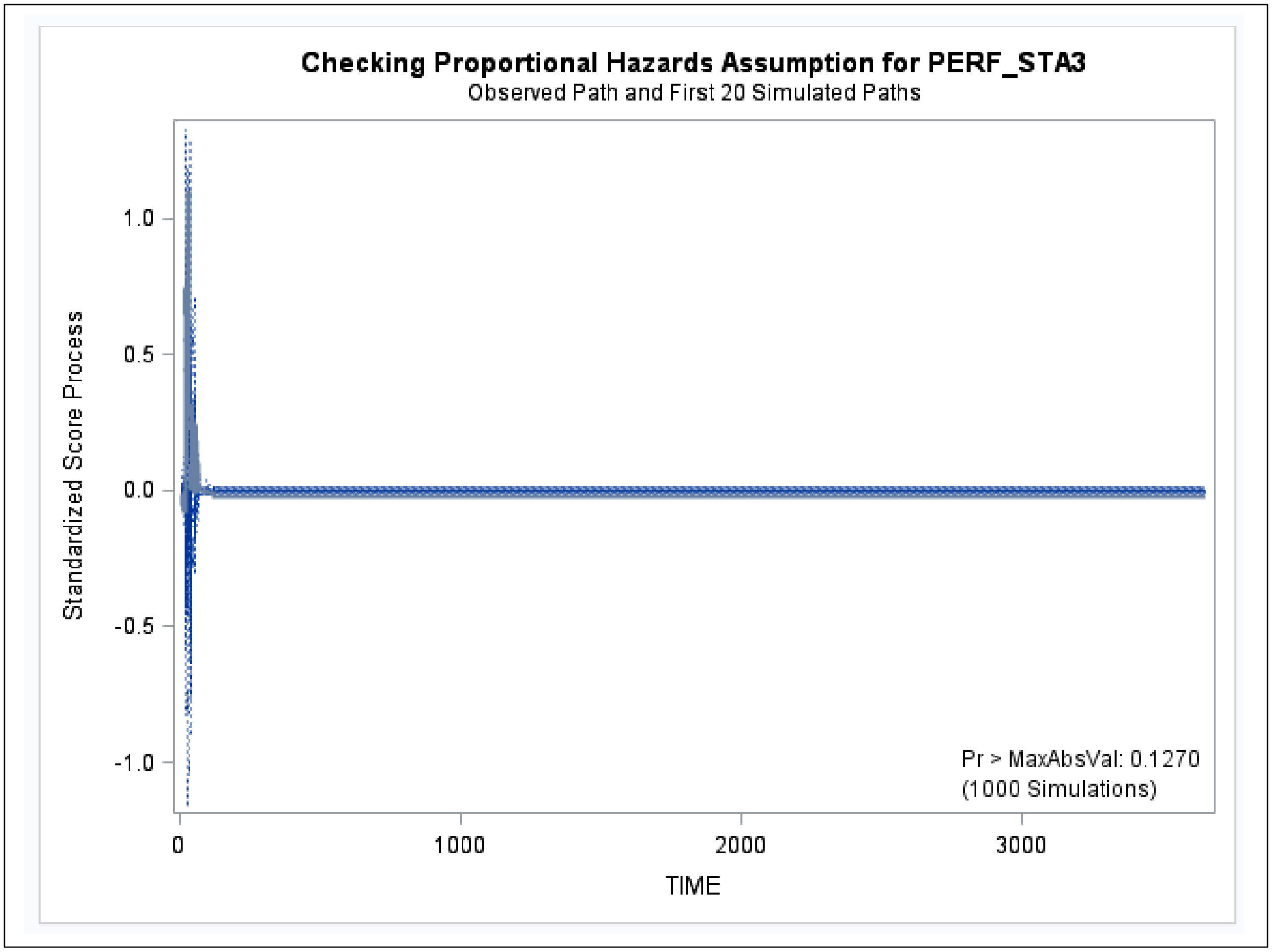

Possible departure from proportional hazards for AGE. The solid blue line represents the observed score. The dotted lines represent 20 simulated scores. Solid line is below lowest dotted line shortly after TIME = 0, but then remains within dotted lines for subsequent follow-up.



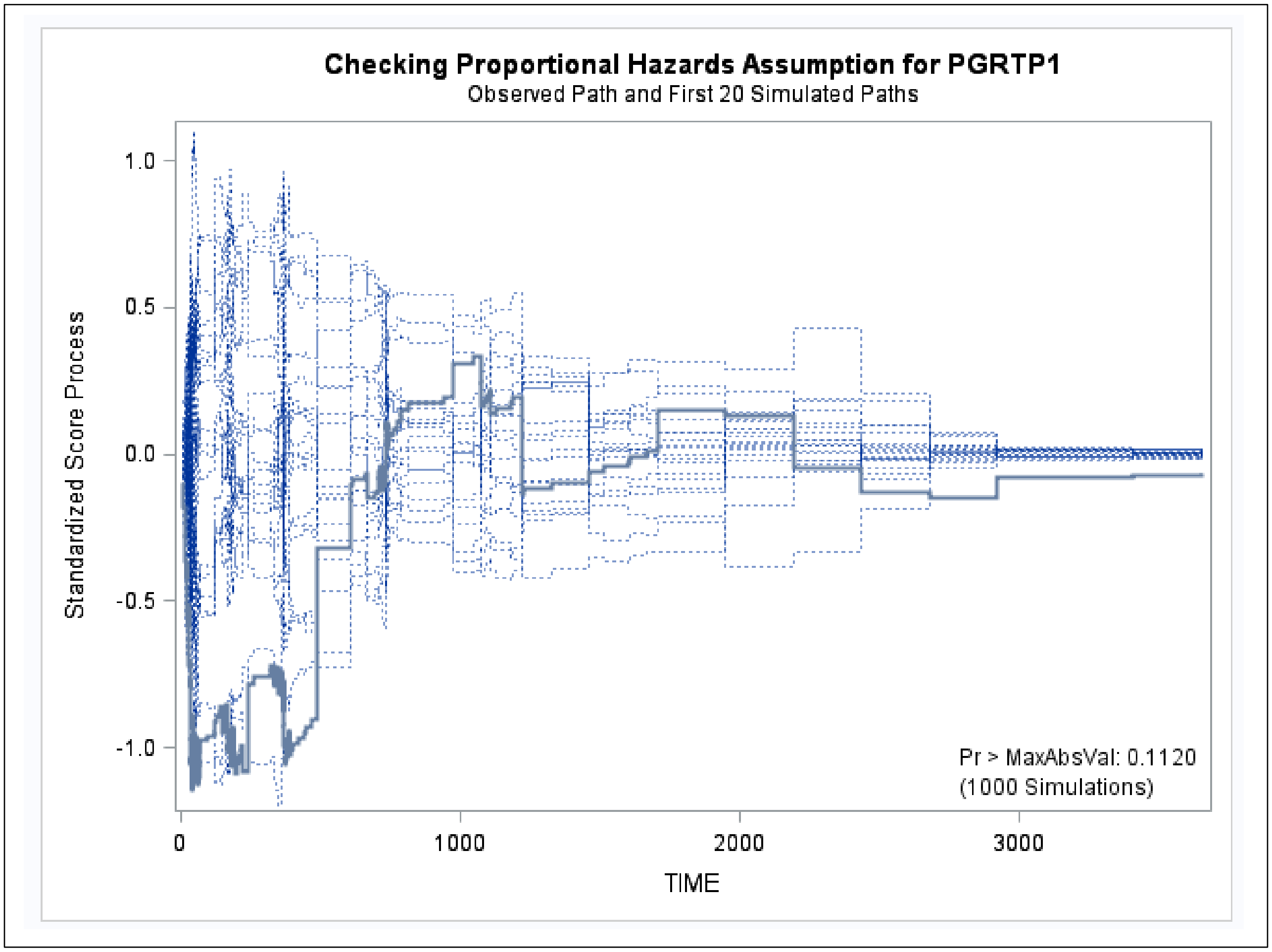

Possible departure from proportional hazards for ECOG = 1. Solid line is above highest dotted line shortly after TIME = 0, is within dotted lines until TIME = 1500 days, then is subsequently above highest dotted line.



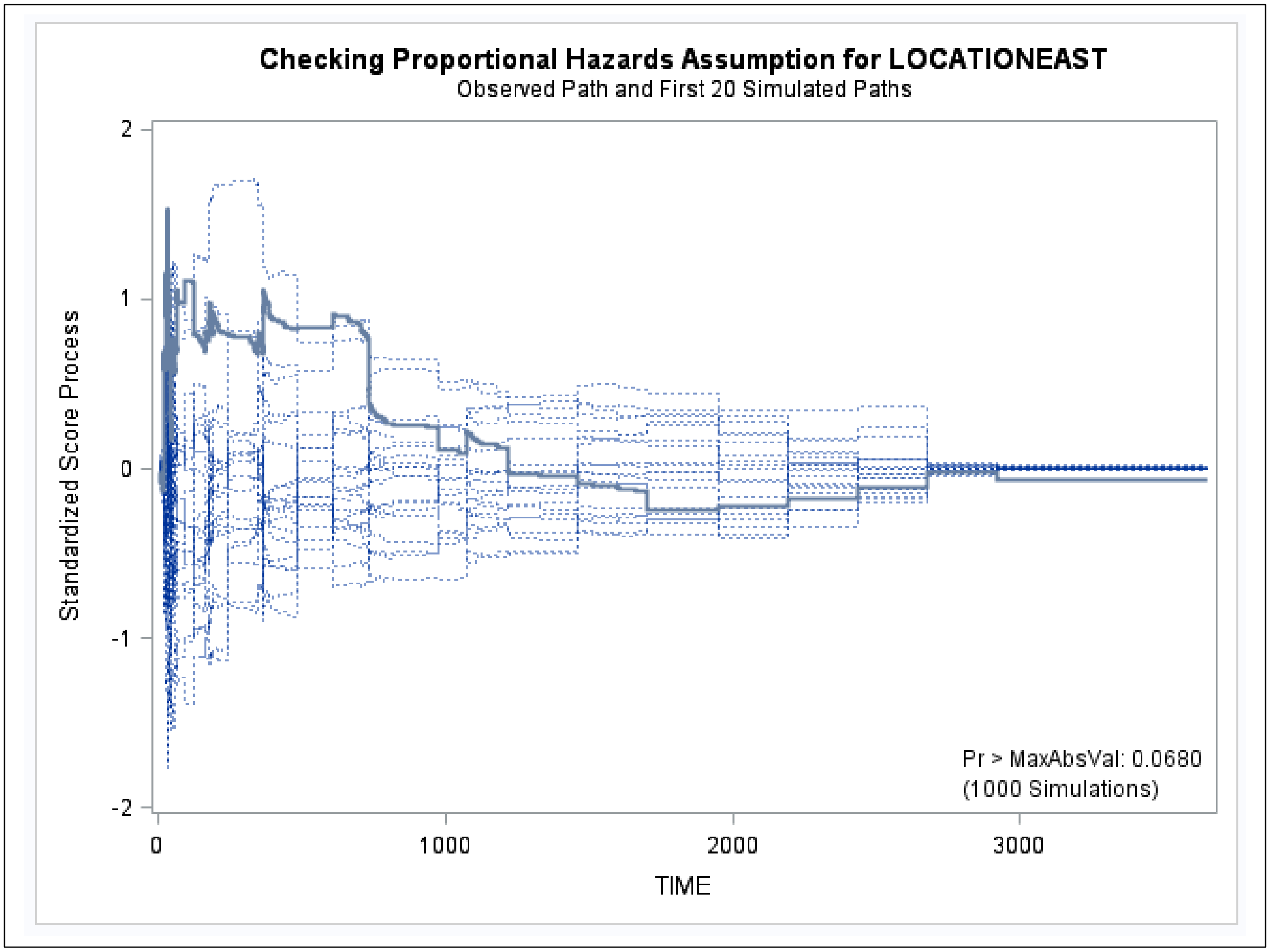

No departure from proportional hazards for ECOG = 2.



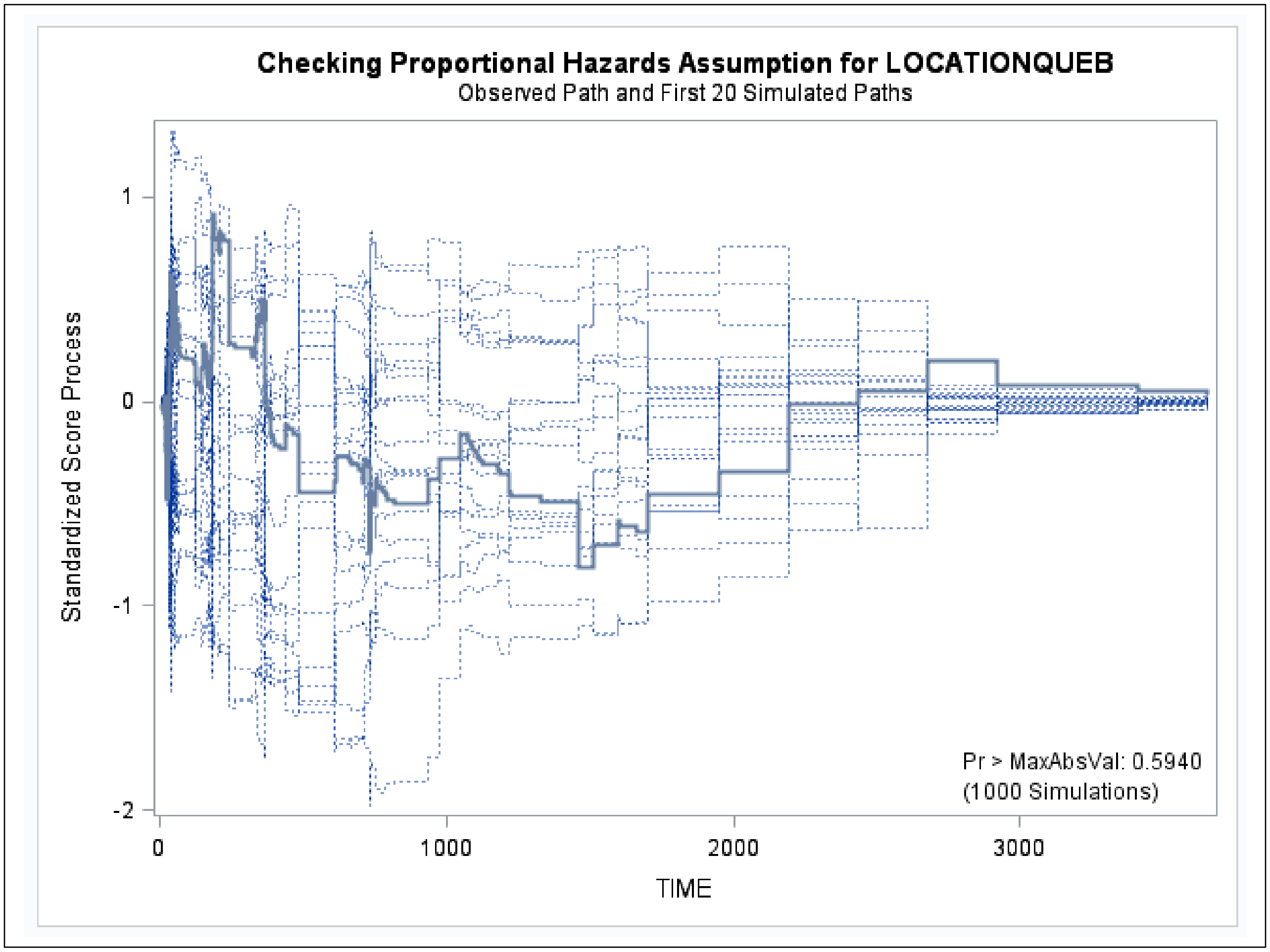

No departure from proportional hazards for ECOG = 3.




No departure from proportional hazards for post-graduate residency training program.

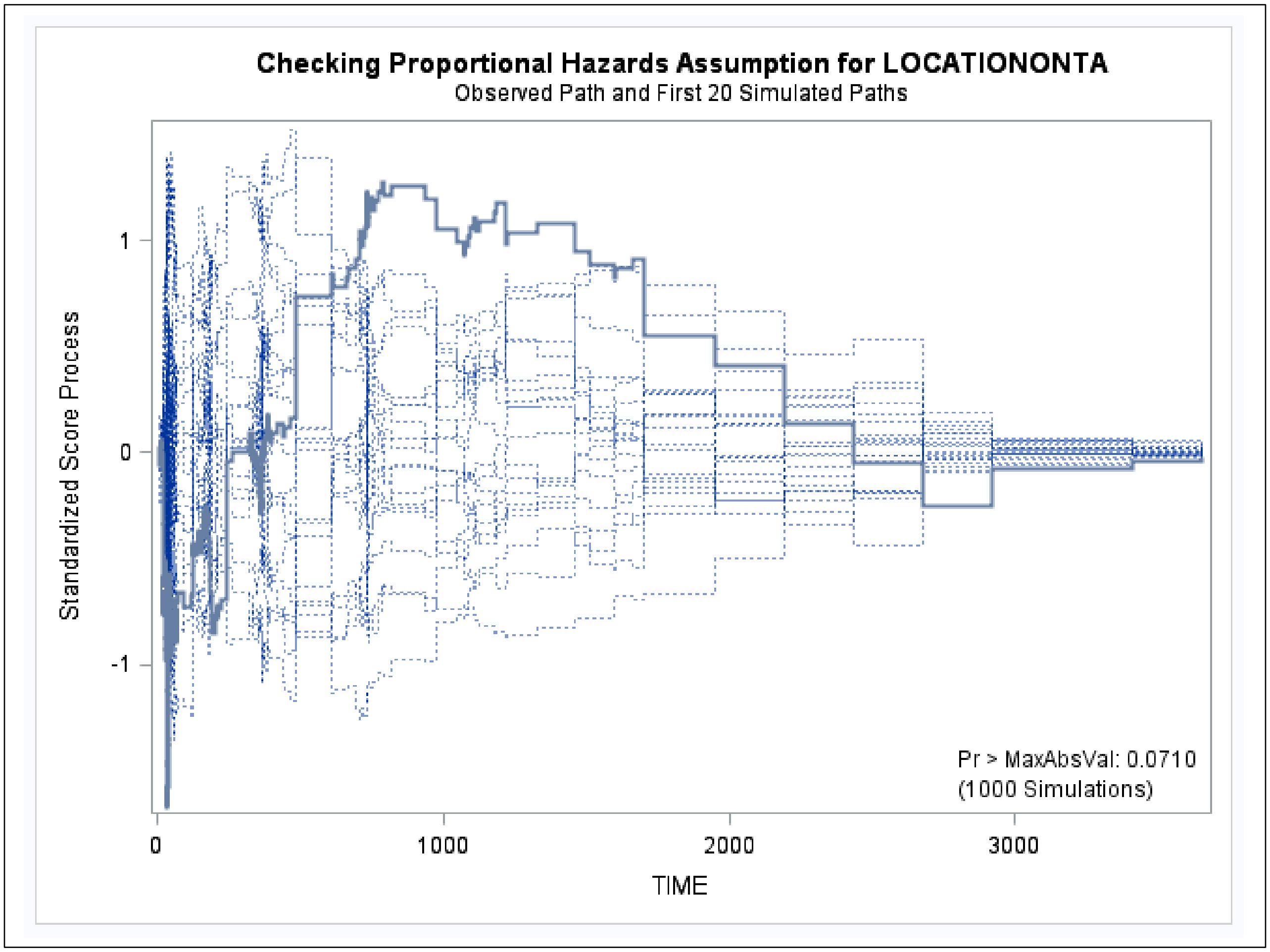



No departure from proportional hazards for centre location in Atlantic provinces.




No departure from proportional hazards for centre location in Quebec.

.




Possible departure from proportional hazards for centre location in Ontario. Solid line is below lowest dotted line shortly after TIME = 0, is within dotted lines until TIME = 800 days, is above highest dotted line until TIME = 1500 days, then is subsequently within dotted lines.

.

| Suprenum Test for Proportional Hazards Assumption | | | |
|---|---|---|---|
| Variable | Maximum Absolute Value | Number of Replications | p value |
| Age | 1.5960 | 1000 | 0.0200 |
| ECOG 1 | 2.1523 | 1000 | <.0001 |
| ECOG 2 | 0.5428 | 1000 | 0.7980 |
| ECOG 3 | 1.0989 | 1000 | 0.1270 |
| Post-graduate residency training program | 1.1463 | 1000 | 0.1120 |
| Atlantic provinces | 1.5360 | 1000 | 0.0680 |
| Ontario | 1.6732 | 1000 | 0.0710 |
| Quebec | 0.9167 | 1000 | 0.5940 |

The supremum test was significant for AGE ($p = 0.020$) and for ECOG performance status = 1 ($p < 0.0001$).

The supremum test was not significant for ECOG = 2 or ECOG = 3, post-graduate residency training program, or centre location in Atlantic provinces, Ontario, or Quebec.

## References

1. Little RL, D'Agostino R, Cohen ML et al. The prevention and treatment of missing data in clinical trials. *N Engl J Med* 2012; 367:1355-1361.

2. Bangdiwala SI, Bhargava A, O'Connor DP, et al. Statistical methodologies to pool across multiple intervention studies. *TBM* 2016;6:228-235.

3. Palmer MJ, Richardson H, Tu D, Brundage M. Factors predicting missing instruments in three cancer randomized clinical trials. *Qual Life Res* 2019; submitted.

4. Whelan TJ et al. Assessment of quality of life in MA.17, a randomized, placebo-controlled trial of letrozole after 5 years of tamoxifen in postmenopausal women. *J Clin Oncol* 2005;23:6931-6940.

5. Crook JM et al. Intermittent androgen suppression for rising PSA level after radiotherapy. *N Engl J Med* 2012;367:895-903.

6. Crump M et al. Randomized comparison of gemcitabine, dexamethasone, and cisplatin versus dexamethasone, cytarabine, and cisplatin chemotherapy before autologous stem-cell transplantation for relapsed and refractory aggressive lymphomas: NCIC CTG LY.12. *J Clin Oncol* 2014;32:3490-3496.

7. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;327:557-560.

8. Sedgwick P. Meta-analyses – what is heterogeneity? *BMJ* 2015;350:h1435.

9. Burke DL, Ensor J, Riley RD. Meta-analysis using individual participant data: one-stage and two-stage approaches, and why they may differ. *Stat Med* 2017;36:855-877.

10. Blettner M, Sauerbrei W, Schlehofer B, Scheuchenpflug T, Friedenreich C. Traditional reviews, meta-analyses and pooled analyses in epidemiology. *Int J Epidemiol* 1999;28:1-9.

11. Friedenreich CM. Commentary: improving pooled analyses in epidemiology. *Int J Epidemiol* 2002;31:86-87.

12. Sedgwick P. How to read a Kaplan-Meier survival plot. *BMJ* 2014;349:g5608.

13. Vittinghoff E, Glidden DV, Shiboski SC, McCulloch CE. *Regression methods in statistics. Linear, logistic, survival, and repeated measures models. Second edition*. Springer, 2012.

14. UCLA Institute for Digital Research and Education. Introduction to survival analysis in SAS. From https://stats.idre.ucla.edu/sas/seminars/sas-survival/

15. Rothman KL, Greenland S, Lash TL. *Modern epidemiology, third edition*. Lippincott Williams & Wilkins: Philadelphia, 2008.

16. Osoba D, Zee B. Completion rates in health-related quality-of-life assessment: approach of the National Cancer Institute of Canada Clinical Trials Group. *Stat Med* 1998;17:603-612.

17. Hahn HA, Webster KA, Cella D, Fairclough DL. Missing data in quality of life research in Eastern Cooperative Oncology Group (ECOG) clinical trials: problems and solutions. *Stat Med* 1998;17:547-559.

18. Barraclough H, Simms L, Govindan G. What a clinician ought to know: hazard ratios. *J Thorac Oncol* 2011;6:978-982.

19. Austin PC. Statistical power to detect violation of the proportional hazards assumption when using the Cox regression model. *J Stat Comput Simul* 2018;88:533-552.

## Appendix D

## Missing items precluded scoring the sexual scale of the menopause-specific quality of life questionnaire for some subjects at baseline.

Michael J Palmer, Harriet Richardson, Michael Brundage

Initial quantitative analyses were performed on baseline individual participant data in Part 2. The results are formatted as a research letter for submission to *Quality of Life Research.*

Missing items occur when a respondent provides answers to some but not all questions on an instrument, potentially leading to loss of power and bias in the analysis of patient-reported outcome (PRO) data from randomized clinical trials [1]. Reports from the scientific literature suggest missing items may associate with specific questions, such as those pertaining to religious issues [2], feelings of self-consciousness [3], or sexual relationships [4, 5].

The presence of missing items may preclude scoring of a scale. For example, the Menopause-Specific Quality of Life Questionnaire (MENQOL) [6] consists of four scales, with each scale including several questions: vasomotor (questions 1 – 3), psychosocial (4-10), physical (11-26), and sexual (27-29). The respondent is expected to indicate with a No or Yes response, for each question, if they experienced the problem in the past month, and if they had experienced the problem (i.e. Yes response), to subsequently rate, on a scale from 0 (not at all bothered) to 6 (extremely bothered), how much they had been bothered by the problem. Obtaining a score for a scale involves assigning a positive integer ranging from 1 (representing subject had not experienced the problem) to 8 (representing subject had experienced the problem and was extremely bothered by it) for each response, adding up the responses in the scale, and calculating the mean. A given scale cannot be scored if more than one half of the items are missing [7]. This convention is especially concerning for the baseline assessment, because subjects without a scale score at time of randomization cannot be included in an analysis of change scores from baseline.

Individual participant data for a sub-group of women with breast cancer in Canada from a randomized clinical trial [8] provided opportunity to explore, in subjects with a non-missing Short Form-36 Health Survey [9] and MENQOL [6] at the time of randomization, the association of different variables with missing items that preclude scoring of a scale. The outcome, defined as more than one-half of the items comprising a scale were missing, was expressed as one (yes) or zero (no) for each scale, and assigned to each subject. A systematic approach to analysis examined the outcome of every scale in the MENQOL. First, the percentage of subjects with each outcome was determined. Then, if more than 10% of subjects had the outcome, log-binomial regression was used, to estimate the relative risk and 95% confidence interval, and to build parsimonious models from six variables for Participant (age, nodal status, performance status, and baseline PRO scores for bodily pain, vitality, and general health from the Short Form-36 Health Survey) and from three variables for Centre (level of activity, presence of a post-graduate residency training program in oncology or hematology, and location). Analyses were performed with Statistical Analysis System version 9.4 (Cary, NC).

**Figure D-1** presents missing items for the MENQOL in the baseline assessment. The highest frequencies of missing items were in the sexual scale, involving question 27 (‘Change in your sexual desire’, n = 110), question 28 (‘Vaginal dryness during intercourse’, n = 194), and question 29 (‘Avoiding intimacy’, n = 176). Missing items that precluded scoring of the sexual scale were observed for 12% of subjects, which surpassed the *a priori* criterion.

Proceeding to log-binomial regression, the model for Participant had two covariates. A one-year increase in age was associated with a 1.06 increase in the relative risk of having at least two missing responses in the sexual scale (95% Confidence Interval = 1.04, 1.07; $p < 0.0001$), while holding bodily pain constant. Stated another way, a 10-year increase in age was associated with a 1.72 increase in the relative risk (95% CI = 1.46, 2.02). A one-unit increase in bodily pain was associated with a 0.99 increase in the relative risk (95% CI = 0.99, 1.00; $p = 0.0564$), while holding age constant.

The model for Centre had one covariate, location. Using subjects at centres in Ontario as the reference, the relative risk for subjects at centres in the western provinces of Canada was 0.6567 (95% CI = 0.4314, 0.9996; $p = 0.0498$). The 95% CIs for subjects at centres in the Atlantic provinces (0.3756, 1.2208; $p = 0.1949$) and for subjects at centres in Quebec (95% CI = 0.4087, 1.0129; $p = 0.0568$) bounded unity.

A final model combined the model for Participant with the model for Centre. Significant independent association with outcome was observed for age ($p < 0.0001$), bodily pain ($p = 0.0457$), and location (subjects at centres in the western provinces versus Ontario, $p = 0.0472$). Interpreting the model in terms of log risks:

$$\ln p(x) := -5.0665(\text{intercept}) + 0.0533\,(\text{age}) - 0.0060\,(\text{bodily pain}) - 0.4167\,(\text{centre in western provinces}) - 0.3515\,(\text{centre in Quebec}) - 0.1679\,(\text{centre in Atlantic provinces})$$

One can use the model to determine the probability of a given subject having two or more missing items in the sexual scale at baseline. As an example, a subject with age 68 years and bodily pain score 72 from a centre in the western provinces would have 10% probability of the outcome ($e^{(-5.0665 + (68 \times 0.0533) - (72 \times 0.0060) - 0.4167)} = 0.1012$). In comparison, a similar subject from a centre in the Atlantic provinces would have 13% probability. These represent a 3% absolute difference but a 30% relative difference.

The Classification Framework [10] includes a factor within the component of Instrument called ‘unacceptable content’, which was based on expert opinion [5], not derived from regression modelling. This study provides new evidence from regression modelling that the age and bodily pain score of subjects at baseline, and the geographic location of the centres, may predict missing items that preclude scoring of the sexual scale in the MENQOL for some subjects at the time of randomization. It might be beneficial when designing a future randomized clinical trial having PROs collected with a similar questionnaire from adjuvant breast cancer patients, to give attention to these three factors so that a potential endpoint is not compromised by missing data.

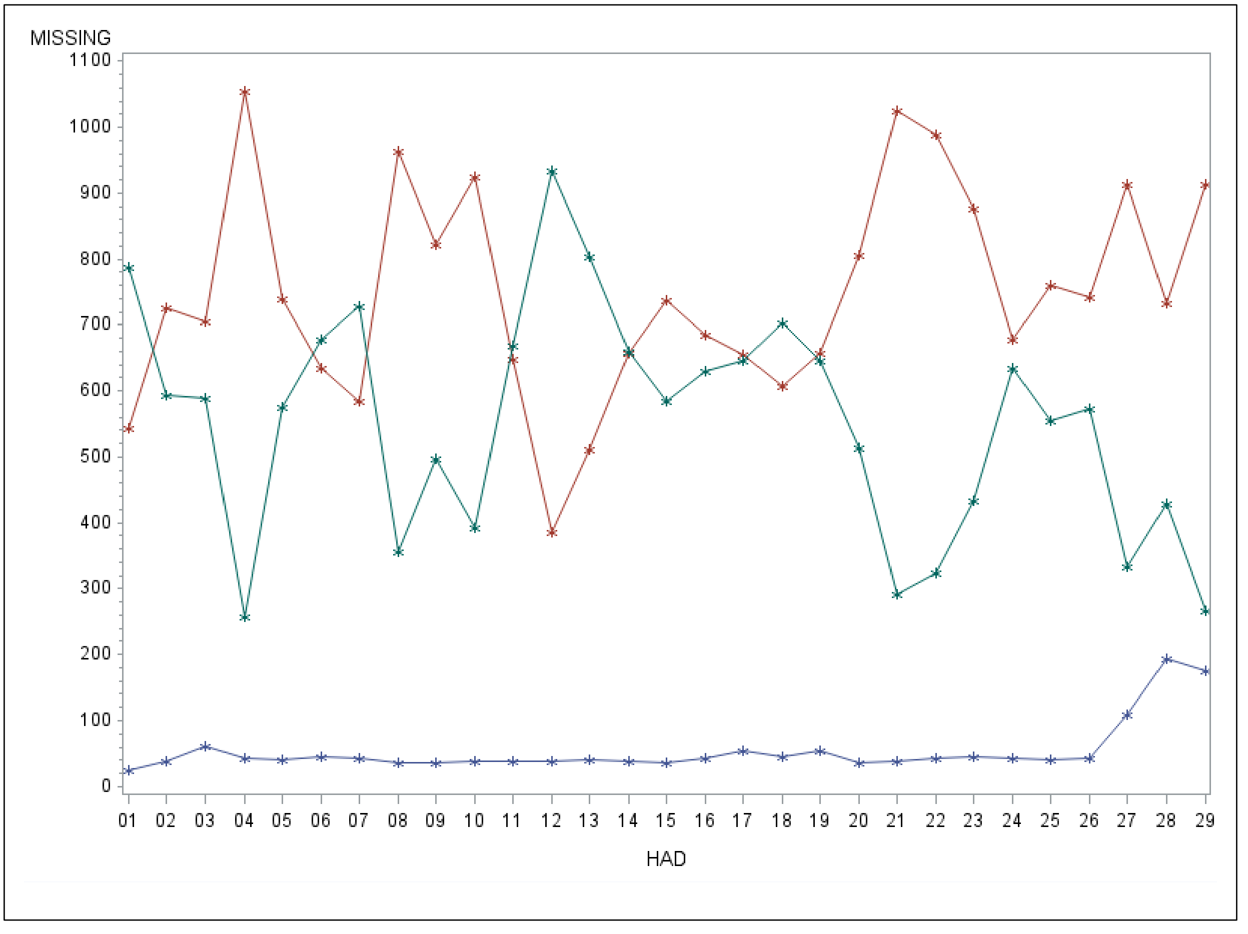



**Figure D-1.** Missing items in the Menopause-Specific Quality of Life Questionnaire for the baseline assessment. Graphs represent frequency of No (Red), Yes (Green), or Missing (Blue) responses among 1356 adjuvant breast cancer patients whether they had experienced a problem with vasomotor (items 1 – 3), psychosocial (4-10), physical (11-26), or sexual (27-29) scales.

**References**


1. Fayers PM, Curran D, Machin D. Incomplete quality of life data in randomized trials: missing items. *Stat Med* 1998;17:679-696.

2. Kassa S, Hjermstad MJ, Jordhoy MS, Wisloff F, Loge JH. Compliance in quality of life data: a Norwegian experience. *Stat Med* 1998;17:623-632.

3. Kiebert GM, Curran D, Aaronson NK. Quality of life as an endpoint in EORTC clinical trials. *Stat Med* 1998;17:561-569.

4. Mosconi P et al. The multi-centre assessment of quality of life: the Interdisciplinary Group for Cancer Care Evaluation (GIVIO) experience in Italy. *Stat Med* 1998;17:577-585.

5. Fallowfield L. Compliance issues in quality of life assessment: experiences of two cancer research campaign sponsored groups. *Stat Med* 1998;17:541-546.

6. Hilditch JR, Lewis JE, Peter A, et al. A menopause-specific quality of life questionnaire: development and psychometric properties. *Maturitas* 1996;24:161-175.

7. Lewis JE, Hilditch JR, Wong CJ. Further psychometric property development of the menopause-specific quality of life questionnaire and development of a modified version, MENQOL-Intervention questionnaire. *Maturitas* 2005;50:209-221.

8. Whelan TJ et al. Assessment of quality of life in MA.17, a randomized, placebo-controlled trial of letrozole after 5 years of tamoxifen in postmenopausal women. *J Clin Oncol* 2005;23:6931-6940.

9. Ware J, Snow KK, Kosinksi M et al. *SF-36 health survey manual and interpretation guide*. The Health Institute: Boston, 1993.

10. Palmer MJ, Mercieca-Bebber R, King M, Calvert M, Richardson H, Brundage M. A systematic review and development of a classification framework for factors associated with missing patient-reported outcome data. *Clinical Trials* 2018;15:95-106.

# Appendix E

## E-Supplements for the three main manuscripts

## Appendix E.1

## E-Supplements for the manuscript presented in Chapter 3

List of the included papers in the systematic review and Summary table of the source journals are provided on the following pages.

The following list identifies the 100 papers that met the eligibility criteria.


1. Aaron LA, Mancl L, Turner JA, et al. Reasons for missing interviews in the daily electronic assessment of pain, mood and stress. *Pain* 2004;109:389-398.

2. Aaronson NK. Quality of life research in cancer clinical trials: a need for common rules and language. *Oncology* 1990;4:59-66.

3. Ahlner-Elmqvist M, Bjordal K, Jordhoy MS, et al. Characteristics and implications of attrition in health-related quality of life studies in palliative care. *Palliat Medicine* 2009;23:432-440.

4. Altman DG, Bland JM. Missing data. *BMJ* 2007;34:424.

5. Badgwell B, Krouse R, Cormier J, et al. Frequent and early death limits quality of life assessment in patients with advanced malignancies evaluated for palliative surgical intervention. *Ann Surg Oncol* 2012;19:3651-3658.

6. Bae K, Bruner W, Baek S, et al. Patterns of missing mini mental status exam (MMSE) in radiation therapy oncology group (RTOG) brain cancer trials. *J Neurooncol* 2011;105:383-395.

7. Basch E. New frontiers in patient-reported outcomes: adverse event reporting, comparative effectiveness, and quality assessment. *Annu Rev Med* 2014;65:307-317.

8. Belle SH, Porayko MK, Hoofnagle JH, et al. Changes in quality of life after liver transplantation among adults. *Liver Transpl Surgery* 1997;3:93-104.

9. Bernhard J, Gusset H, Hurny C. Quality-of-life assessment in cancer clinical trials: an intervention by itself? *Support Care Cancer* 1995;3:66-71.

10. Bernhard J, Cella D, Coates AS, et al. Missing quality of life data in cancer clinical trials: serious problems and challenges. *Stat Med* 1998;17:517-532.

11. Bernhard J, Peterson HF, Coates AS, et al. Quality of life assessment in International Breast Cancer Study Group (IBCSG) trials: practical issues and factors associated with missing data. *Stat Med* 1998;17:587-601.

12. Bernhard J, Gusset H, Hurny C. Practical issues in quality of life assessment in multicentre trials conducted by the Swiss Group for Clinical Cancer Research. *Stat Med* 1998;17:633-639.

13. Bezjak A, Lee CW, Ding K, et al. Quality-of-life outcomes for adjuvant chemotherapy in early-stage non-small-cell lung cancer: results from a randomized trial, JBR10. *J Clin Oncol* 2008;26:5052-5059.

14. Cirrincione R, Taliana G, Caporaso N, et al. Quality of life assessment in chronic liver disease. *Hepato Gastroenterol* 2002;49:813-816.

15. Courvoisier DS, Eid M, Lischetzke T. Compliance to a cell phone-based ecological momentary assessment study: the effect of time and personality characteristics. *Psychol Assess* 2012;24:713-720.

16. Curran D, Bacchi M, Hsu Schmitz F, et al. Identifying the types of missingness in quality of data from clinical trials. *Stat Med* 1998;17:739-756.

17. De Graaf R, Van Dorsselaer S, Tuithof M, et al. Sociodemographic and psychiatric predictors of attrition in a prospective psychiatric epidemiological study among the general population. Result of the Netherlands Mental health Survey and Incidence Study-2. *Compr Psychiatry* 2013;54:1131-1139.

18. DeSousa PCB, Mendes FM, Imparato JCP, et al. Differences in responses to the Oral Health Impact Profile (OHIP 14) used as a questionnaire or in an interview. *Braz Oral Res* 2009;23:358-364.

19. Di Gessa G, Grundy E. The relationship between active ageing and health using longitudinal data from Denmark, France, Italy and England. *J Epidemiol Community Health* 2014;68:261-267.

20. Di Maio M, Perrone F. Quality of life in elderly patients with cancer. *Health Qual Life Outcomes* 2003;1:44.

21. Eich D, Ajdacic-Gross V, Condrau M, et al. The Zurich study: participation patterns and Symptom Checklist 90-R scores in six interviews, 1979-99. *Acta Psychiatr Scand* 2003;108 (Suppl 418):11-14.

22. Fallowfield L. Compliance issues in quality of life assessment: experiences of two cancer research campaign sponsored groups. *Stat Med* 1998;17:541-546.

23. Fang F-M, Tsai W-L, Chien C-Y, et al. Changing quality of life in patients with advanced head and neck cancer after primary radiotherapy or chemoradiation. *Oncology* 2005;68:405-413.

24. Fayers PM, Hopwood P, Harvey A, et al. Quality of life assessment in clinical trials – guidelines and a checklist for protocol writers: the U. K. Medical Research Council experience *Eur J Cancer* 1997;33:20-28.

25. Fernandez-Ballesteros R, Zamarron MD, Diez-Nicholas J, et al. Mortality and refusal in the longitudinal 90+ project. *Arch Gerontol Geriatr* 2011;53:e203-e208.

26. Gaertner J, Elsner F, Pollmann-Dahmen K, et al. Electronic pain diary: a randomized crossover study. *J Pain Symptom Manage* 2004;28:259-267.

27. Ganz PA, Haskell CM, Figlin RA, et al. Estimating the quality of life in a clinical trial of patients with metastatic lung cancer using the Karnofsky performance status and Functional Living Index – Cancer. *Cancer* 1988;61:849-856.

28. Ganz PA, Day R, Costantino J. Compliance with quality of life data collection in the National Surgical Adjuvant Breast and Bowel Project (NSABP) breast cancer prevention trial. *Stat Med* 1998;17:613-622.

29. Ganz PA, Guadagnoli E, Landrum MB, et al. Breast cancer in older woman: quality of life and psychosocial adjustment in the 15 months after diagnosis. *J Clin Oncol* 2003;21:4027-4033.

30. Garcia M, Fernandez E, Schiaffino A, et al. Attrition in a population-based cohort eight years after baseline interview: the Cornella health interview survey follow-up (CHIS.FU) study. *Ann Epidemiol* 2005;15:98-104.

31. Greenhalgh J. An assessment of the feasibility and utility of the MS Symptom and Impact Diary (MSSID). *Qual Life Res* 2005;14:1363-1374.

32. Groenvold M. Methodological issues in the assessment of health-related quality of life in palliative care trials. *Acta Anaesthesiol Scand* 1999;43:948-953.

33. Guadagnoli E, Cleary PD. Age-related item nonresponse in surveys of recently discharged patients. *J Gerontol Psychol Sci* 1992;47:206-212.

34. Hahn HA, Webster KA, Cella D, et al. Missing data in quality of life research in Eastern Cooperative Oncology Group (ECOG) clinical trials: problems and solutions. *Stat Med* 1998;17:547-559.

35. Hansten ML, Downey L, Rosengren DB, et al. Relationship between follow-up rates and treatment outcomes in substance abuse research: more is better but when is ‘enough” enough? *Addiction* 2000;95:1403-1416.

36. Hewitt CE, Kumaravel B, Dumville JC, et al. Assessing the impact of attrition in randomized controlled trials. *J Clin Epidemiol* 2010;63:1264-1270.

37. Higginson I, Wade A, McCarthy M. Palliative care: views of patients and their families. *BMJ* 1990;301:277-281.

38. Hopwood P. Quality of life assessment in chemotherapy trials for non-small cell lung cancer: are theory and practice significantly different? *Semin Oncol* 1996;23 (5, Suppl 10):60-64.

39. Hopwood P, Harvey A, Davies J, et al. Survey of the administration of quality of life questionnaires in three multicentre randomised trials in cancer. *Eur J Cancer* 1997;34:49-57.

40. Huntington JL, Dueck A. Handling missing data. *Curr Prob Cancer* 2005;29:317-325.

41. Hurny C, Bernhard J, Joss R, et al. Feasibility of quality of life assessment in a randomized phase III trial of small cell lung cancer. *Ann Oncol* 1992;3:825-831.

42. Jansen ME, Kollbaum PS, McKay FD, et al. Factors influencing the electronic capture of patient-reported contact lens performance data. *Cont Lens Anterior Eye* 2013;36:130-135.

43. Judson TJ, Bennett AV, Rogak LJ, et al. Feasibility of long-term patient self-reporting of toxicities from home via the internet during routine chemotherapy. *J Clin Oncol* 2013;31:2580-2585.

44. Kaasa S, Hjermstad MJ, Jordhoy MS, et al. Compliance in quality of life data: a Norwegian experience. *Stat Med* 1998;17:623-632.

45. Kopp I, Lorenz W, Rothmund M, et al. Relation between severe illness and non-completion of quality-of-life questionnaires by patients with rectal cancer. *J R Soc Med* 2003;96:442-448.

46. Kramer JA, Curran D, Piccart M, et al. Randomised trial of paclitaxel versus doxorubicin as first-line chemotherapy for advanced breast cancer: quality of life evaluation using the EORTC QLQ-C30 and the Rotterdam Symptom Checklist. *Eur J Cancer* 2000;36:1488-1497.

47. Lamers F, Hoogendoorn AW, Smit JH, et al. Sociodemographic and psychiatric determinants of attrition in the Netherlands study of depression and anxiety (NESDA). *Compr Psychiatry* 2012;53:63-70.

48. Land SR, Ritter MW, Costantino JP, et al. Compliance with patient-reported outcomes in multicenter clinical trials: methodologic and practical approaches. *J Clin Oncol* 2007;25:5113-5120.

49. Land SR. Missing patient-reported outcome data in an adjuvant lung cancer study. *J Clin Oncol* 2008;26:5018-5019.

50. Lindahl-Jacobsen L, Hansen DG, La Cour K, et al. Evaluation of a complex intervention to improve activities of daily living of disabled cancer patients: protocol for a randomised controlled study and feasibility of recruitment and intervention. *BMC Health Serv Res* 2014;14:194.

51. Little RJ, D'Agostino R, Cohen ML, et al. The prevention and treatment of missing data in clinical trials. *N Engl J Med* 2012;367:1355-1360.

52. Little RJ, Cohen ML, Dickersin K, et al. The design and conduct of clinical trials to limit missing data. *Stat Med* 2012;31:3433-3443.

53. Lundstrom M, Strenevi U, Thorburn W, et al. Catquest questionnaire for use in cataract surgery care: assessment of surgical outcomes. *J Cataract Refract Surg* 1998;24:968-974.

54. Marcus AC, Telesky CW. Non-participation in telephone follow-up interviews. *Am J Public Health* 1983;73:72-77.

55. Matthews FE, Chatfield M, Brayne C, et al. An investigation of whether factors associated with short-term attrition change or persist over ten years: data from the Medical Research Council Cognitive Function and Ageing Study (MRC CFAS). *BMC Public Health* 2006:6:185.

56. McEvoy PM, Burgess MM, Nathan P. The relationship between interpersonal problems, negative cognitions, and outcomes from cognitive behavioural group therapy for depression. *J Affect Disord* 2013;150:266-275.

57. McIntosh S, Pardoe H, Brown K. Effect of colorectal cancer surgery on female sexual function: a prospective cohort study. *Gastroenterol Nurs* 2013;11:29-35.

58. McQuellon RP, Thaler HT, Cella D, et al. Quality of life (QOL) outcomes from a randomized trial of cisplatin versus cisplatin plus paclitaxel in advanced cervical cancer: a Gynecologic Oncology Group study. *Gynecol Oncol* 2006;101:296-304.

59. McWhinney IR, Bass MJ, Donner A. Evaluation of a palliative care service: problems and pitfalls. *BMJ* 1994:309:1340-1342.

60. Mein G, Johal S, Grant RL, et al. Predictors of two forms of attrition in a longitudinal health study involving ageing participants: an analysis based on the Whitehall II study. *BMC Med Res Methodol* 2012;12:164.

61. Menn P, Weber N, Holle R. Health-related quality of life in patients with severe COPD hospitalized for exacerbations – comparing EQ-5D, SF-12 and SGRQ. *Health Qual Life Outcomes* 2010;8:39.

62. Meraner V, Gamper E-M, Grahmann A, et al. Monitoring physical and psychosocial symptom trajectories in ovarian cancer patients receiving chemotherapy. *BMC Cancer* 2012;12:77.

63. Moinpour CM. Do quality of life assessments make a difference in the evaluation of cancer treatments? *Control Clin Trials* 1997;18:311-317.

64. Moinpour CM, Lovato LC. Ensuring the quality of quality of life data: the Southwest Oncology Group experience. *Stat Med* 1998;17:641-651.

65. Moinpour CM, Triplett JS, McKnight B, et al. Challenges posed by non-random missing quality of life data in an advanced-stage colorectal cancer clinical trial. *Psychooncology* 2000;9:340-354.

66. Morgan K, Thompson J, Dixon S, et al. Predicting longer-term outcomes following psychological treatment for hypnotic-dependent chronic insomnia. *J Psychosom Res* 2003;54:21-29.

67. Mosconi P, Torri V, Cifani S, et al. The multi-centre assessment of quality of life: the Interdisciplinary Group for Cancer Care Evaluation (GIVIO) experience in Italy. *Stat Med* 1998;17:577-585.

68. Movsas B, Hunt D, Watkins-Bruner D, et al. Can electronic web-based technology improve quality of life data collection? Analysis of Radiation Therapy Oncology Group 0828. *Pract Radiat Oncol* 2014;4:187-191.

69. Moyer A, Knapp-Oliver SK, Sohl SJ, et al. Lessons to be learned from 25 years of research investigating psychosocial interventions for cancer patients. *Cancer J* 2009;15:345-351.

70. Moynihan C. Patient ‘non-compliance’ and ‘missing data’ in quality of life research: where does the problem lie? *Eur J Cancer* 1998;34:9-11.

71. Olschewski M. Compliance with QOL assessment in multi-centre German breast cancer trials. *Stat Med* 1998;17:571-575.

72. Osoba D. The quality of life committee of the Clinical Trials Group of the National Cancer Institute of Canada: organization and functions. *Qual Life* Res 1992;1:211-218.

73. Osoba D, Dancey J, Zee B, et al. Health-related quality-of-life studies of the National Cancer Institute of Canada Clinical Trials Group. *Monogr Natl Cancer Inst* 1996;20:107-111.

74. Osoba D, and Zee B. Completion rates in health-related quality-of-life assessment: approach of the National Cancer Institute of Canada Clinical Trials Group. *Stat Med* 1998;17:603-612.

75. Peat G, Thomas E, Handy J, et al. The Knee Clinical Assessment Study – CAS(K). A prospective study of knee pain and knee osteoarthritis in the general population: baseline recruitment and retention at 18 months. *BMC Musculoskelet Disord* 2006:7:30.

76. Peruselli C, Paci E, Franceschi P, et al. Outcome evaluation in a home palliative care service. *J Pain Symptom Manage* 1997;13:158-165.

77. Phan NQ, Blome C, Fritz F, et al. Assessment of pruritis intensity: prospective study on validity and reliability of the visual analogue scale, numerical rating scale and verbal rating scale in 471 patients with chronic pruritis. *Acta Derm Venereol* 2012;92:502-507.

78. Pinchon DJ, Stanworth SJ, Doree C, et al. Quality of life and use of red cell transfusion in patients with myelodysplastic syndromes. A systematic review. *Am J Hematol* 2009;84:671-7.

79. Powell DA, Furchtgott E, Henderson M, et al. Some determinants of attrition in prospective studies on aging. *Exp Aging Res* 1990;16:17-24.

80. Psaty BM, Cheadle A, Koepsell TD, et al. Race- and ethnicity-specific characteristics of participants lost to follow-up in a telephone cohort. *Am J Epidemiol* 1994;140:161-171.

81. Reimer T, Gerber B. Quality-of-life considerations in the treatment of early-stage breast cancer in the elderly. *Drugs Aging* 2010;27:791-800.

82. Reisine S, Fifield J, Winkelman DK. Characteristics of rheumatoid arthritis patients: who participates in long-term research and who drops out? *Arthritis Care Res* 2000;13:3-10.

83. Ross L, Thomsen BL, Boesen EH, et al. In a randomized controlled trial, missing data led to biased results regarding anxiety. *J Clin Epidemiol* 2004;57:1131-1137.

84. Sadura A, Pater J, Osoba D, et al. Quality-of-life assessment: patient compliance with questionnaire completion. *J Natl Cancer Inst* 1992;84:1023-1026.

85. Schamber EM, Takemoto SK, Chenok KE, et al. Barriers to completion of patient reported outcome measures. *J Arthroplasty* 2013;28:1449-1453.

86. Scott RB, Farmer E, Smiton A, et al. Methodology of neuropsychological research in multicentre randomized clinical trials: a model derived from the international subarachnoid aneurysm trial. *Clin Trials* 2004;1:31-39.

87. Sherman DW, McSherry CB, Parkas V, et al. Recruitment and retention in a longitudinal palliative care study. *Appl Nurs Res* 2005;18:167-177.

88. Shumaker SA, Moinpour CM, Aaronson NK, et al. Part 1: health-related quality of life assessment in clinical research. *Quality of Life Assessment: Practice, Problems, and Promise.* National Institutes of Health, Bethesda MD, 1990, 27-45.

89. Sloan JA, Varricchio C. Quality of life endpoints in prostate chemoprevention trials. *Urology* 2001;57(Suppl 4a):235-240.

90. Slymen DJ, Drew JA, Wright BL, et al. Item non-response to lifestyle assessment in an elderly cohort. *I J Epidemiol* 1994:23:583-591.

91. Solberg TK, Sorlie A, Sjaavik K, et al. Would loss to follow-up bias the outcome evaluation of patients operated for degenerative disorders of the lumbar spine? *Acta Orthopaedica* 2011;82:56-63.

92. Taphoorn MJB, van den Bent MJ, Mauer MEL, et al. Health-related quality of life in patients treated for anaplastic oligodendroglioma with adjuvant chemotherapy: results of a European Organisation for Research and Treatment of Cancer clinical trial. *J Clin Oncol* 2007;25:5723-5730.

93. Taphoorn MJB, Sizoo EM, Bottomley A. Review on quality of life issues in patients with primary brain tumors. *Oncologist* 2010;15:618-26.

94. Vickers AJ. Multiple assessment in quality of life trials: how many questionnaires? How often should they be given? *J Soc Integr Oncol* 2006;4:135-138.

95. Walker M, Brown J, Brown K, et al. Practical problems with the collection and interpretation of serial quality of life assessments in patients with malignant glioma. *J Neurooncol* 2003;63:179-186.

96. Watkins-Bruner D, Scott C, Lawton C, et al. RTOG's first quality of life study RTOG 90-20: a phase II trial of external beam radiation with etanidazole for locally advanced prostate cancer. *Int J Radiation Oncology Biol Phys* 1995;33:901-906.

97. Williams TA, Leslie GD. Challenges and possible solutions for long-term follow-up of patients surviving critical illness. *Aust Crit Care* 2011;24:175-185.

98. Yancik R, Edwards BK, Yates JW. Assessing the quality of life of cancer patients: practical issues in study implementation. *J Psychosoc Oncol* 1989;7:59-74.

99. Zhivan NA, Ang A, Amaro H, et al. Ethnic/race differences in the attrition of older American survey respondents: implications for health-related research. *Health Serv Res* 2012;47:241-254.

100. Zunzunegui MV, Beland F, Gutierrez-Cuadra P. Loss to follow-up in a longitudinal study on aging in Spain. *J Clin Epidemiol* 2001;54:501-510.

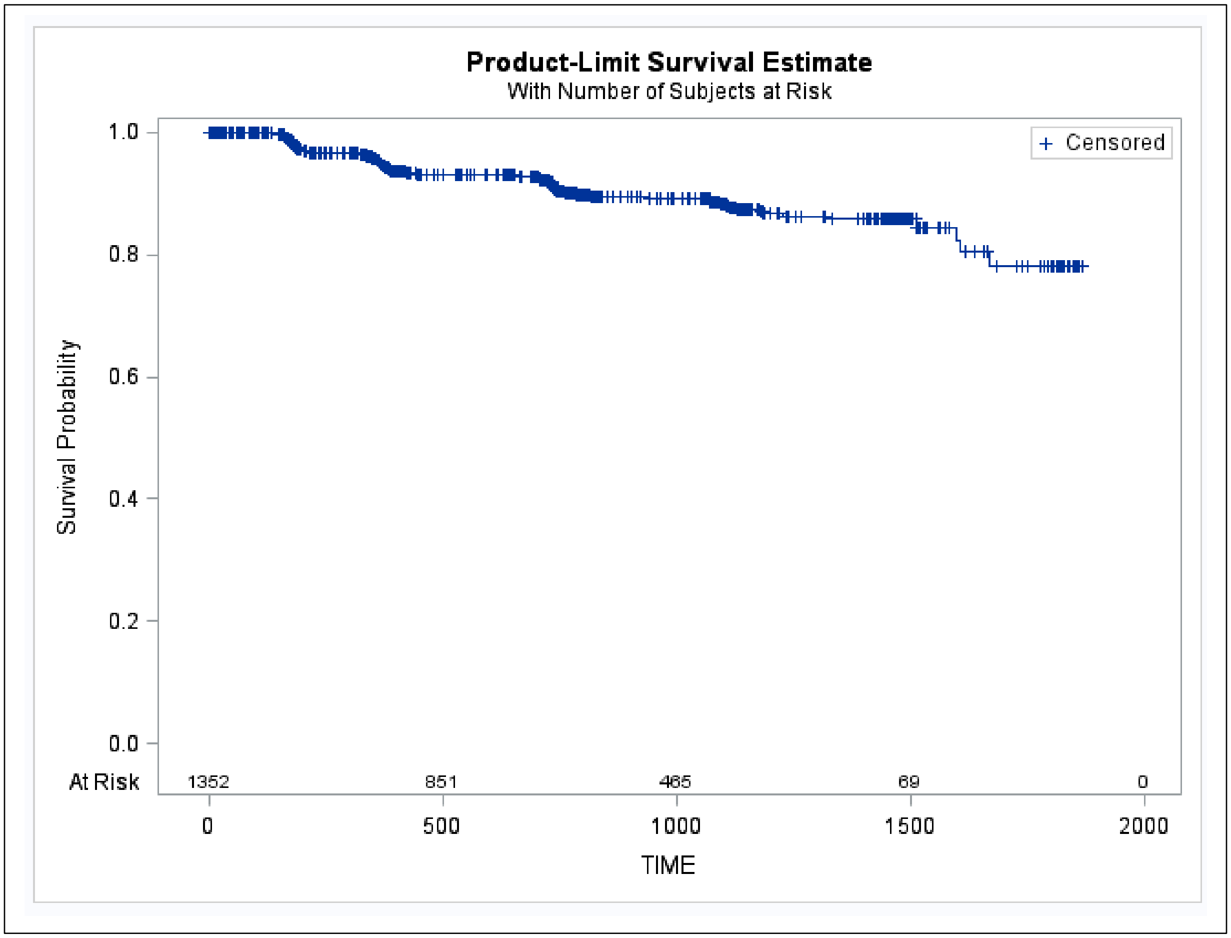
The following table summarizes the source journals for the 100 papers that met the eligibility criteria.

| Number of Papers from Journal | Journal | Number of Journals |
|---|---|---|
| 13 | *Stat Med* | 1 |
| 6 | *J Clin Oncol* | 1 |
| 5 | - | - |
| 4 | *Eur J Cancer* | 1 |
| 3 | *BMJ; J Clin Epidemiol* | 2 |
| 2 | *Compr Psychiatry; Health Qual Life Outcomes; J Neurooncol; J Pain Symptom Manage; Oncology; Qual Life* Res | 6 |
| 1 | *Acta Anaesthesiol Scand; Acta Derm Venereol; Acta Orthopaedica; Acta Psychiatr Scand; Addiction; Am J Epidemiol; Am J Hematol; Am J Public Health; Ann Epidemiol; Ann Oncol; Ann Surg Oncol; Annu Rev Med; Appl Nurs Res; Arch Gerontol Geriatr; Arthritis Care Res; Aust Crit Care; BMC Cancer; BMC Health Serv Res; BMC Med Res Methodol; BMC Musculoskelet Disord; BMC Public Health; Braz Oral Res; Cancer; Cancer J; Clin Trials; Cont Lens Anterior Eye; Control Clin Trials; Curr Prob Cancer; Drugs Aging; Exp Aging Res; Gastroenterol Nurs; Gynecol Oncol; Health Serv Res; Hepato Gastroenterol; I J Epidemiol; Int J Radiation Oncology Biol Phys; J Affect Disord; J Arthroplasty; J Cataract Refract Surg; J Epidemiol Community Health; J Gerontol Psychol Sci; J Natl Cancer Inst; J Psychosoc Oncol; J Psychosom Res; J Soc Integr Oncol; J R Soc Med; Liver Transpl Surgery; Monogr Natl Cancer Inst; N Engl J Med; Oncologist; Pain; Palliat Medicine; Pract Radiat Oncol; Psychol Assess; Psychooncology; Quality of Life Assessment: Practice, Problems, and Promise; Semin Oncol; Support Care Cancer; Urology* | 59 |

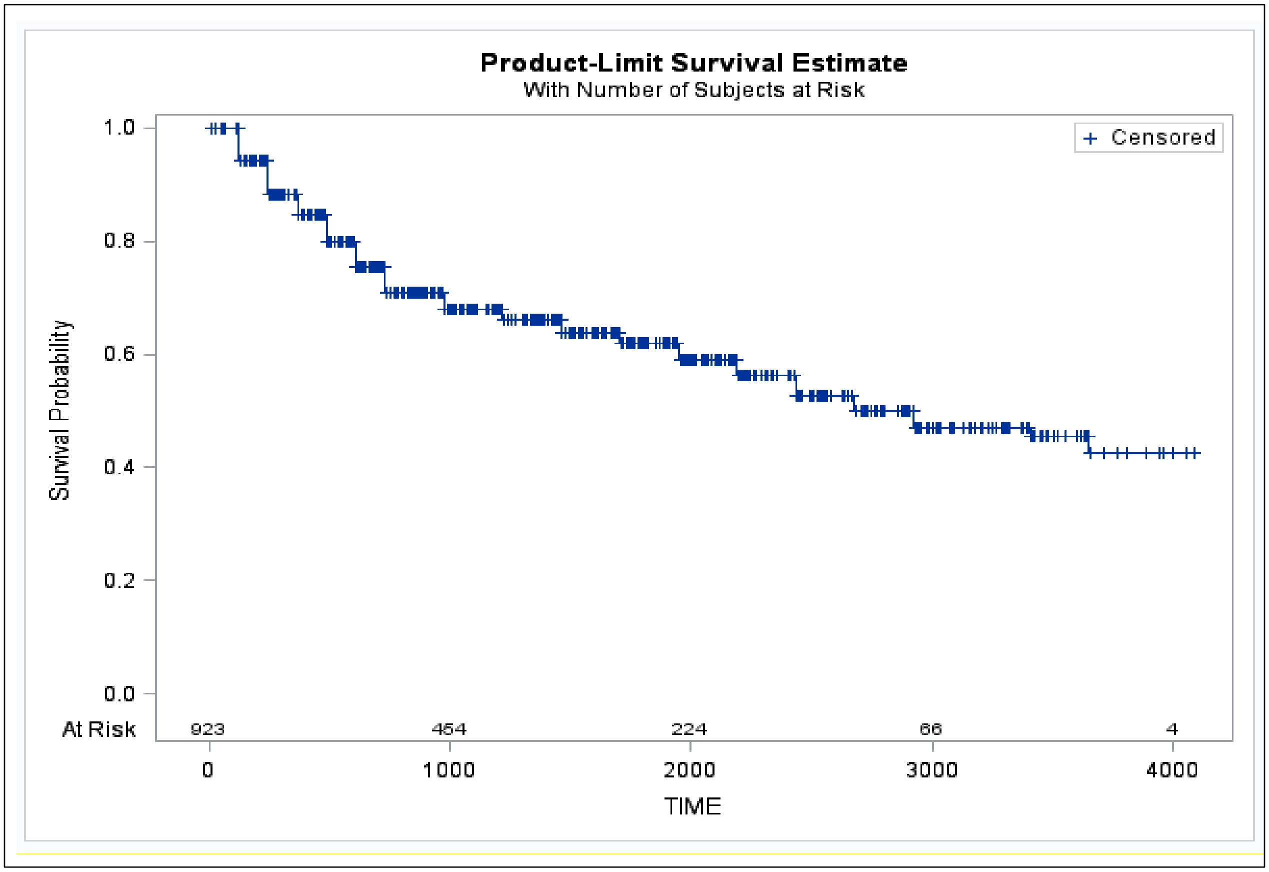
## Appendix E.2

## E-Supplements for the manuscript presented in Chapter 4

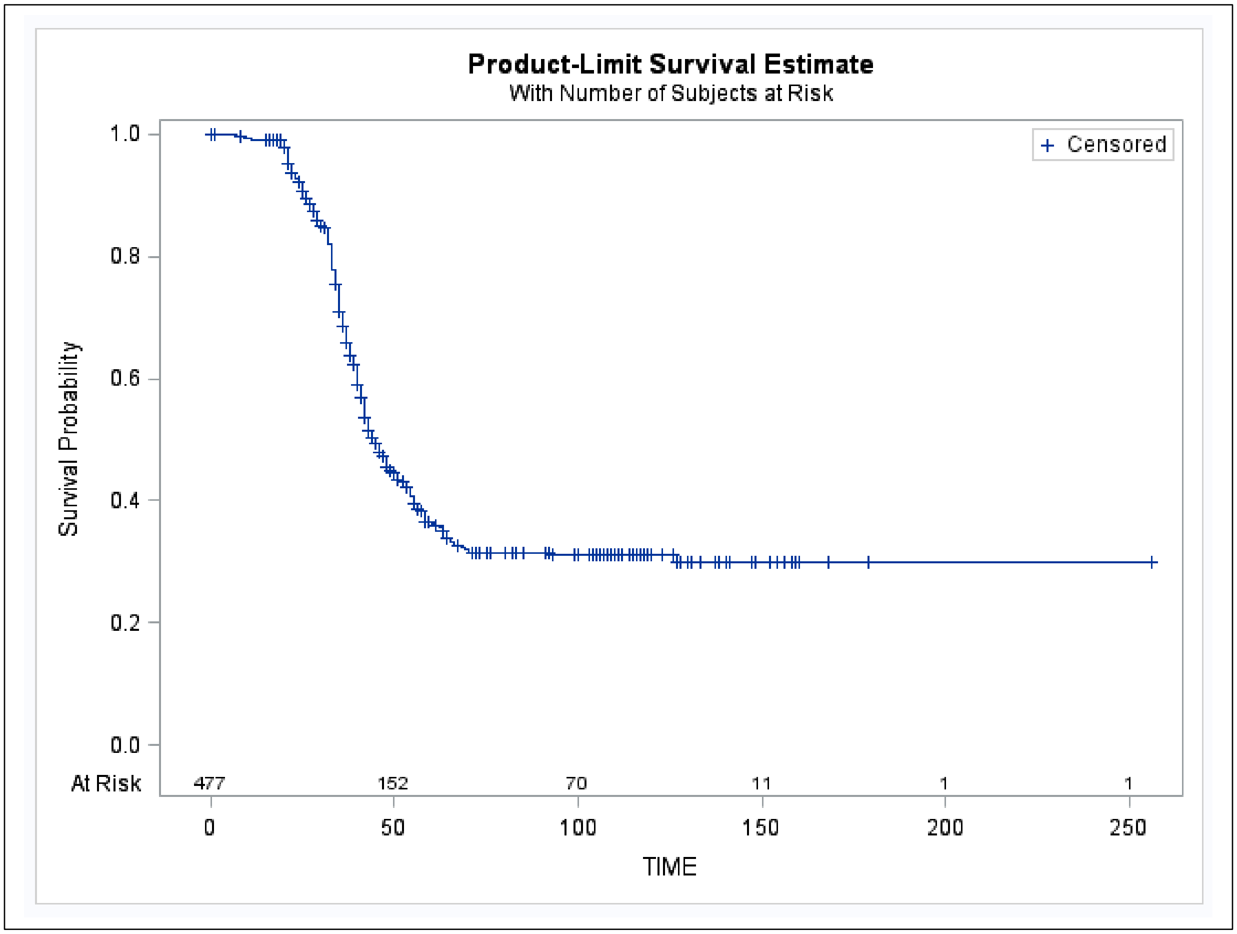
Kaplan-Meier graphs for MA17, PR7, and LY12 are provided on the following pages.




Graph is Kaplan-Meier estimate of survival function for MA17. Step function drops when a subject had a missing instrument. Graph remains flat between failure times. Vertical ticks indicate censored observations. 25% of women had a missing instrument within 4.6 years of randomization. The median time to first missing instrument was not observed for MA17.




Graph is Kaplan-Meier estimate of survival function for PR7. Step function drops when a subject had a missing instrument. Graph remains flat between failure times. Vertical ticks indicate censored observations (castration-resistant disease, off-protocol, death). The curve drops most steeply at the start. About 20% of men had a missing instrument within the first year. 50% of men had a missing instrument within 7.3 years (= 2678 days) of randomization.




Graph is Kaplan-Meier estimate of survival function for LY12. Step function drops when a subject had a missing instrument. Graph remains flat between failure times. Vertical ticks indicate censored observations. The curve drops most steeply at the start. About three-quarters of subjects had a missing instrument within the first 50 days. 50% of subjects had a missing instrument within 0.12 years of randomization.

## Appendix E.3

## E-Supplements for the manuscript presented in Chapter 5

Letter of Invitation, Informed Consent Form, Interview Guide for Clinical Research Associates, Background and Demographic Information Form, Example of Field Note, and Summary of the Triangulation and Face-validity Exercises are provided on the following pages.

Letter of Invitation

Subject: Request your participation in an interview

Hello ___________,

My name is Michael Palmer. I am a PhD Candidate at Queen's University. With the guidance of my co-supervisors, Dr. Michael Brundage and Dr. Harriet Richardson, I would like to understand why information referred to as patient-reported outcomes is sometimes missing in randomized clinical trials.

Anna Sadura provided your contact information to me and is aware of my research.

I am organizing individual interviews with Clinical Research Associates. I would like to invite you to participate in an interview that I anticipate will last one hour. Through discussion of general questions, I hope to learn about your experience with missing PRO data.

Ethics approval has been received from Queen's University HSREB. There is an informed consent form. The interview will be audiotaped and transcribed. It is considered confidential and your privacy will be respected.

Would you be interested in participating?

Thank you for considering this request. I look forward to your reply.

Kind regards,

Michael

Michael J. Palmer
PhD Candidate
Cancer Care and Epidemiology
Department of Public Health Sciences
Queen's University
Kingston, ON, Canada

Email: Michael.Palmer@kingstonhsc.ca

Informed Consent Form

**Faculty**

Michael Brundage, MD
Radiation Oncologist
Director

Kristan J. Aronson, PhD
Epidemiologist

Chris Booth, MD
Medical Oncologist

Deb Feldman-Stewart, PhD
Cognitive Psychologist

Jennifer Flemming, MD
Gastroenterologist

Patti A. Groome, PhD
Epidemiologist

Stephen Hall, MD
Otolaryngologist

Timothy Hanna, MD
Radiation Oncologist

Ana P. Johnson, PhD
Health Economist

William J. Mackillop, MD
Radiation Oncologist

Paul Peng, PhD
Biostatistician

Harriet Richardson, PhD
Epidemiologist

Hugh Walker, PhD
Health Economist

Jenn O'Donnell
Administrative Assistant
*cce@queensu.ca*

**Supported by**
Queen's University
Cancer Care Ontario

CCE
Cancer Care and Epidemiology
Cancer Research Institute
Queen's University

Informed Consent Form

Consent to participate in an interview for research purposes

You are being invited to participate in an interview organized by Michael Palmer, a graduate student at Queen's University. The purpose of this research is to collect data (that will be used by Michael Palmer for his Ph.D. thesis) to understand why information referred to as 'patient reported outcomes' is sometimes missing in randomized clinical trials.

**What is involved?**

You are being asked to participate in one interview. The interview will be audio-recorded and transcribed. In total twenty people will be interviewed.

**Confidentiality**

The interview is considered confidential and your privacy will be respected. All identifying information will be removed from the transcription. Only Michael Palmer, his supervisors (Dr. Harriet Richardson and Dr. Michael Brundage), and a member of his thesis committee (Dr. Terry Krupa), will review the audio-recording and transcript but will not be able to identify you.

**Risks and Benefits**

There should be no risks to your participation in this activity but if stress arises, you will have opportunity to debrief with the audio-recording turned off. Your participation will benefit future patients by helping researchers improve the design of randomized clinical trials. This activity focuses on research and the interview may be included in a publication.

**Voluntary nature of study / Freedom to withdraw or participate**

Your participation in this activity is voluntary. You may withdraw from this activity at any time without any negative consequences.

*Achieving the Achievable in Cancer Control*

10 Stuart Street, Level 2, Kingston, Ontario, Canada K7L 3N6 Tel: 613-533-6895
Fax: 613-533-6794 www.queensu.ca/cce email: cce@queensu.ca

**Informed consent**

I have read and understand this consent form. I have had the purposes, procedures and technical language of this research activity explained to me. I have been given sufficient time to consider the above information and seek advice if I chose to do so. I have had the opportunity to ask questions which have been answered to my satisfaction. I am voluntarily signing this form.

I will receive a copy of this consent form for my information.

If at any time I have further questions, problems or adverse events, I can contact:

Michael Palmer or Dr. Harriet Richardson or Dr. Michael Brundage, Division of Cancer Care and Epidemiology within Queen's University Cancer Research Institute, at telephone number 613-533-6895.

If I have any concerns about this study I can contact:

Dr. Will Pickett, Head of the Department of Public Health Sciences within Queen's University, at telephone number 613-533-6000 extension 79551.

If I have questions regarding my rights as a research participant I can contact:

Dr. Albert Clark, Chair, Queen's University Health Sciences and Affiliated Teaching Hospitals Research Ethics Board, at telephone number 1-844-535-2988.

By signing this consent form, I am indicating that I agree to participate in this interview.

| | |
|---|---|
| ______________________ | ______________________ |
| Signature of Participant | Date |
| ______________________ | ______________________ |
| Signature of Person explaining the consent process | Date |

Achieving the Achievable in Cancer Control

10 Stuart Street, Level 2, Kingston, Ontario, Canada K7L 3N6 Tel: 613-533-6895
Fax: 613-533-6794 www.queensu.ca/cce email: cce@queensu.ca

Interview Guide for Clinical Research Associates

| Heading | Content | Rationale |
|---|---|---|
| Preamble | I would like to ask you some questions. I will listen to what you say. I will make a few notes but please continue to talk. | To focus attention |
| Start | Are you ready to begin? Okay, I will turn on the audio-recorder. | |
| Behaviour | 1. I'm interested in your experience with asking a cancer patient to complete a PRO instrument on a RCT. What is a typical encounter like? Take me through the steps.<br>Prompt:<br>Describe the usual process. | To initiate CRA thinking about the process and to obtain information on how it works. |
| Experience | 2. Tell me about some of your memorable experiences with missing PRO data.<br>Prompt:<br>Take your time. Reflect on your experiences. | To capture what the CRA experienced with missing PRO data, for the textual description, and to capture the contexts in which the CRA experienced the missing PRO data, for the structural description. |
| | 3. Can you share one or two stories about missing PRO data?<br>Prompt:<br>Think of your own practice and describe an experience with missing PRO data that was meaningful to you. | |
| | 4. In your experience what are the important influences leading to missing PRO data?<br>Prompt:<br>What are the things that most influence missing PRO data?<br>What situations especially influence missing PRO data?<br>What influences do you feel are the most important? | |
| | 5. It has been said that the important influences leading to missing PRO data may depend on whether it is before or after randomization. I am interested in your experience. Can you give me one or two examples?<br>Prompt:<br>Think back to when an assessment was expected but the patient did not complete all or part of the instrument. It may be at baseline, or it may be during treatment or follow-up. Describe your experience with the missing PRO data.<br>Have you had experience with each of these? | |

| | | |
|---|---|---|
| | This next section is about a number of things. I will ask you about each of them, one at a time.<br><br>6. Tell me what your experiences are as to whether or not <u>instruments</u> influence missing PRO data.<br><br>Prompt:<br>Quality of life questionnaires are examples of instruments. Think about your experience with the instruments. Was there anything about the instruments that contributed to missing PRO data? Tell me about it.<br><br>7. Tell me what your experiences are as to whether or not <u>cancer patients</u> influence missing PRO data.<br><br>Prompt:<br>Think about your experience with the cancer patients. Was there anything about the cancer patients that contributed to missing PRO data? Tell me about it.<br><br>8. Tell me what your experiences are as to whether or not the <u>centre</u> influences missing PRO data.<br><br>Prompt:<br>The <u>centre</u> is the facility where you work as a CRA. Think about your experience with the centre. Was there anything about the centre that contributed to missing PRO data? Tell me about it.<br><br>9. Tell me what your experiences are as to whether or not the <u>staff</u> influences missing PRO data.<br><br>Prompt:<br>The staff is research personnel working at the centre. Think about your experience with the staff. Was there anything about the staff that contributed to missing PRO data? Tell me about it.<br><br>10. Tell me what your experiences are as to whether or not the <u>RCTs</u> influence missing PRO data.<br>Prompt:<br>The <u>RCTs</u> are phase III studies. Think about your experience with the RCTs. Was there anything about the RCTs that contributed to missing PRO data? Tell me about it. | |

| | | |
|---|---|---|
| Final question | 11. Is there anything else you would like to tell me about your experience with missing PRO data?<br><br>Prompt:<br><br>I have no more questions, but I'd like to give you opportunity to share anything else. | To provide opportunity for the CRA to contribute more information |

Background and Demographic Information Form

Please provide some information about yourself.

1. Have you worked as a CRA at another centre?

__ No.
__ Yes. Which centre(s)? ______________________________________________________________

2. Have you worked as a CRA on RCTs for an organization other than CCTG?

__ No.
__ Yes. Which organization(s)? __ Another cooperative group __ Industry __ Within centre __ Other

3. What is your age?

__ <21 years
__ 21-30 years
__ 31-40 years
__ 41-50 years
__ 51-60 years
__ >60 years

4. Please indicate the highest level of education achieved.

__ Secondary (high) school diploma or certificate
__ Registered Apprenticeship or other trades certificate or diploma
__ College, CEGEP, or other non-university certificate or diploma
__ University bachelor's degree
__ University master's degree
__ University doctorate

5. How do you identify yourself?

__ Certified Clinical Research Associate [CCRA® (ACRP)]
__ Certified Clinical Research Coordinator [CCRC® (ACRP)]
__ Certified Clinical Research Professional [CCRP® (SOCRA)]
__ Clinical Research Associate
__ Clinical Research Nurse
__ Registered Nurse
__ Other. Please describe. ______________________________________________________________

6. How many years of experience do you have with collection of PROs on RCTs?

__ <1 year
__ 1–5 years
__ 6–10 years
__ 11-15 years
__ 16–20 years
__ 21-25 years
__ >25 years

7. What training have you received for collection of PROs on RCTs? Please briefly describe.

______________________________________________________________________________

______________________________________________________________________________

______________________________________________________________________________

Acronyms

| | |
|---|---|
| ACRP | Association of Clinical Research Professionals |
| CCTG | Canadian Cancer Trials Group |
| CEGEP | Collège d'enseignement général et professionnel |
| CRA | Clinical Research Associate |
| PROs | Patient Reported Outcomes |
| RCTs | Randomized Clinical Trials |
| SOCRA | The Society of Clinical Research Associates |

Example of Field Note

Today, at 2:00 pm, is my second interview. I have never met B, and look forward to interviewing [redacted].

In preparation, I reviewed the interview guide, survey of background and demographic information, and consent form. The audio-recorder is functional and has a fresh battery. The small meeting room is booked and I will post a sign (Interview in Progress) on the door.

I had the following e-mail exchange with B, when scheduling a convenient time and date for the interview:

B: I am sorry I don't really understand what you are looking for, is it from my perspective, personal experience?

MJP: I am interested in your own experience. I will ask you to describe your experience with cancer patients. I will be interested in learning about your experience with a few patients who participated on a randomized clinical trial but didn't answer all of the questions in a questionnaire, or didn't complete the questionnaire at all.

B: Thanks for the info, that gives me more of idea of what you are looking for and I can think of patients in these 2 situations prior to our interview. I do not have a lot experience with the paper PRO CCTG studies but I do with ePRO data using a tablet for various pharmaceutical studies I work on. So they can't miss questions on the ePRO since it won't let them go to the next page. I have had patients decline during the study for a few reasons or have me work on them with them since they aren't great at reading or working with electronic device.

My focus is on B and [redacted] lived experience with missing PRO data. Given the above information, it will involve EDC. In the previous interview, A described [redacted] experience with EDC so it will be interesting to hear the experience of B. Clearly, there is opportunity with B to illicit depth in responses about missing instruments with EDC and, as per RHBS-932 note of 2015FEB19, I will keep a Devil's advocate question in my back pocket:

> 'In the experience of some people, modern technology is so good that with EDC, there are no missing questions because the patient has to answer everything on the screen. However, the experience of other people is that the technology is so bad that with EDC, there are lots of missing instruments. What is your experience? '

I acknowledge my past work within CCTG and at McGill University, and current research observing a focus group of CRAs, conducting a systematic review of the literature and creating a classification framework, and analyzing data from a few RCTs. However, I will not think about what I may know regarding the association of factors with missing PRO data, nor will I think about responses provided by A. Rather, I will actively listen to what B says during the interview. B has expertise. I don't, and can learn from B's description of meaningful events.

I met B [redacted] and walked to [redacted]. B asked if I had previously worked at CCTG, and if one of my RCTs was MA17. Yes. B previously worked in [redacted], then transitioned to cancer (experience with PRO on RCTs). B thought I may be [redacted] at CCTG and wondered if purpose was to find out why PROs aren't administered/completed. No – I have no affiliation with CCTG and purpose of my research is to understand why PROs are sometimes missing on RCTs.

Provided overview, obtained informed consent, asked if B had any questions (no), turned on audiotape and began interview at 2:10.

1. Good overview of ePRO with tablet, visits and baseline and after randomization. Sometimes the tablet contains 3-4 questionnaires.

2. Lung cancer. One patient: elderly, old school, 70 years. Another patient: different trial, B reads questions and records his answers, 60 years. Compared with paper, with tablet, answer all questions.

3. ESAS – completion mandated for all patients, done electronically in clinic before ePRO on tablet or cell phone, but if ESAS not done, paper copy provided in chart outside room. PRO-alert described. Patient burden: same questions asked.

4. Lung cancer: PRO after progression. [redacted] blinded study: stop at safety visit. Assessment at discontinuation. Two males had progressed, were unblinded, had both received treatment not placebo, so study treatment was not an option. One man: angry, upset, would not do PRO. Other man: accepting, would do PRO. Only difference was pathology (squamous versus non-squamous): not reason for not doing PRO.

   Lung cancer patients are accepting of their diagnosis (smoking). Breast cancer patients are harder on B personally as asked different PRO questions: body image, attractiveness, sex life, "those questions", "what kind of pervert would ask those questions?". One woman with breast cancer would answer paper questionnaire but would refuse to answer group of questions starting Q.41. Missing questions with paper but not tablet. I played devil's advocate. In [redacted] experience with tablet, no missing questions because tablet works. Depends on vendor.

5. Not asked.

6. Length. MA32 – metformin study in breast cancer. Really long questionnaire and dietary questionnaire. Too much. 20 pages.

7. Stage. Adjuvant – highly motivated, more likely to complete PRO. Metastatic – more of a problem completing.

8. No.

9. Sometimes forget to administer questionnaire or forget to provide access to tablet for back-up CRA. Personally embarrassing.

10. Protocol. Discontinuation.

11. Patients get asked so many questions.

Turned off audiotape at 3:30 pm.

B said that stress was not experienced, and no question was inappropriate. However, [redacted] was surprised the interview was audiotaped. [redacted] did volunteer, with audiotape turned off, that [redacted] job involves front-line work with real people who may be dying of cancer. [redacted] boss had an in-house session reminding everyone of importance of collecting urine sample, PRO and other information from all patients even when palliative. B asked how they could do this and said, in some situations and with some patients, [redacted] will not do it. [redacted] has been in room, crying with a patient who was coping with bad news, and takes this home with [redacted]. Lung cancer patients often deteriorate quickly and B is aware of how each patient is doing. Sometimes [redacted] can't ask a patient to complete a PRO. The role of a CRA involves humanity and an emotional burden. I commented I have no clinical experience, but wonder about the dual responsibility to care for a person and to collect data.

B completed survey on background and demographic information. [redacted] pointed out [redacted] accreditation is [redacted] which is competitor of [redacted]. I said I would revise Q.5 to include this as option.

End 3:40 pm

General impressions:

A very good interview. Articulate, excellent examples, informative. Good description of EDC. B likes tablet and no body language with arms crossed. B laughed and smiled on several occasions. I recall one occasion where [redacted] changed from description of patient who had been informed of progression, to a similar patient who completed questionnaire. I was better at following the discussion, but could improve my formulation of questions. I especially fumbled trying to articulate Q.6.

After about one hour, the first page of the interview guide had been covered. B asked, if I was going to interview more CRAs, [redacted] could connect me with some [redacted] colleagues. I explained I was going to interview 20 CRAs in total, all experienced with cancer patients and PROs on RCTs, from different centres across Canada. I asked B if [redacted] was comfortable with a few more questions – yes, so I turned the page and asked Q.6 – Q.10. Moving forward with interviews, I should consider not asking these questions if the influence of (instruments/patients/centre/staff/RCTs) have been described in depth previously.

Additional information:

B sent me an e-mail (entitled 'Interview') after the interview with the following text and an enclosed figure.

B: It was nice meeting with you last week. As discussed I am sending you a copy of the ESAS questions that patients answer/rate when they come into the cancer center. Below is the electronic print out they receive from the machine after they complete the questions at the kiosk. It shows how they previously have answered at visits. If they miss this, the nurse asks them to complete a paper copy and someone inputs that information so it appears on the print out at their next visit so each visit can be compared. If a patient rates high the nurse is to document her conversation about that particular symptom as to why it is high so help can be offered. Let me know if you have follow-up questions, I would happy to answer.

MJP: I appreciate receiving a copy of the ESAS questions and additional information about it - this is very helpful!

I will contact you if I have additional questions.

It was a pleasure to meet you and I learned a lot from the interview!!

Best regards and many thanks,

Summary of the Triangulation and Face-validity Exercises

MP and TK participated in an investigator triangulation exercise by independently identifying relevant statements in the first five interviews then meeting to talk about it, including thoughts on organizing initial clusters in a meaningful way. MP brought the next level of analysis to co-authors with the supporting data and discussed emerging themes. HR and MB expressed confidence in the ability of MP to identify clusters from the data.

TK, HR and MB participated in a face-validity exercise. For each theme, participants were asked: Is the theme an accurate reflection of the clusters of meaning statements? Is the wording of the theme clear? For each cluster, participants were asked: Is the cluster relevant to missing PROs? Is the wording of the cluster clear?

Participants independently reviewed every theme and the associated clusters of meaning statements, indicated their choices, and provided insightful comments. MP collated the responses. A ‘No’ response and/or inquisitive comment prompted MP to flag nine themes and 26 clusters for review. MP carefully considered the provided information, checked the supporting data as required, and made revisions. One theme was revised to more accurately reflect the clusters of meaning statements and to improve clarity, eight themes were revised to improve clarity, and one theme was kept as is. Twenty-six of the 64 clusters were flagged for review. Two clusters were deleted (because of redundancy and data do not support a link to missing PRO data), two clusters were moved to another theme, and 20 clusters were revised (to improve relevance to the theme, relevance to missing PRO data, and clarity). Two clusters were not changed because the statements accurately reflect the data. Minor suggestions to improve wording were accepted for 13 clusters, and 24 clusters were kept as is.

Participants were provided with a detailed summary of the previous round, and the revised wording in track changes for every theme and associated clusters, clean wording, and a few notes clarifying some of the changes. Participants were asked to review the revised wording and to provide comments if further revision was required.

Participants independently reviewed the content and provided insightful comments. MP collated the responses, carefully considered the provided information, checked the supporting data as required, and made revisions. One theme and eight clusters were flagged for review. In the 'Resource' theme, two words were deleted because of redundancy. In the 'Manifested' theme, four clusters were collapsed into one cluster to simplify wording, and another cluster regarding emotional reaction was kept as is because the supporting data were insufficiently rich to justify creation of a new theme. In the 'Format' theme, one cluster was revised to improve clarity. In the 'Instrument' theme, one cluster was revised to improve clarity, and a new cluster was added.

Consensus was reached on the final themes and on the final clusters of meaning statements.

## Appendix F

**Design, implementation and reporting strategies to reduce the instance and impact of missing patient-reported outcome (PRO) data: a systematic review.**

Rebecca Mercieca-Bebber, Michael J Palmer, Michael Brundage, Melanie Calvert, Martin R Stockler, Madeleine T King. 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:e010938. doi:10.1135/bmjopen-2015-010938

The abstract is provided on the following pages.

**Abstract**

**Objectives** Patient-reported outcomes (PROs) provide important information about the impact of treatment from the patients' perspective. However, missing PRO data may compromise the interpretability and value of the findings. We aimed to report: (1) a non-technical summary of problems caused by missing PRO data; and (2) a systematic review by collating strategies to: (A) minimise rates of missing PRO data, and (B) facilitate transparent interpretation and reporting of missing PRO data in clinical research. Our systematic review does not address statistical handling of missing PRO data.

**Data sources** MEDLINE and Cumulative Index to Nursing and Allied Health Literature (CINAHL) databases (inception to 31 March 2015), and citing articles and reference lists from relevant sources.

**Eligibility criteria** English articles providing recommendations for reducing missing PRO data rates, or strategies to facilitate transparent interpretation and reporting of missing PRO data were included.

**Methods** 2 reviewers independently screened articles against eligibility criteria. Discrepancies were resolved with the research team. Recommendations were extracted and coded according to framework synthesis.

**Results** 117 sources (55% discussion papers, 26% original research) met the eligibility criteria. Design and methodological strategies for reducing rates of missing PRO data included: incorporating PRO-specific information into the protocol; carefully designing PRO assessment

schedules and defining termination rules; minimising patient burden; appointing a PRO coordinator; PRO-specific training for staff; ensuring PRO studies are adequately resourced; and continuous quality assurance. Strategies for transparent interpretation and reporting of missing PRO data include utilising auxiliary data to inform analysis; transparently reporting baseline PRO scores, rates and reasons for missing data; and methods for handling missing PRO data.

**Conclusions** The instance of missing PRO data and its potential to bias clinical research can be minimised by implementing thoughtful design, rigorous methodology and transparent reporting strategies. All members of the research team have a responsibility in implementing such strategies.

## Appendix G

## Guidelines for Inclusion of Patient-Reported Outcomes in Clinical Trial Protocols. The SPIRIT-PRO Extension.

Melanie Calvert, Derek Kyte, Rebecca Mercieca-Bebber, Anita Slade, An-Wen Chan, Madeleine T. King, and the SPIRIT-PRO Group: Amanda Hunn, Andrew Bottomley, Antoine Regnault, Carolyn Ellis, Daniel O'Connor, Dennis Revicki, Donald Patrick, Doug Altman, Ethan Basch, Galina Valikova, Gary Price, Heather Draper, Jane Blazeby, Jane Scott, Joanna Coast, Josephine Norquist, Julia Brown, Kirstie Haywood, Laura Lee Johnson, Lisa Campbell, Lori Frank, Maria von Hildebrand, Michael Brundage, **Michael Palmer**, Paul Kluetz, Richard Stephens, Robert M. Golub, Sandra Mitchell, Trish Groves.

Reproduced with permission from *Journal of the American Medical Association*. 2018. 319(5):483-494. doi:10.1001/jama.2017.21903 

The abstract is provided on the following pages.

**Abstract**

**Importance** Patient-reported outcome (PRO) data from clinical trials can provide valuable evidence to inform shared decision making, labeling claims, clinical guidelines, and health policy; however, the PRO content of clinical trial protocols is often suboptimal. The SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) statement was published in 2013 and aims to improve the completeness of trial protocols by providing evidence-based recommendations for the minimum set of items to be addressed, but it does not provide PRO-specific guidance.

**Objective** To develop international, consensus-based, PRO-specific protocol guidance (the SPIRIT-PRO Extension).

**Design, Setting, and Participants** The SPIRIT-PRO Extension was developed following the Enhancing Quality and Transparency of Health Research (EQUATOR) Network's methodological framework for guideline development. This included (1) a systematic review of existing PRO-specific protocol guidance to generate a list of potential PRO-specific protocol items (published in 2014); (2) refinements to the list and removal of duplicate items by the International Society for Quality of Life Research (ISOQOL) Protocol Checklist Taskforce; (3) an international stakeholder survey of clinical trial research personnel, PRO methodologists, health economists, psychometricians, patient advocates, funders, industry representatives, journal editors, policy makers, ethicists, and researchers responsible for evidence synthesis (distributed by 38 international partner organizations in October 2016); (4) an international Delphi exercise (n = 137 invited; October 2016 to February 2017); and (5) consensus meeting (n = 30 invited;

May 2017). Prior to voting, consensus meeting participants were informed of the results of the Delphi exercise and given data from structured reviews evaluating the PRO protocol content of 3 defined samples of trial protocols.

**Results** The systematic review identified 162 PRO-specific protocol recommendations from 54 sources. The ISOQOL Taskforce (n = 21) reduced this to 56 items, which were considered by 138 international stakeholder survey participants and 99 Delphi panelists. The final wording of the SPIRIT-PRO Extension was agreed on at a consensus meeting (n = 29 participants) and reviewed by external group of experts during a consultation period. Eleven extensions and 5 elaborations to the SPIRIT 2013 checklist were recommended for inclusion in clinical trial protocols in which PROs are a primary or key secondary outcome. Extension items focused on PRO-specific issues relating to the trial rationale, objectives, eligibility criteria, concepts used to evaluate the intervention, time points for assessment, PRO instrument selection and measurement properties, data collection plan, translation to other languages, proxy completion, strategies to minimize missing data, and whether PRO data will be monitored during the study to inform clinical care.

**Conclusions and Relevance** The SPIRIT-PRO guidelines provide recommendations for items that should be addressed and included in clinical trial protocols in which PROs are a primary or key secondary outcome. Improved design of clinical trials including PROs could help ensure high-quality data that may inform patient-centered care.

## Appendix H

**Missing patient-reported outcome data: understanding associated factors and strategies to reduce its instance and impact through careful design, quality assurance and reporting.**



An oral communication was presented at International Society for Quality of Life Research in Dublin. The abstract is provided on the following pages.

## Abstract


**Aims:** Longitudinal studies with patient-reported outcomes (PROs) almost inevitably suffer from missing data, which can compromise the interpretability and value of findings, particularly when whole assessments are unavailable for analysis. High rates of missed assessments reduce study power and precision, and may lead to biased estimates if not handled appropriately in the analysis—especially when missing data is ‘‘informative’’ of poorer health outcomes. This presentation integrates and synthesises factors associated with missing PRO data and presents novel guidance for minimising the impact of missing PRO data.

**Methods:** Medline and CINAHL were systematically searched (inception-March 2015) for articles on missing PRO data. Two investigators independently screened abstracts against pre-defined criteria and extracted data. References and citing articles were hand-searched. A framework was developed to classify factors associated with missing PRO data into five components (instrument, participant, centre, staff, study). Strategies for minimising the instance and impact of missing data were organised into design and implementation strategies and further mapped to the classification framework of associated factors. Reporting strategies were presented separately.

**Results:** We screened 6083 abstracts. Ultimately 100 articles informed the conceptual framework, and 117 offered strategies to reduce the instance of ‘‘avoidable’’ types of missing PRO data, and reduce the impact of both avoidable and informative types. In total, 184 unique articles were included in the final guidance. Design strategies most commonly comprised study factors (e.g. ensuring baseline PROs were completed as part of the eligibility criteria, aligning

assessments to clinic visits), and instrument factors (e.g. reducing burden). Common implementation strategies included study factors (e.g. recording reasons for missing PRO data), staff (e.g. appointing a study coordinator responsible for PROs, checking completed questionnaires), centre (e.g. sending staff reminders of upcoming assessments, monitoring) and participant factors. Reporting strategies included: reporting rates, types, and analysis methods for handling missing data; discussing the potential for missing data bias; and presenting baseline PRO scores and sample characteristics.

**Conclusion:** This guidance enhances our understanding of factors associated with missing data and is recommended for use by clinical trial teams to improve the standard of PRO research and minimise the negative impact of missing PRO data.