ASSOCIATION BETWEEN PFS AND HRQoL IN CANCER.

EVALUATING THE ASSOCIATION BETWEEN PROGRESSION-FREE SURVIVAL AND HEALTH-RELATED QUALITY OF LIFE IN CANCER.

By BRUNO KOVIC, BSc, MSc

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

DOCTOR OF PHILOSOPHY (2017)

Department of Health Research Methods, Evidence and Impact
McMaster University
Hamilton, Ontario, Canada

TITLE: Evaluating the association between progression-free survival and health-related quality of life in cancer.

AUTHOR: Bruno Kovic, BSc (University of Toronto), MSc (Universitat Pompeu Fabra)

SUPERVISOR: Dr. Feng Xie

NUMBER OF PAGES: xii, 128

ABSTRACT

Background: In oncology, progression-free survival (PFS) is a surrogate outcome measure and trial end point, which is increasingly being used to determine the efficacy and implementation in patient care of new drugs. The goal of patient-centred cancer care is to extend overall survival (OS) or improve health-related quality of life (HRQoL), however, drugs with PFS benefit are more commonly being approved for use in the absence of OS, when patient benefit would have to arise from improved HRQoL. The association between PFS and HRQoL in oncology has been poorly studied, and this association remains unknown. The objective of this thesis was to thoroughly evaluate the PFS-HRQoL association in oncology.

Methods: We published a protocol outlining the design of a highly comprehensive systematic review, and a new analytical approach to optimally explore the PFS-HRQoL association from published oncology randomized controlled trials (RCTs). We recruited an international team of reviewers to conduct the systematic review across three HRQoL domains, and performed a quantitative analysis to find the PFS-HRQoL association in oncology. We examined our database of eligible RCTs for methodological issues through a descriptive exploration of risk of bias, to further inform on the design and conduct of future RCTs.

Results: We failed to find an association between PFS and HRQoL in the absence of OS in oncology. Very few published oncology RCTs measure and report HRQoL

information, and among those that do, design and conduct issues related to blinding, and especially attrition are common.

Conclusion: Oncology RCTs must either be adequately powered for OS, or designed to properly measure HRQoL, so patients can receive treatments offering benefit in one of these two patient-centered outcomes, and not solely based on PFS. There is a lack of high quality RCTs informing on the PFS-HRQoL association, and more of these types of trials are needed in the future for further analysis to confirm our findings.

ACKNOWLEDGEMENTS

This thesis work is the culmination of enormous effort and dedication on the part of many individuals throughout a long journey. I would like to express my many thanks to all of them.

I would like to express my deepest gratitude and appreciation to my supervisor, Dr. Feng Xie. During these 4 years of study he has provided me with support, guidance, and opportunities for which I will forever be indebted.

A most special thanks also goes out to Dr. Gordon Guyatt. His incredible and never-ending support, guidance, teachings, and mentorship will never be forgotten. His dedication and insight throughout this process helped me achieve my goals. His friendship and inspiration helped me succeed in the toughest of times. Thank you for giving me more support than I could ever imagine. I hope we can continue to collaborate and be friends for many years to come.

I want to sincerely thank Dr. Lehana Thabane for his support and guidance on projects throughout my PhD. Through his teachings he helped me on my journey of personal growth. Thanks for believing in me.

I also want to thank Dr. Michael Brundage for his endless support, guidance, and encouragement. It is difficult to imagine how this thesis work would have been possible if not for his immeasurable insights.

I want to thank my colleagues who worked with me along the way for their hard work, support, and dedication to contributing to my work.

Thank you to my family for your love and encouragement. Last but definitely not least, I want to thank my beautiful wife Adriana and my sunshine Alessandro. Thank you for the infinite physical, emotional, and mental support throughout the entire time of my studies. It definitely would not have been possible without your unconditional love and encouragement. It is hard to imagine a life without your love through the good times and the bad, and the constant inspiration that motivates me to be my best.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AHRQ: Agency for Healthcare Research and Quality

AUC: Area under the curve

CI: Confidence interval

ECOG: Eastern Cooperative Oncology Group

EORTC-QLQ: European Organization for Research and Treatment of Cancer quality of life questionnaire

FACT: Functional Assessment of Cancer Therapy

HR: Hazard ratio

HRQoL: Health-related quality of life

LASA: Linear analog self-assessment

LTFU: loss to follow-up

OS: Overall survival

PFS: Progression-free survival

PRISMA: Preferred Reporting Items for Systematic reviews and Meta-Analyses

RCT: Randomized controlled trials

SD: Standard deviation

TTP: Time to progression

PREFACE

This thesis is a "sandwich" thesis, which combines three individual studies prepared for publication in peer-reviewed journals. At the time of writing this thesis, the first study has been published, the second has been submitted to a journal for publication, and the third has not been submitted yet. Bruno Kovic is the first author for all studies included in this dissertation, and his contributions to all the papers are as follows: study conception, developing the research question, designing the statistical analysis plan, designing the search strategy, designing screening and data abstraction forms, pilot testing, calibration and management of the co-author team, data collection and management, conducting the statistical analyses, designing figures and tables, writing all manuscripts, submitting the manuscripts and responding to reviewers' comments. The co-authors contributed to acquiring the datasets, provision of clinical expertise, and preparing the manuscripts for publication. The work in this thesis was conducted between September 2013 and August 2017.

CHAPTER 1:

INTRODUCTION

Progression-free survival (PFS) in oncology is defined as the time from random assignment in a clinical trial to objective tumour progression or death.[1 2] PFS was originally developed as a measurement tool describing change in tumour burden to identify signals of activity in early drug development.[3] However, in recent years it has been increasingly used as a primary endpoint in randomized controlled trials (RCTs) to evaluate the efficacy, and denote the clinical benefit of new drugs. This increase is evident in both RCT publications as well as drug regulatory approvals, with the proportion of oncology RCT publications using PFS going from 0% during 1975–1984 to 26% during 2005–2009, and 29% of RCTs used to approve drugs between 2000-2010 having PFS as a primary end point.[3] This increased use of PFS is due to advantages such as smaller sample size and shorter follow-up when compared with studies using overall survival (OS).[1 4]

The goal of patient-centred cancer care is to extend survival (i.e. increase OS) or improve health-related quality of life (HRQoL).[3 5 6] OS is widely considered to be the gold standard for demonstrating clinical benefit in cancer RCTs.[6 7] However, regardless of its clinical importance there has been a decrease in its use as the primary endpoint in oncology RCTs in recent years,[8] due to its associated disadvantages including increased trial length, patient requirements, and expense.[4 9] HRQoL, another important outcome

measure in oncology, reflects patients' subjective feeling about their own health,[5 6] although study data related to this outcome is rarely provided.[10]

As well as being used as a primary endpoint in cancer RCTs to approve treatments, PFS is considered a surrogate measure in oncology. A surrogate end point, which is a measure validated as an adequate substitute for an outcome of intrinsic value to patients (i.e. how a patient feels, functions or survives),[11 12] should reliably and precisely predict treatment effect on the outcome being replaced - these outcomes being OS or HRQoL in cancer.[13 14] Unfortunately, in the case of PFS there is uncertainty of its validity as a surrogate measure in oncology, and of its importance to patients, which raises major concerns regarding its use in oncology RCTs.[3-5]

The relationship between PFS and OS is both variable and unpredictable, with PFS only being a valid surrogate for OS in limited scenarios.[3 13 15] Furthermore, many drugs have been approved based on PFS benefit in the absence of OS benefit, with at least a dozen regulatory approvals between 2005 and 2010.[13] The lack of OS benefit would not be a problem given the presence of HRQoL benefit, however, this is by no means certain since there is a lack of data examining the PFS-HRQoL association.[10] Only one study, a qualitative review by the Agency for Healthcare Research and Quality (AHRQ), has thus far attempted to gather the evidence on the PFS-HRQoL association, but its results were inconclusive.[13]

The question which is therefore raised is what benefits these approved drugs can bring to patients if PFS does not represent a good surrogate for HRQoL? Considering the possible absence of OS benefit, it is critically important to evaluate the PFS-HRQoL association in order to know the expected benefits in HRQoL when evidence of PFS benefits are found, and whether these treatments provide important benefit to patients.[16]

This thesis is composed of a series of studies, all the first of their kind sharing the same overarching objective of setting out to evaluate the association between PFS and HRQoL in oncology. We designed chapters 2 to 4 as independent stand-alone manuscripts that examined different information from the same data set. Therefore, there is some overlap in their introductory sections, eligibility criteria, search strategy, and study selection.

Chapter 2 is the proposal for our evaluation, and thus lays the groundwork for future analyses by detailing the results of all scoping work that occurred, including all details on the methodology used. Our protocol, which outlines the methods of our systematic review, also allowed us to explain the statistical methodology and overall analytical approach we developed to analyze the PFS-HRQoL association, since no methodology existed for this. Our new approach that combined area under the curve (AUC) with other statistical properties including expectation of variance, allowed us to quantitatively analyze this association, for the first time in the literature.

Chapter 3, progresses on the methodology built in the previous chapter by performing the systematic review to gather the latest evidence from the literature, and quantitatively analyzes the PFS-HRQoL association, in the absence of OS benefit, using the abstracted data from this evidence. In this chapter we are able to report, for the first time in the literature, on the PFS-HRQoL association.

Chapter 4 aims to expand on the PFS-HRQoL association found by examining the methodological issues and limitations that exist in studies reporting HRQoL in oncology. This study specifically focuses on risk of bias, outlines the impact of methodological issues on the PFS-HRQoL association, and informs the field on future RCT design and conduct.

Chapter 5 ends my thesis by summarizing the main findings and limitations of Chapters 2 to 4. The implications and future direction on the issue of PFS-HRQOL association in oncology are also discussed in Chapter 5.

References

1. Guidance for Industry: Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics.: U.S. Department of Health and Human Services Food and Drug Administration, 2007.

2. McKee AE, Farrell AT, Pazdur R, et al. The role of the U.S. Food and Drug Administration review process: clinical trial endpoints in oncology. *Oncologist* 2010;15 Suppl 1:13-8. doi: 10.1634/theoncologist.2010-S1-13

3. Booth CM, Eisenhauer EA. Progression-free survival: meaningful or simply measurable? *J Clin Oncol* 2012;30(10):1030-3. doi: 10.1200/JCO.2011.38.7571

4. Fallowfield LJ, Fleissig A. The value of progression-free survival to patients with advanced-stage cancer. *Nat Rev Clin Oncol* 2011;9(1):41-7. doi: 10.1038/nrclinonc.2011.156

5. Peppercorn JM, Smith TJ, Helft PR, et al. American society of clinical oncology statement: toward individualized care for patients with advanced cancer. *J Clin Oncol* 2011;29(6):755-60. doi: 10.1200/JCO.2010.33.1744

6. Pazdur R. Endpoints for assessing drug activity in clinical trials. *Oncologist* 2008;13 Suppl 2:19-21. doi: 10.1634/theoncologist.13-S2-19

7. Beckman M. More clinical cancer treatments judged by progression-free rather than overall survival. *J Natl Cancer Inst* 2007;99(14):1068-9. doi: 10.1093/jnci/djm073

8. Martell RE, Sermer D, Getz K, et al. Oncology drug development and approval of systemic anticancer therapy by the U.S. Food and Drug Administration. *Oncologist* 2013;18(1):104-11. doi: 10.1634/theoncologist.2012-0235

9. Fiteni F, Westeel V, Pivot X, et al. Endpoints in cancer clinical trials. *J Visc Surg* 2014;151(1):17-22. doi: 10.1016/j.jviscsurg.2013.10.001

10. Hotte SJ, Bjarnason GA, Heng DY, et al. Progression-free survival as a clinical trial endpoint in advanced renal cell carcinoma. *Curr Oncol* 2011;18 Suppl 2:S11-9.

11. Kim C, Prasad V. Strength of Validation for Surrogate End Points Used in the US Food and Drug Administration's Approval of Oncology Drugs. *Mayo Clin Proc* 2016 doi: 10.1016/j.mayocp.2016.02.012

12. Biomarkers Definitions Working G. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther* 2001;69(3):89-95. doi: 10.1067/mcp.2001.113989

13. Gutman SI, Piper M, Grant MD, et al. Progression-Free Survival: What Does It Mean for Psychological Well-Being or Quality of Life? Rockville (MD)2013.

14. NCI Dictionary of Cancer Terms: National Cancer Institute; [Available from: https://www.cancer.gov/publications/dictionaries/cancer-terms?expand=S accessed 28 july 2017.

15. Shi Q, Sargent DJ. Meta-analysis for the evaluation of surrogate endpoints in cancer clinical trials. *Int J Clin Oncol* 2009;14(2):102-11. doi: 10.1007/s10147-009-0885-4

16. Guyatt G, Montori V, Devereaux PJ, et al. Patients at the center: in our practice, and in our use of language. *ACP J Club* 2004;140(1):A11-2.

CHAPTER 2:

Association between Progression-Free Survival and Health-Related Quality of Life in Oncology: A Systematic Review Protocol

Bruno Kovic,[1] Gordon Guyatt,[1,2] Michael Brundage,[3] Lehana Thabane,[1]

Neera Bhatnagar,[4] Feng Xie [1,5]

Author affiliations

[1]Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario, Canada

[2]Faculty of Health Sciences, McMaster University, Hamilton, Ontario, Canada

[3]Department of Oncology, Queen's University, Kingston, Ontario, Canada

[4]Health Sciences Library, McMaster University, Hamilton, Ontario, Canada

[5]The Research Institute of St. Joseph's Healthcare, Hamilton, Ontario, Canada

Contributors BK and FX are the guarantors. BK drafted the manuscript. BK, GG, MB and FX developed the selection criteria. BK, GG and MB developed the data extraction criteria and risk of bias assessment strategy. BK and NB developed the search strategy. LT, FX and MB provided statistical expertise. LT and BK developed the statistical methodology. MB provided oncology-specific expertise. All authors read, provided feedback and approved the final manuscript.

Competing interests None declared.

Provenance and peer review Not commissioned; externally peer reviewed.

Correspondence to Dr Feng Xie; fengxie@mcmaster.ca

Kovic B, Guyatt G, Brundage M, et al. Association between progression-free survival and health-related quality of life in oncology: a systematic review protocol. BMJ Open 2016;6:e012909. doi:10.1136/bmjopen-2016-012909

Published by the BMJ Publishing Group Limited.

2.1 Abstract

Introduction

There is an increasing number of new oncology drugs being studied, approved and put into clinical practice based on improvement in progression-free survival, when no overall survival benefits exist. In oncology, the association between progression-free survival and health-related quality of life is currently unknown, despite its importance for patients with cancer, and the unverified assumption that longer progression-free survival indicates improved health-related quality of life. Thus far, only 1 study has investigated this association, providing insufficient evidence and inconclusive results. The objective of this study protocol is to provide increased transparency in supporting a systematic summary of the evidence bearing on this association in oncology.

Methods and analysis

Using the OVID platform in MEDLINE, Embase and Cochrane databases, we will conduct a systematic review of randomised controlled human trials addressing oncology issues published starting in 2000. A team of reviewers will, in pairs, independently screen and abstract data using standardised, pilot-tested forms. We will employ numerical integration to calculate mean incremental area under the curve between treatment groups in studies for health-related quality of life, along with total related error estimates, and a 95% CI around incremental area. To describe the progression-free survival to health-related quality of life association, we will construct a scatterplot for incremental health-related quality of life versus incremental progression-free survival. To estimate the

association, we will use a weighted simple regression approach, comparing mean incremental health-related quality of life with either median incremental progression-free survival time or the progression-free survival HR, in the absence of overall survival benefit.

Discussion

Identifying direction and magnitude of association between progression-free survival and health-related quality of life is critically important in interpreting results of oncology trials. Systematic evidence produced from our study will contribute to improvement of patient care and practice of evidence-based medicine in oncology.

2.2 Introduction

Cancer is a devastating disease, causing high morbidity and mortality. Over time, a variety of anticancer treatments have emerged to deal with the many different forms of cancer, all having the same guiding principle – they should only be considered for use when they provide important benefit (i.e. improved quantity and/or quality of patient survival).[1 2] Unfortunately, it is not currently always known whether treatment does provide important benefit to patients, which is critical to clinical decision making.[3]

In oncology, overall survival (OS) is an objective end point since it represents survival time, a direct and unambiguous patient benefit. OS has long been regarded by the oncology community at large, as well as drug regulatory bodies such as the U.S. Food and Drug Administration (FDA), as the gold standard for demonstrating clinical benefit.[4] Using OS as an endpoint has its limitations, in particular the need for larger sample sizes and longer follow-up.[2 4] The other well-recognized primary consideration in cancer therapy is the improvement of health-related quality of life (HRQoL),[1 4] which in oncology generally refers to symptomatic improvement. HRQoL is a patient-reported outcome that reflects the subjective feeling of patients about their own health. HRQoL has been recognized as an important outcome measure to patients, since basing decisions on only survival ignores other vital dimensions of great concern to patients. Its importance is highlighted by the American Society of Clinical Oncology

recommendations for patients, which identifies research into the maximization of HRQoL as a research priority in oncology.[1]

Although important clinical benefit should ultimately be defined in oncology by prolonged survival / improved HRQoL, treatments are also accepted by regulatory authorities on the basis of established surrogates.[4] One currently accepted surrogate endpoint being used is progression-free survival (PFS), defined as the time from randomization in a clinical trial to objective tumor progression or death. Sometimes time to progression (TTP) is used interchangeably with PFS.[5]

Cancer disease progression is typically assessed via medical imaging at scheduled intervals, and determined based on one of the four changes: appearance of one or more new lesions, increase in size of target measurable lesions, clear increase in non-target disease, and worsening of symptoms of disease.[5] Progression is an outcome that was originally developed as a measurement tool to describe change in tumor burden during therapy, thus it was intended for use in phase II screening trials of new drugs, and not used to denote clinical benefit. Essentially this use was focused on identifying signals of activity in early drug development, and not rooted in benefit for patients.[2]

In spite of the first use of PFS as an indicator of treatment effect on tumours, PFS use as an important end point in phase III randomized controlled trials (RCTs) has increased over time, with this being reflected in RCT publications and drug regulatory approvals.

For example the proportion of breast, colorectal, and lung cancer RCT publications using PFS/TTP as the primary endpoint in the *Journal of Clinical Oncology* went from 0% during 1975-1984 to 26% during 2005-2009, and the approval rates of drug indications based on PFS/TTP end points, as reported in different studies, was found to be 23% during 2005-2007 versus 29% during 2000-2010.[2]

Increasing the use of PFS as the outcome measure of choice in oncology trials stems from various sources. First, in contrast to the limitations of the gold standard OS, PFS studies can be shorter and have fewer patients, providing results faster and at less expense.[5] Second, new drugs are now being targeted towards cytostatic rather than cytotoxic molecular mechanisms of action, which makes using PFS to measure cytostatic effects on tumours a logical choice.[6]

There are also major disadvantages to using PFS as a primary outcome, reflecting the uncertainty of the importance of the outcome to patients. If prolonged PFS was associated with prolonged OS, it would clearly be important. However, since prolonged PFS is not necessarily associated with prolonged OS, the only reason it would be of importance to patients is if it were associated with improved HRQoL, which is by no means certain. For instance, any HRQoL benefit of PFS may be eliminated or even reversed by HRQoL impairment as a result of adverse events (AEs) of the treatment required to achieve prolonged PFS.[5]

This fundamental problem with PFS is reflected by the various recent publications from oncological experts who are concerned with the validity of using PFS as a primary outcome for the evaluation of new treatments.[1 2 5] As Booth highlights, the growing use of PFS as a primary end point is not based on its surrogacy for either OS/HRQoL, but on the conveniences of shorter and faster trials and evaluations, offering little benefit to patients when this is the basis for drug approvals, since there should be good evidence for PFS as a surrogate for OS/HRQoL.[2]

A surrogate endpoint is defined as a measure validated as an adequate substitute for an outcome of intrinsic value to patients: how a patient feels, functions, or survives. Treatment effect observed on a valid surrogate endpoint should reliably and precisely predict treatment effect on the outcome being replaced.[6] In terms of PFS being a proper surrogate for how a patient survives (i.e. OS), data suggests that PFS is only a valid surrogate for OS in colorectal cancer, and certain types of ovarian cancer, with data for other types of cancer such as breast, prostate, and lung cancer not supporting the surrogate relationship.[2 6] Indeed, although there is some evidence for the surrogacy of PFS for OS, correlation between these is variable and unpredictable. PFS failing surrogacy for OS, however, may not be a problem if it is a valid surrogate for HRQoL (i.e. how a patient feels/functions).

Unfortunately, there is a paucity of data examining the surrogacy of PFS for HRQoL, however, many drugs are approved based on PFS benefit in the absence of OS benefit.

The question which is therefore raised is what benefits these approved drugs can bring to patients if PFS does not represent a good surrogate for HRQoL? There are insufficient studies directly measuring the value and benefits of PFS for patients. Many novel cancer treatment RCTs report outcomes in PFS and few trials collect good HRQoL data, even though HRQoL is important to patient care, since outcomes may not provide benefits to patients experiencing AEs.[5] An example of this scenario occurred with bevacizumab, which was initially approved for breast cancer by the FDA based on 5.9 months of PFS gain, with no accompanying OS or HRQoL gain and many associated AEs. Further trials of this treatment revealed smaller PFS gains and more associated AEs than originally found, and ultimately lead the FDA to revoke their approval for this indication.[5]

Advocates for PFS claim that delaying progression delivers patient-important benefit since being progression-free results in stability of disease and symptoms that leads to a reduction in physical and psychological morbidity, and thus improves HRQoL. However, as Fallowfield noted, the HRQoL benefits of increased PFS must be balanced against the toxicity of drug therapy.[5] The limitations in measurement of HRQoL outcomes often leave the matter of net benefits of increased PFS versus increased treatment toxicity uncertain.[5]

Until now there has only been one study, a report published by the Agency for Healthcare Research and Quality (AHRQ),[6] that has attempted to gather the evidence and perform a systematic analysis of the association between PFS and HRQoL. The AHRQ report

examined the association of PFS with HRQoL and related outcomes, such as disease symptoms, for agents where PFS was the primary outcome used to assess treatment benefit. Unfortunately, only four studies proved eligible for the AHRQ analysis, and the quality of the evidence was deemed insufficient to make any conclusion about the PFS-HRQoL association.

In evaluating the PFS-HRQoL association, the AHRQ report had some important limitations. The search strategy used in the AHRQ report was not sufficiently comprehensive. First, their search was limited to only the specific set of oncology drugs that were approved based on PFS, ignoring all other studies containing PFS and HRQoL data. Second, they limited their search to the 2004-2012 timeframe, a timeframe consistent with drugs receiving approval for a primary outcome of PFS, but again ignoring a few years of data where PFS was reported in RCTs. Finally, and most importantly, they only chose studies that included a direct quantitative statistical comparison of PFS to HRQoL, limiting their available data given that it is very rare that a direct comparison between PFS and HRQoL is made in published studies.

If a broader search with fewer limits to capture more of the available PFS and HRQoL data was performed, more studies would be found. Greater available studies would allow for a quantitative analysis, including sensitivity and subgroup analyses, which would lead to more robust conclusions on the PFS-HRQoL association.

Given the increased use of PFS as the primary outcome of importance in oncology trials of new drugs, the importance of HRQoL to patients in oncology treatments and trials, and the lack of a consistent surrogate relationship between PFS and OS, it is important to all stakeholders including patients, clinicians, payers, and all decision makers to identify the currently unknown PFS-HRQoL association. Without this evidence it remains possible that patients are receiving toxic and expensive treatments on the basis of PFS prolongation, but are not experiencing any benefit of value to them in the form of prolonged survival or HRQoL.

In order to address this important gap in the current oncology literature, we will perform a systematic review and quantitative analyses to determine the PFS-HRQoL association in studies of oral or intravenous or intraperitoneal or intrapleural chemotherapy and biological therapy designed to improve disease related outcomes among patients with cancer in a RCT setting.

2.3 Methods and Analysis

Study Overview

We will conduct a systematic review of RCTs using the standard methodology as described by the Cochrane Collaboration.[7] Our protocol adheres to the Preferred Reporting Items for Systematic reviews and Meta-Analyses for Protocols 2015 (PRISMA-P 2015)[8 9] guidelines. Additionally, in order to increase the availability and accessibility of the a-priori methods of our systematic review, prior to the initiation of data abstraction we plan on registering our study protocol on the PROSPERO (International Prospective Register of Ongoing Systematic Reviews)[10 11] international register.

Literature Search

With the help of an experienced research librarian, using the OVID platform, we will design and run searches in Medline, Embase, and Wiley's Cochrane Library (i.e. Cochrane Central Register of Controlled Trials - CENTRAL) databases using medical subject headings (MeSH) and text words. The Medline search will allow for the capture of RCTs published in the core clinical journals, or *Abridged Index Medicus*, as defined by the National Library of Medicine.[12] References of retrieved publications will be scanned to identify any other relevant publications. In order to increase the comprehensiveness of our review, our search will have no language limitations.

Search strategies will be built that combine terms from the three key areas of the disease, HRQoL, and cancer treatment, with "AND" modifiers. For the terms in the disease key area, the five major categories of cancer will be used, as identified by major cancer organizations such as the National Cancer Institute (NCI) and others.[13 14] The treatment key area terms will be constructed around chemotherapy and biological type cancer therapies acting as primary agents in cancer treatment. Chemotherapy and biological therapies identified in the search will be defined as stated by the NCI.[15 16] To be even more comprehensive, in our treatment key area terms, we will include the list of specific drugs that have been approved by the FDA based only on PFS outcome for cancer as identified in the AHRQ report,[6] since we know these drugs are of interest to our research question. Additionally, to be able to form conclusions using higher quality studies that are the standard of approval for new interventions and therefore our context of interest, the focus will be only on RCTs. Therefore, a fourth key area with an AND modifier will be used in the form of an RCT filter in order to focus the search results further. The Medline search has been constructed first, and will be used as the baseline search strategy on which to construct the Embase and Cochrane searches. The Medline strategy can be found in online supplementary Appendix 1.

The outcome of PFS was not included in the search strategy, and instead will be screened on in order to adequately capture HRQoL post-hoc analysis publications corresponding to RCTs reporting PFS outcomes. HRQoL post-hoc analysis publications might not mention PFS, and HRQoL data are frequently reported in these later types of publications, instead

of at the time of publication on primary outcomes of efficacy and safety. Various key articles known to be eligible for data abstraction were identified a-priori, with some of these being the final four from the AHRQ report, and these were used to perform a quality check on the Medline search, to check for inclusion of these key articles in the search strategy.

Eligibility Criteria

The following are our eligibility criteria:

Inclusion criteria:

1. The study is a RCT
2. Only human participants
3. The study is examining only cancer disease (i.e. malignant neoplasm), with a focus on cancer treatment
4. The study was published during 2000-2016 (deemed by expert opinion to include most RCTs using PFS as a primary outcome)
5. The study reported PFS outcome estimates for either all trial arms or the difference between arms, in terms of median PFS times or HR difference
6. Report of HRQoL measures (generic, specific, or utility)
7. Oral or intravenous or intraperitoneal or intrapleural chemotherapy/biologicals used as the primary agent in treating cancer including: immunotherapy (cytokines & antibodies), gene therapy, and targeted therapies (Tyrosine kinase inhibitors, Apoptosis-inducing drugs, treatment vaccines, anti-angiogenesis inhibitors)

Exclusion criteria:

1. The presence of statistically significant OS benefits
2. Median or HR PFS outcome measures not provided for all arms
3. Fewer than 2 discrete HRQoL assessments/measurements provided in study (baseline and at least one follow-up assessment must be provided)
4. Error estimates, in the form of variances or SD not reported for all HRQoL measurements
5. All time durations between successive HRQoL measurements not reported
6. All sample sizes at each HRQoL measurement point not reported
7. All non-oral or non-intravenous or non-intraperitoneal or non-intrapleural chemotherapy/biologicals such as surgical therapy or radiotherapy
8. Any agents used against cancer that are used as supportive, secondary, or preventive agents such as hormone therapy or preventive vaccines

Review Process

A team of screeners and abstractors will be recruited to assist with the review process. Teams of two trained reviewers will perform citation and full text screening and data abstraction, in duplicate and independently, for citations distributed to each team at random. Each team will attempt to resolve all discrepancies by consensus through discussion with their respective team member, but if unsuccessful, adjudication to resolve the discrepancy will be performed by one of two arbitrators (BK, FX). The arbitrator will independently review the study and respective discrepancy, before discussing their final decision with the team members.

We will use electronic forms developed with DistillerSR® (Evidence Partners Inc., Ottawa, Canada) and Microsoft Excel software for study screening and data abstraction, respectively. Level 1 title and abstract screening and level 2 full text screening forms will be constructed using eligibility criteria. All forms will be pilot-tested before distribution and implementation in the review process. Before the review formally starts, we will conduct calibration exercises to ensure consistency across reviewers. These exercises will consist of project and role review meetings, as well as sample data review exercises. For the purposes of evaluating reproducibility, K will be calculated and reported for level 2 screening.

Since PFS is not being included in the search strategy, the presence of PFS outcome measures being reported for all arms in each study will be screened on during level 1 screening, if possible, in order to reduce the large amount of studies expected to not contain this required data.

Since the relationship between PFS and HRQoL is important in the absence of OS benefit, and HRQoL reporting would be expected to be less thorough when OS benefit is present, the focus will be on studies reporting no statistically significant OS benefit.

Data Abstraction

We will extract information on our primary outcomes of PFS and HRQoL, regardless of statistically significant differences reported in studies. Median PFS times and/or PFS HRs, will be extracted for every intervention and control group. Mean HRQoL point values for the overall physical domain base case will be extracted for every group at every time point in the study. Error estimates around all mean HRQoL point values, and if available, around PFS values will also be extracted. For the HRQoL measurements, time durations between each successive measurement in each group, and the sample size at each measurement point will also be recorded.

To perform sensitivity and subgroup analyses the following additional data from each study will also be recorded:

- HRQoL measurements for the overall global and overall emotional HRQoL domains.
- potential bias as per the Cochrane risk of bias instrument,[17] with a focus on critical sources (i.e. presence of loss to follow-up (LTFU) issues defined as >20% attrition or problem self-identified by study, failure to follow patients after progression, and the absence of blinding)
- cancer types and/or stages
- the presence or absence of industry funding

Data Analysis

Primary Analysis:

Our data analysis will be informed by observation of the independent point values of median incremental PFS or PFS HR, and the incremental HRQoL values for one intervention versus another (control) provided by each study. All individual HRQoL and PFS point value and error estimate measurements will be standardized across studies to ensure that comparisons for incremental areas and progression-free times are commensurate with each other.

To calculate incremental PFS, we will use either one of two measurements, depending on available data. The first method consists of taking the difference in median PFS time duration between arms as reported in the study, or the second consists of using the PFS HR as reported in the study.

In order to determine the overall PFS-HRQoL association for the base case, a HRQoL domain must be selected that can summarize HRQoL effects across different cancer types and maintain the same types of comparisons across different instruments (since instruments can vary in the domains they capture). Therefore, for the base case scenario analysis of the association of PFS with HRQoL, we will use the overall physical HRQoL domain from each study.

Data analyses for incremental HRQoL will be performed by using an area under the curve (AUC) approach, following the trapezoidal rule, to calculate the treatment effect on HRQoL which requires at least two measures of HRQoL. Incremental AUC differences between intervention and control arm group curves for consecutive HRQoL measurements per group in each RCT, regardless of statistical significance, will be calculated for the entire timeframe duration of HRQoL measurements, up to the shortest HRQoL arm group duration measurement provided (no imputation assumptions will be made). The whole curve with the lower HRQoL baseline score will be adjusted upwards to correct for the imbalance at baseline HRQoL. Refer online supplementary appendix 2 for a more detailed description of the formulas and calculations that were developed for the data analysis.

To describe the PFS-HRQoL association (i.e. whether or not an improvement in PFS corresponds with an improvement in HRQoL), we will construct a scatterplot, with the axes consisting of incremental HRQoL versus incremental PFS, and with each point on the plot representing the data extracted from each individual study. Through the visual observation of this plot we may be able to notice/describe an overall trend of how one variable changes versus another. Figure 1 shows two possible scatterplot example scenarios.

To estimate the PFS-HRQoL association (i.e. how much the dependent variable HRQoL varies with a corresponding increase in the independent variable PFS), we will use a

weighted simple regression approach with an assumption of no error for the independent incremental PFS variable. This assumption aligns with the potential unavailability of error estimates for PFS, since to our knowledge it is not usually reported in studies. The regression will be performed by using the dependent incremental HRQoL area point estimate, versus the independent incremental median PFS time estimate or PFS HR for each study. We will run this simple regression using a regression formula (i.e. $y = \alpha + \beta(x)$, where x is measured without error). Finally, each study point in the regression analysis will be weighted by the inverse of the total variance (incremental area variance), to account for the influence of each study estimate on the regression line, originating from the uncertainty due to different sample sizes across studies.

Sensitivity/subgroup analyses and a-priori hypotheses:

Sensitivity and subgroup analyses will be performed on the base case scenario of the association of PFS with HRQoL. First, we will perform subgroup analyses for overall global and emotional HRQoL domains, and hypothesize a-priori that the overall global HRQoL domain will have the same direction of association as the overall physical HRQoL domain, although the magnitude of the associations may vary due to other domains having lesser or greater impact on overall global HRQoL.

The AHRQ report identified the three most critical sources of potential bias, which are especially related to the association between PFS and HRQoL, these being: LTFU issues defined as >20% attrition, failure to follow patients after progression, and failure to blind. Studies will be evaluated for potential bias with the Cochrane risk of bias instrument,[17] with a focus on the three critical sources, and sensitivity analyses will be performed by excluding studies deemed to have high risk of bias on any of these critical variables. Our a-priori hypothesis regarding study quality is that the inclusion of studies not measuring past median progression, or with significant LTFU/attrition problems (>20% attrition), will decrease the positive PFS-HRQoL association, since incremental HRQoL benefits derived from increased PFS between arms are expected to be seen more the longer the follow-up, due to more time to get over adverse treatment effects and for the positive impact of staying free of recurrence to manifest itself.

Finally, to explore some other potential predictors for their potential effect on the PFS-HRQoL association, oncology type subgroup analysis and a sensitivity analysis excluding industry funded RCTs may be performed, depending on available data. Table 1 presents a summary of the main methodological aspects of this study.

2.4 Discussion

Strengths and Limitations

Our study has several strengths. First, we will employ rigorous systematic review methodology including well designed and comprehensive search strategies, screening that is guided by explicit eligibility criteria, and the use of standardized and pilot-tested screening and abstraction forms. Second, we will maximize reproducibility of our review methods by conducting calibration exercises throughout the review process to ensure consistency across reviewers, and we will report on the K statistic for level 2 screening to identify any potential problems with regards to reproducibility. Our comprehensive search will provide an informative estimate on the PFS-HRQoL association. Our sophisticated analytic strategy will optimally explore the association, and we will conduct sensitivity analyses to address the robustness of the results, and a subgroup analysis to determine whether the association varies across specific domains of HRQoL.

We anticipate several limitations to our analysis. First, reported median PFS will not always match for time duration with HRQoL measurements provided, resulting in a time mismatch between these outcomes. The greater the difference in the proportion of patients in intervention and control groups without a recurrence, and the longer time in which patients with and without a recurrence live with their conditions, the greater the net HRQoL difference between the two groups is likely to be. The difference in the proportion of recurrence, as well as the time lived with or without a recurrence, and thus

the net difference in HRQoL in intervention and control groups will increase with the duration of follow-up. Thus studies with briefer or truncated follow-up with respect to HRQoL, but longer follow-up with respect to PFS – and we have prior knowledge that such differences in follow-up will occur – are liable to underestimate the PFS-HRQoL association. Since, to the extent that this problem occurs, we will be underestimating the PFS-HRQoL association, and this association can only increase in positive magnitude with longer HRQoL follow-up. Finding a strong positive HRQoL-PFS association will be definitive. However, if in many studies we find a large difference in duration of follow-up in HRQoL and PFS, a finding of a weak or absent association will not be definitive.

Second, association does not necessarily imply causality, and with a plethora of unknown potential confounding factors linking PFS to HRQoL, there could be unknown underlying mechanisms having an effect on PFS and HRQoL. Additionally, we may not be able to distinguish between two explanations of an association between PFS and HRQoL: one in which the difference in tumor burden in those with and without recurrence directly results in difference in HRQoL, versus knowledge of having progression compared to being progression-free possibly reduces anxiety and increases HRQoL by that mechanism. It is possible that domain by domain analyses (for instance, if emotional function but not physical function improves, it might suggest the latter mechanism of effect) may help elucidate this issue. Also, we plan on including all HRQoL assessments, most of which will be prior to disease progression, in order to minimize the bias inherent in HRQoL collected on the day of documentation of progression.

Finally, although ideally we would stratify our HRQoL data by all known confounding factors, such as by type of progression (i.e. radiographic vs clinical) or censoring due to death, we cannot perform these stratifications because of the nature of published HRQoL data. However, since many of these confounders are expected to be relatively well balanced between trial arms, we do not anticipate this being a major concern.

Study Implications

Thus far, only one study has attempted to evaluate and provide systematic evidence on the PFS-HRQoL association. The AHRQ report conducted a qualitative analysis of a very limited amount of data, resulting in inconclusive results. Our more comprehensive search will result in a greater amount of retrieved data, which will allow for quantitative analyses providing stronger and more robust conclusions regarding the PFS-HRQoL association.

The PFS-HRQoL association is of primary concern to all stakeholders and decision makers including patients, clinicians, and payers. This is especially true given the ever increasing number of regulatory approved oncology drugs based on PFS benefit, even when no OS advantages are evident. This study will provide systematic research on a current evidence gap that exists in the oncology world. In order for PFS based oncology treatments to conform to evidence based medicine, systematic research on the PFS-HRQoL association needs to be produced so it can be integrated with individual clinical expertise in making decisions about the care of individual patients.

One area that this study will most likely challenge will be the conduct of oncology RCTs. Given the nature of our data, it is likely that one of the conclusions of this study will be that RCT authors need to design and run studies in a way as to minimize attrition and LTFU problems, as well as follow patients well past median progression, in order to, for the time being, measure what is most important to patients, HRQoL, and then only substitute PFS when its validity as a surrogate is definitively established.

The findings of this study may produce evidence that challenges conventional thinking regarding the surrogacy of PFS for HRQoL, or it may produce evidence that validates PFS as a surrogate for HRQoL. Regardless, it will contribute evidence that will support improved patient care, either through potentially changing research/clinical practices, guidelines, and policies grounded in PFS, or through providing knowledge that supports the current practices being employed.

2.5 Figures and Tables

Figure 1: Scatterplots for incremental HRQoL (y-axis) versus incremental PFS (x-axis)

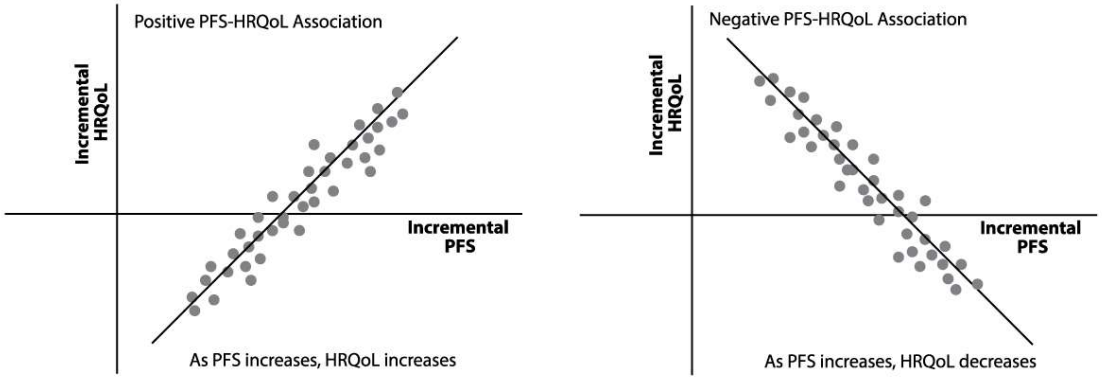


Table 1: Summary of Objective, Sensitivity Analyses, A-Priori Hypotheses, and Analysis Method

Objective/Question	Outcomes	Sensitivity Analyses for Potential Predictors	A-Priori Hypotheses	Analysis Method
•To determine the association between the effect on PFS and effect on HRQoL of oral or intravenous or intraperitoneal or intrapleural chemotherapy and biologicals designed to improve disease related outcomes among patients with cancer in a RCT setting for which these therapies are being compared with another control therapy	•Incremental HRQoL units measured by AUC difference for intervention versus control (continuous) •Incremental PFS measured by difference in median time or HR for intervention versus control (continuous)	•Subgroup analyses for overall global and emotional domains •Sensitivity analysis for where LTFU/attrition, lack of measurement past progression, and lack of blinding is a concern •Oncology type subgroup analyses •Sensitivity analysis excluding industry funded RCTs	•Overall global HRQoL evidence has same directional association as overall physical HRQoL domain analysis •The inclusion of studies not measuring past median progression and with LTFU problems decrease the positive association	•Incremental AUC HRQoL units between treatment groups, adjusted at baseline for HRQoL, calculated using trapezoidal rule •Error estimates of variance and SE will be calculated for incremental AUC point estimate •To describe the PFS-HRQoL association a scatter plot will be constructed •To estimate the PFS-HRQoL association simple weighted regression will be performed by using incremental HRQoL area versus incremental PFS estimate

PFS = progression-free survival, HRQoL = health-related quality of life, RCT = randomized controlled trial, AUC = area under the curve, HR = hazard ratio, LTFU = loss to follow-up, SE = standard error

2.6 References

1. Peppercorn JM, Smith TJ, Helft PR, et al. American society of clinical oncology statement: toward individualized care for patients with advanced cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2011;29(6):755-60. doi: 10.1200/JCO.2010.33.1744

2. Booth CM, Eisenhauer EA. Progression-free survival: meaningful or simply measurable? *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2012;30(10):1030-3. doi: 10.1200/JCO.2011.38.7571

3. Guyatt G, Montori V, Devereaux PJ, et al. Patients at the center: in our practice, and in our use of language. *ACP J Club* 2004;140(1):A11-2.

4. Pazdur R. Endpoints for assessing drug activity in clinical trials. *The oncologist* 2008;13 Suppl 2:19-21. doi: 10.1634/theoncologist.13-S2-19

5. Fallowfield LJ, Fleissig A. The value of progression-free survival to patients with advanced-stage cancer. *Nature reviews Clinical oncology* 2012;9(1):41-7. doi: 10.1038/nrclinonc.2011.156

6. Gutman SI, Piper M, Grant MD, et al. Progression-Free Survival: What Does It Mean for Psychological Well-Being or Quality of Life? Rockville (MD)2013.

7. Higgins JPT, Green S, eds. *Cochrane Handbook for Systematic Reviews of Interventions*. The Cochrane Collaboration, 2011.

8. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic reviews* 2015;4:1. doi: 10.1186/2046-4053-4-1

9. Shamseer L, Moher D, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *Bmj* 2015;349:g7647. doi: 10.1136/bmj.g7647

10. Booth A, Clarke M, Dooley G, et al. The nuts and bolts of PROSPERO: an international prospective register of systematic reviews. *Systematic reviews* 2012;1:2. doi: 10.1186/2046-4053-1-2

11. Booth A, Clarke M, Ghersi D, et al. An international registry of systematic-review protocols. *Lancet* 2011;377(9760):108-9. doi: 10.1016/S0140-6736(10)60903-8

12. Abridged Index Medicus (AIM or "Core Clinical") Journal Titles. 09 July 2003 ed: U.S. National Library of Medicine.

13. SEER Training Modules, Review: Categories of Cancer: U. S. National Institutes of Health, National Cancer Institute; [Available from: http://training.seer.cancer.gov/disease/categories/review.html accessed 25 November 2015.

14. Types of cancer: Cancer Research UK; [updated 28 October 2014. Available from: http://www.cancerresearchuk.org/about-cancer/what-is-cancer/how-cancer-starts/types-of-cancer#main accessed 25 November 2015.

15. NCI Dictionary of Cancer Terms - Chemotherapy: National Cancer Institute.

16. NCI Dictionary of Cancer Terms - Biological Therapy: National Cancer Institute.

17. Chapter 8: Assessing risk of bias in included studies. [updated March 2011. The Cochrane Collaboration, 2011.:[Available from: www.cochrane-handbook.org.

2.7 Appendix 1

Database: Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) <1946 to Present>

Search Strategy:

---

1 cancer*.mp.

2 exp Neoplasms/

3 neoplasm*.mp.

4 exp Carcinoma/

5 carcinoma*.mp.

6 exp Sarcoma/

7 sarcoma*.mp.

8 exp Lymphoma/

9 lymphoma*.mp.

10 exp Leukemia/

11 leukemia*.mp.

12 myeloma.mp.

13 tumor*.mp.

14 tumour*.mp.

15 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14

16 exp "Quality of Life"/

17 quality of life.mp.

18 QOL.mp.

19 Health Related Quality of Life.mp.

20 HRQOL.mp.

21 HRQL.mp.

22 patient reported outcome*.mp. or exp "Outcome Assessment (Health Care)"/

23 ((pro or pros) and outcome*).ti,ab.

24 Patient Preference.mp. or exp Patient Preference/

25 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24

26 drug*.ti,ab.

27 pharmaceutical preparations/ or controlled substances/ or exp dosage forms/ or exp drug combinations/ or exp drugs, essential/ or exp drugs, generic/ or exp drugs, investigational/ or exp materia medica/ or exp nonprescription drugs/ or exp pharmaceutic aids/ or exp placebos/ or exp prescription drugs/ or exp prodrugs/ or exp solutions/ or exp "vaginal creams, foams, and jellies"/ or exp xenobiotics/

28 drug therapy/ or antineoplastic protocols/ or exp chemoradiotherapy/ or chemotherapy, adjuvant/ or consolidation chemotherapy/ or exp administration, intravenous/ or exp administration, oral/ or exp chemotherapy, cancer, regional perfusion/ or exp infusions, parenteral/ or exp injections/ or exp drug carriers/ or exp drug prescriptions/ or drug therapy, combination/ or antineoplastic combined chemotherapy protocols/ or exp fluid therapy/ or home infusion therapy/ or induction chemotherapy/ or maintenance chemotherapy/ or molecular targeted therapy/

29 chemotherapy.mp.

30 growth substances/ or exp angiogenesis modulating agents/ or exp growth inhibitors/ or exp immunologic factors/ or exp adjuvants, immunologic/ or exp interferon inducers/

31 exp Tumor Necrosis Factor-alpha/ or biologic*.ti,ab. or exp Antibodies, Monoclonal/

32 Immunotherapy/

33 Immunotherapy.mp. or Immunotherapy/

34 Cytokines.mp. or exp Cytokines/

35 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34

36 (Bevacizumab or avastin).mp.

37 (ixabepilone or ixempra).mp.

38 (lapatinib or Tykerb).mp.

39 (panitumumab or vectibix).mp.

40 (doxorubicin or doxil).mp.

41 (gemcitabine or gemzar).mp.

42 (trabectedin or yondelis).mp.

43 (sorafenib or nexavar).mp.

44 (pazopanib or votrient).mp.

45 (sunitinib or sutent).mp.

46 (erlotinib or tarceva).mp.

47 (docetaxel or doxetaxel or taxotere).mp.

48 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47

49 35 or 48

50 15 and 25 and 49

51 randomized controlled trial$.mp.

52 randomized controlled trial.pt.

53 double-blind method/

54 single-blind method/

55 controlled clinical trial.pt.

56 ((double$ or trebl$ or tripl$) adj25 (blind$ or mask$)).mp.

57 random$.mp.

58 51 or 52 or 53 or 54 or 55 or 56 or 57

59 50 and 58

60 animals/ not humans/

61 59 not 60

62 limit 61 to yr="2000 -Current"

2.8 Appendix 2

AUC for each group / arm, and the incremental AUC between groups for each study will be calculated using the trapezoidal rule. This rule from the family of formulas for numerical integration, or quadrature, is used for approximating the definite integral, and works by approximating the region under the graph of the function f(x) as a trapezoid and calculating its area. The trapezoid rule instructs to divide the curve into a series of trapezoids, each with area equaling average height (i.e. determined by the two adjacent HRQoL measures) multiplied by width, and to sum the areas of the strips.[1] Because area is a function of differences between means at different time points between curves, or intervention groups, incremental AUC (ΔAUC) is denoted by:

$$1)\ \Delta AUC = Area_I - Area_C = (A1_I + A2_I + \cdots) - (A1_C + A2_C + \cdots),$$

where the brackets containing the "I" subscript represent the total area of the treatment intervention group, and the brackets containing the "C" subscript represent the total area of the control group.

For the purposes of our analysis, this formula is expanded as follows:

$$2)\ \Delta AUC = \left[\frac{(Q1_I + Q2_I)}{2} \times (TD2_I - TD1_I) + \frac{(Q2_I + Q3_I)}{2} \times (TD3_I - TD2_I) + \cdots\right] - \left[\frac{(Q1_C + Q2_C)}{2} \times (TD2_C - TD1_C) + \frac{(Q2_C + Q3_C)}{2} \times (TD3_C - TD2_C) + \cdots\right],$$

where Q represents the mean HRQoL measurements provided for each treatment group, and half the sum between the Q's is the average height of the trapezoid, and where TD represents Time Duration as reported for each HRQoL measurement, and the difference between the TDs is the width of the trapezoid. Formula (2) can be used to find the incremental AUC point estimate difference between intervention and control groups, where we are given consecutive mean HRQoL measurements across time.

In order to perform our data analyses, we not only need the total incremental area point estimates, but also the total error estimates around these point values. We therefore need to calculate two separate variances to get the totals; one for within each treatment group, and the other for between the groups. We will start by calculating within group variances by taking each half of incremental area formula (2), first for the intervention group and then for control group, and once we simplify and rearrange for each group, our formula is as follows:

$$3)\ V(Area_{I\ or\ C}) = d_1^2 \times v(Q_1 + Q_2) + d_2^2 \times v(Q_2 + Q_3) + \cdots ,$$

where d represents the constant of half the time duration difference, and where v($Q_1 + Q_2$) represents the pooled variance within a group.

Because variance within a group has repeated measures, we will account for covariance between measurements by applying the pooled standard deviation for dependent samples equation to formula (3) as follows:

$$4)\ V(Area_{I\ or\ C}) = d_1^2 \times (\sigma_1^2 + \sigma_2^2 + 2 \times \sigma_1 \times \sigma_2 \times p) + d_2^2 \times (\sigma_2^2 + \sigma_3^2 + 2 \times \sigma_2 \times \sigma_3 \times p) + \cdots\ ,$$

where σ represents the standard deviation for each HRQoL measurement as extracted from the study, and where p will equal 0.5, an assumption of a positive midpoint correlation of +0.5. The rationale for this assumption originates from an expectation of increasing HRQoL over time for each particular group with subsequent HRQoL readings, due to treatment being provided in the RCT leading to increased HRQoL over time compared to baseline. Formula (4) will be used to find the within group area variances for the intervention and control group.

In order to find total variance, which is the combination of within group plus between group variances, or variance of the incremental area, we will combine all within group variances calculated. For this we return to formula (1), but as opposed to mean incremental area being a function of differences between curves or groups, total variance is instead a function of the addition of all within group variances as follows:

$$5)\ \mathrm{V}(\Delta\mathrm{Area}) = \mathrm{V}(Area_I) + V(Area_C)$$

Once total variance for incremental HRQoL area is calculated, we will use this in combination with the total sample size of all measurements to calculate the standard error (SE), since SE reflects the variability between means,[2] and we are calculating a difference between means for the incremental area. The SE will be used to derive the 95% confidence interval around the HRQoL incremental area for each study, and get a sense of the inherent uncertainty present in the dependent HRQoL variable.

2.8.1 References

1. Mysovskikh IP. Trapezium formula: Encyclopedia of Mathematics; [updated 7 February 2011. Available from: http://www.encyclopediaofmath.org/index.php?title=Trapezium_formula&oldid=12696 accessed 25 November 2015.

2. Streiner DL. Maintaining standards: differences between the standard deviation and standard error, and when to use each. *Canadian journal of psychiatry Revue canadienne de psychiatrie* 1996;41(8):498-502.

CHAPTER 3:

The association between progression-free survival and health-related quality of life in oncology: a systematic review and quantitative analysis.

Bruno Kovic,[1] Xuejing Jin,[1] Sean A. Kennedy,[2] Mathieu Hylands,[3] Michał Pędziwiatr,[4,5] Akira Kuriyama,[6] Huda Gomaa,[7] Yung Lee,[8] Morihiro Katsura,[9] Masafumi Tada,[10] Brian Y. Hong,[11] Sung Min Cho,[12] Patrick Jiho Hong,[11] Ashley Yu,[11] Yasmin Sivji,[13] Augustin Toma,[14] Li Xie,[15,16] Ludwig Tsoi,[17] Marcin Waligora,[18] Manya Prasad,[19] Neera Bhatnagar,[20] Lehana Thabane,[1] Michael Brundage,[21] Gordon Guyatt,[1,8] Feng Xie [1,22]

Author affiliations

[1]Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Ontario, Canada

[2]Diagnostic Radiology Residency Program, University of Toronto, Toronto, Ontario, Canada

[3]Department of Surgery, Université de Sherbrooke, Sherbrooke, Québec, Canada

[4]2nd Department of General Surgery, Jagiellonian University, Krakow, Poland

[5]Centre for Research, Training and Innovation in Surgery (CERTAIN Surgery). Krakow, Poland

[6]Department of General Medicine, Kurashiki Central Hospital, Kurashiki, Japan

[7]Department of pharmacy, drug information center, Tanta Chest Hospital, Gharbia, Egypt

[8]Faculty of Health Sciences, McMaster University, Hamilton, Ontario, Canada

[9]Department of Surgery, Okinawa Prefectural Nanbu Medical Center & Children's Medical Center, Okinawa, Japan

[10]Department of Health Promotion and Human Behavior, Kyoto University Graduate School of Medicine / School of Public Health, Kyoto, Japan

[11]Faculty of Medicine, University of Ottawa, Ottawa, Ontario, Canada

[12]Faculty of Medicine, University of Toronto, Toronto, Ontario, Canada

[13]Department of Medicine, McMaster University, Hamilton, Ontario, Canada

[14]Michael G. DeGroote School of Medicine, McMaster University, Hamilton, Ontario, Canada

[15]Department of Cancer Prevention, Fudan University Shanghai Cancer Center, Shanghai, China

[16]Department of Oncology, Shanghai Medical College, Shanghai, China

[17]A&E Department, Queen Mary Hospital, Hong Kong

[18]Department of Philosophy and Bioethics, Jagiellonian University Medical College, REMEDY, Research Ethics in Medicine Study Group, Krakow, Poland

[19]Department of Community Medicine, Postgraduate Institute of Medical Sciences, Rohtak, Haryana, India

[20]Health Sciences Library, McMaster University, Hamilton, Ontario, Canada

[21]Department of Oncology, Queen's University, Kingston, Ontario, Canada

[22]The Research Institute of St. Joseph's Healthcare, Hamilton, Ontario, Canada

Correspondence to Dr Feng Xie; fengxie@mcmaster.ca

Research Support None

Prior Research Presentation

Meeting: 2017 ASCO Annual Meeting

Type: Poster Session

Time: Monday June 5, 2017

Abstract No: 6574

Citation: J Clin Oncol 35, 2017 (suppl; abstr 6574)

Web Link: http://abstracts.asco.org/199/AbstView_199_192701.html

3.1 Abstract

Background

The goal of cancer care is to improve survival duration or health-related quality of life (HRQoL). Progression-free survival (PFS) has become an important surrogate outcome in assessing efficacy of new drugs, especially in absence of overall survival (OS) benefit. However, the relationship between improved PFS and HRQoL is unclear. The objective of this study was to examine the relationship between PFS and HRQoL through a systematic review and analysis of published evidence.

Design

We searched MEDLINE, Embase, and Cochrane databases for randomized controlled trials (RCTs) published since 2000 addressing chemotherapy or biological treatments reporting PFS or HRQoL with no OS benefits. We utilized difference in median PFS time duration between treatment groups, and compared HRQoL between groups using difference in standardized mean incremental area under the curve adjusted to per month values. We used weighted simple regressions to examine the PFS-HRQoL association for physical, global, and emotional HRQoL domains.

Results

Of 35,960 records identified, 42 articles reporting on 30 RCTs involving 10,731 patients across 12 cancer types using 5 different HRQoL instruments proved eligible. Of these 30 RCTs, 67% demonstrated improved PFS, and 56%, 54%, and 62% demonstrated

improved physical, global, and emotional HRQoL, respectively. The associations proved weak and easily explained by chance: PFS and physical HRQoL domain (n=18, regression coefficient β=-0.21, 95% CI; -0.65 to 0.24), global domain (n=24, β=0.09, 95% CI; -0.27 to 0.46), or emotional domain (n=13, β=0.78, 95% CI; -0.05 to 1.60).

Conclusions

In clinical trials reporting PFS outcomes with no benefit in OS, longer PFS is not significantly associated with better HRQoL. These findings raise questions regarding the assumption that interventions prolonging PFS also improve quality of life in patients with cancer.

3.2 Introduction

The goal of patient-centred cancer care is to extend survival or improve health-related quality of life (HRQoL).[1-3] Overall Survival (OS), an objective end point representing survival duration, is widely regarded as the gold standard for demonstrating clinical benefit.[3 4] HRQoL reflects patients' subjective feeling about their own health, and has also been recognized as an important outcome measure,[1 3] as well as a research priority by organizations including the American Society of Clinical Oncology.[1] Although important benefit should ultimately be established by improved OS/HRQoL, regulatory authorities have also approved cancer treatments on the basis of surrogates such as progression-free survival (PFS),[3] or time to progression (TTP).[5]

Originally developed as a measurement tool to identify signals of activity in early drug development,[2] PFS has become a widely used surrogate outcome to assess efficacy of new cancer drugs. The appeal of PFS use in randomized controlled trials (RCTs),[2] reflects limitations associated with using OS,[2 3 5] including higher cost, longer follow-up, and larger sample sizes. The increasing use of PFS is also reflected in drug regulatory approvals,[2] with at least a dozen drugs being approved by the US Food and Drug Administration (FDA) since 2005 using PFS as the primary endpoint.[6]

Although the use of PFS is increasing, the uncertainty of its importance to patients, and its validity as a surrogate measure, have raised major concerns regarding the use of PFS as a primary outcome. A valid surrogate end point should reliably and precisely predict

treatment effect on the outcome being replaced, either survival or how a patient feels or functions.[6] This definition is currently implemented by regulatory agencies (i.e. FDA) that grant oncology drugs either accelerated or traditional marketing approvals.[7]

Thus far, evidence suggests that PFS serves as a valid surrogate for OS only in limited scenarios, being both variable and unpredictable.[2 6] The relationship between improved PFS and HRQoL is also not clear, particularly in the absence of OS benefit. The association of prolonged PFS with improved HRQoL is far from self-evident since HRQoL is likely to be impaired by adverse events (AEs) patients experience as a result of treatments delivering prolonged PFS.[5] In response to the uncertainty regarding the relation between PFS and HRQoL, oncological experts have raised concerns regarding the validity of using PFS as a primary outcome for evaluating new treatments,[1 2 5] noting that the growing use of PFS as a primary end point is not based on its surrogacy for either OS or HRQoL, but on the conveniences of shorter and faster trials and evaluations.[2]

Resolving the question of whether PFS is a satisfactory surrogate for HRQoL is challenging given that few studies measure the value and benefits of PFS for patients, and few trials collect and report HRQoL data.[5 8] To date, only one study[6] has performed a systematic analysis of the PFS-HRQoL association. Unfortunately, that report has important limitations relating to lack of a comprehensive search for relevant evidence, and has thus left the question of the association unresolved.

Given the increased use of PFS as the primary outcome in oncology trials of new drugs, and uncertainty whether PFS is an adequate surrogate for either OS or HRQol, it remains possible that patients are receiving toxic and expensive treatments without experiencing any benefit of value. We have therefore examined the relationship between PFS and HRQoL through a systematic review and quantitative analysis of published studies of oral or intravenous or intraperitoneal or intrapleural chemotherapy or biological therapy designed to improve disease-related outcomes among patients with cancer in RCT settings.

3.3 Methods

Our study protocol, detailing the design and analysis, was previously published,[9] and registered on PROSPERO (International Prospective Register of Ongoing Systematic Reviews)[10 11] with registration number CRD42016047162. We conducted a review of human cancer RCTs published from 2000 to May 04, 2016, utilizing standard methodology as described by the Cochrane Collaboration.[12] Our systematic review adheres to the PRISMA Statement[13 14] guidelines to ensure transparent and complete reporting. We included trials examining chemotherapy or biological type cancer therapies as the primary agent in treating cancer (i.e. primary investigational treatment), and that reported on PFS outcome estimates and HRQoL measures, but excluded the trials with statistically significant OS benefits.

Literature Search

We utilized comprehensive search strategies developed with the help of an experienced research librarian (see protocol online supplementary appendix 1 for MEDLINE strategy).[9] We used the OVID platform to search in MEDLINE, Embase, and Wiley's Cochrane databases, and used both MeSH terms as well as free text words to capture RCTs published in the *Abridged Index Medicus.*[15] Our search strategies were built by combining terms from disease key areas (cancer categories), HRQoL, cancer therapies, and an RCT filter. To increase comprehensiveness, we included a list of drugs approved by the FDA based only on PFS benefit,[6] and set no language limitations. We utilized a previous report[6] to pre-identify key articles eligible for the review, and performed a

search strategy validation check by verifying the inclusion of these key articles in the search results. All references were managed using EndNoteX v.7.0.2 (Thomson Reuters, Philadelphia, USA).

Study Screening and Data Abstraction

Ten pairs of reviewers, working independently, conducted eligibility screening and data abstraction. The international team allowed for screening and abstraction of articles in a variety of languages, thus reducing language limitations. Reviewers resolved disagreement by either discussion or adjudication by an arbitrator (BK).

Screening was performed using pilot-tested electronic forms in DistillerSR (Evidence Partners, Ottawa, Canada), constructed as per eligibility criteria in the protocol.[9] Through team meetings and email communications, reviewers received training and detailed written instructions to perform title and abstract and full-text screening. Meetings included calibration exercises with sample data review exercises. We measured reviewer agreement on full text screening using Kappa (κ) as per guidelines proposed by Landis and Koch,[16] by utilizing the VassarStats Kappa online calculator.[17]

Oncology trials often report PFS results first, and subsequently and separately publish HRQoL results. Therefore, we classified trial publications into three categories of outcome reporting: only PFS outcome data (i.e. typical oncology trial publication), only HRQoL data, or a publication reporting both PFS and HRQoL. All categories of

publications are potentially relevant for our review, we therefore performed trial level publication matching and searching prior to data abstraction for publications reporting only PFS or HRQoL. We first cross referenced categorized publication types against each other (i.e. trial level matching) for articles passing full-text screening, to identify and capture trials reporting required data across multiple publications. Furthermore, we performed an additional supplemental search through OVID for any additional unmatched publications using author, intervention, and cancer type keywords in MEDLINE and Embase databases.

Data abstraction was performed using pilot-tested electronic forms in Microsoft Excel software, by reviewers undergoing similar training and calibration as in the screening stage. Reviewers extracted all data relevant to primary outcomes of PFS and HRQoL including median PFS time and/or PFS hazard ratios (HRs) for every intervention and control group, HRQoL scores and corresponding error estimates for the physical domain, time durations between each successive HRQoL measurement in each group, and sample size at each HRQoL measurement point. In order to perform pre-specified sensitivity analyses,[9] global and emotional HRQoL domain scores were also abstracted if available. Additional data originally intended for further sensitivity and subgroup analyses (i.e. risk of bias, cancer types and stages, industry funding) within each HRQoL domain were abstracted;[9] however, insufficient quantity of trials within each HRQoL domain precluded these analyses.

Data Analyses

We performed our analysis using Microsoft Excel and SPSS software (Version 22). We utilized the reported difference in independent point values of median incremental PFS, and the calculated incremental HRQoL values for one intervention versus a control for each trial. Supplementary appendix 2 of our protocol provides a detailed description of our analyses.[9]

All included trials reported median progression times for intervention and control groups, and we calculated incremental PFS for each study by taking the difference in median PFS time duration between the two arms. Approximately 25% of trials (8 of 30) used TTP, so we assumed TTP to be equal to PFS since the majority of trials report PFS, and the two terms are sometimes used interchangeably.[4 5]

HRQoL for the duration of provided measures in each arm was calculated by using an area under the curve (AUC) approach. The difference in AUC between intervention and control arm groups was calculated by adjusting for HRQoL baseline imbalance and for different durations in measuring/reporting HRQoL between arms. Although per trial instrument selection had no effect on calculation of HRQoL scores due to score standardization, we established a hierarchy of measures, using the highest in the hierarchy for our analyses as follows: FACT-G,[18 19] then the EORTC-QLQ-C30,[20 21] and finally any other instrument available. All HRQoL scores across instruments were standardized on a scale of 0 to 100, with higher scores representing better HRQoL. Similarly, all

corresponding error estimates were standardized accordingly by using properties of expectation and variance of a random variable (refer to Appendix 1). Finally, we adjusted incremental HRQoL benefit for all studies to per month values by utilizing the same properties of expectation and variance approach as in the instrument scale error estimate standardization (Appendix 1). This served to both reduce overall heterogeneity of the y-axis by reducing variability due to differences in HRQoL time durations measured across trials, and allowed for a facilitated visual comparison of relative HRQoL benefits across trials in constructed PFS-HRQoL scatterplots.

The physical HRQoL domain scores from each trial were used for the base case analysis of the PFS-HRQoL association. This domain was selected as the base case in order to maximize the comparability across cancer types by utilizing this commonly measured specific domain of perceived importance. The global and emotional HRQoL domain scores were used in the sensitivity analyses. These domains were analyzed similarly to the base case, and against the a-priori hypothesis of sharing similar directional association as the physical HRQoL domain.

We constructed a scatterplot of incremental HRQoL (y-axis) versus incremental PFS (x-axis), with each study point representing an individual trial. To estimate the PFS-HRQoL association, we used weighted simple (i.e. weighted least squares - WLS) regression; to account for different sample sizes across studies, each study point in the regression was weighted by the inverse of the total variance.

3.4 Results

We initially identified 35,960 citations in MEDLINE, Embase, and Cochrane databases, with a total of 30,296 articles after duplicate removal. We removed 27,944 articles after title and abstract screening, leaving 2352 full-text articles assessed for eligibility, with 2317 being excluded. In addition to the 35 articles finally identified in the primary search, we included 7 more articles identified through a supplemental search. Our review included a final total of 30 trials reported in 42 articles.[22-63] Figure 1 presents the article selection process.[64] We calculated $\kappa = 0.75$, 95% confidence interval (CI) (0.73 to 0.78), indicating substantial agreement[16] between reviewers in assessments of study inclusion during full text screening.

Of the 42 articles representing 30 trials, 18 reported both PFS and HRQoL data in a single article. Trial level matching identified another 24 articles representing 12 eligible RCTs.

The 30,296 potentially eligible articles included 1607 non-English language publications in Chinese, French, German, Japanese, Korean, Polish, and Spanish. Two Chinese articles,[35 60] the largest non-English language group, were finally included in our quantitative synthesis.

Table 1 presents trial characteristics organized alphabetically by cancer type and publication year, with Trial ID format corresponding to number of the trial out of 30, followed by cancer type and trial number for the specific cancer type.

The 30 eligible trials, published between 2000 and 2014, involved 10,731 patients across 12 cancer types. Seven cancer types were studied in multiple trials. Trials enrolled from 42 to 1248 adult patients aged 18 to 91. Intervention treatments varied widely, both across and within cancer types. The variability of comparators was lower, with some repetition within cancer type. Reported median follow-up ranged from 10.5 to 66 months across trials. Median PFS of trial interventions ranged from 1.8 to 33.7 months, and the duration of reported or measured HRQoL ranged from 1.5 to 34 months. Out of 30 trials, 19 had shorter HRQoL follow-up than median PFS for the intervention. The studies included in our analysis utilized five different HRQoL instruments: EORTC-QLQ-C30 (17 of 30 trials), FACT (9), Lung Cancer Symptom Scale (2), the eight-item linear analog self-assessment (LASA) questionnaire (1), and clinician-reported Karnofsky score (1). Of the 30 trials, 26 (87%) had high risk of bias in at least one of four domains (i.e. loss to follow-up, failure to follow patients after progression, and absence of blinding in participants, personnel, or outcome assessment),[9] with 11 of 30 trials failing to follow patients after progression.

Table 2 presents a summary of study data results used in our regression analyses. Of 30 trials, 20 (67%) had improved PFS for intervention versus comparator. Physical HRQoL was reported in 18 of 30 trials (60%), while global HRQoL was the most common domain, reported in 24 of 30 trials (80%); emotional HRQoL was reported in 13 of 30 RCTs (43%). Of the RCTs reporting HRQoL across different domains, 10 of 18 (56%),

13 of 24 (54%), and 8 of 13 (62%) trials demonstrated improved physical, global, and emotional HRQoL for the intervention versus the comparator.

Figures 2, 3, and 4 present scatterplots of the relationship between PFS and each of the HRQoL domains. For the association between physical HRQoL and PFS (n=18), the regression coefficient (β=slope) was -0.21 (95% CI -0.65 to 0.24). For the analysis of global HRQoL and PFS (n=24), the regression coefficient (β) was 0.09 (95% CI -0.27 to 0.46). For the analysis of emotional HRQoL and PFS (n=13) the regression coefficient (β) was estimated 0.78 (95% CI -0.05 to 1.60).

3.5 Discussion

Our review and analysis found all three HRQoL domains to have weak and non-significant associations with PFS in the absence of OS benefit, demonstrating no consistent PFS-HRQoL directional association, and supporting our a-priori hypothesis of similar results across different HRQoL domains.

Our study, the first of its kind to quantitatively evaluate the PFS-HRQoL association in oncology, has several strengths. First, we conducted an exhaustive search with no language limitations. Second, we developed explicit eligibility criteria and conducted duplicate assessment of eligibility and data abstraction that included use of standardized and pilot-tested screening and data abstraction forms, review team meetings and communications to ensure resolution of reviewer concerns, and thus achieved excellent agreement. Third, we developed a quantitative analysis methodology that allowed to include the widest possible range of relevant publications. Finally, our trial dataset had widely distributed patient and trial characteristics, ensuring optimally generalizable results.

Our study also has limitations. First, since over 60% of trials have shorter HRQoL follow-up than median PFS for the interventions, and 11 of 30 trials fail to follow patients after progression, we may have failed to capture some HRQoL benefit attributable to PFS that could have occurred later in trials, and thus may have underestimated the association between PFS and HRQol. Second, with only 30 eligible RCTs we could not perform

some planned sensitivity and subgroup analyses. Third, the results do not show significant PFS-HRQoL, and this could be in part from a lack of statistical power arising from the limited number of trials included in the analysis.

Thus far, the question of PFS-HRQoL association in oncology has only been examined in one previous study by the Agency for Healthcare Research and Quality (AHRQ).[6] Unfortunately, only four studies proved eligible for the AHRQ qualitative review, with the quality of the evidence being deemed insufficient to make any conclusion about the PFS–HRQoL association. The most important limitations in the AHRQ report were related to the limited comprehensiveness of their search, which included choosing only studies with a direct quantitative statistical comparison of PFS to HRQoL, resulting in a limited eligible study pool.

The increasing use of PFS in oncology RCTs and drug regulatory approvals over recent years has been based on several advantages of using PFS over OS to evaluate drugs, in particular lower requirements for both sample size and extended follow-up and thus greater speed of trial completion. Aside from these practical advantages, PFS advocates believe that being progression-free indicates disease control and stabilization, leading to reduction in disease symptoms, thus implying clinical benefit through improvement in HRQoL.[6 8] Our results cast serious doubts on such assertions.

These results have important implication for the design and conduct of oncology RCTs. One possible conclusion is that trials with a primary PFS endpoint also need to be designed to provide high-quality findings regarding whether the interventions impacts on HRQoL, including adequate power and data quality (duration of follow-up and high patient compliance). This approach will avoid assumptions[65] and would make the relationship between PFS outcomes and HRQoL outcomes clear, allowing clinical judgement to assess benefits versus risks should the HRQoL outcomes be compromised by treatments that increase PFS. In contrast to the solution for the apparent inadequacy of PFS as a surrogate by ensuring trials are powered to definitively establish impact on OS, ensuring results demonstrate the impact on HRQoL will not require larger studies. Oncology-specific instruments are responsive to small but important changes in HRQoL,[66 67] and require sample sizes of the same order as those powered to establish PFS. Our finding of approximately the same proportion of trials showing HRQoL benefit as those showing PFS benefit supports this observation.

An increase in the measurement and reporting of HRQoL in RCTs would constitute a substantial change in practice: we found a small number of articles, as observed in Figure 1, reporting only HRQoL (26) in comparison to those reporting only PFS (650). Our findings are consistent with previous research reporting that only 4-10% of oncology trials report the impact of interventions on HRQoL.[68 69]

The necessary measurement of HRQoL in RCTs will come with logistical challenges. Investigators will have to implement strategies to minimize the missing HRQoL data, a frequent problem in current RCTs that do address HRQoL. Such strategies would include requiring baseline measurement of HRQoL prior to randomization into trials,[68 69] careful monitoring to ensure measurement takes place at each patient visit, obtaining contact information for a number of individuals with whom patients are not living but who are likely to be aware of the their whereabouts, ensuring resources are available for tracking patients who prove hard to follow, and use of electronic administration of HRQoL instruments completed by patients themselves. Further, investigators will have to ensure adequate logistical planning and institutional staff training that will further minimize missing data and ensure optimal administration of HRQoL instruments.[6 68 70 71]

3.6 Conclusion

The findings of our systematic review and quantitative analysis seriously challenge claims of increased PFS being clearly associated with improved HRQoL. The implications are profound: trials must either be adequately powered for OS, or designed to ensure rigorous and trustworthy measurement of HRQoL.

3.7 Acknowledgments

We would like to thank Maria Cecilia Arechabala for her assistance with citation screening, Nora (Zhiyuan) Chen for her assistance with literature searching and article retrieval, as well as Daniel Shi, Leon Li, Kailai (Kevin) Zhang, and Valerie (Seungyeon) Kim for their assistance with article retrieval.

3.8 Figures and Tables

Figure 1: PRISMA Flow Diagram

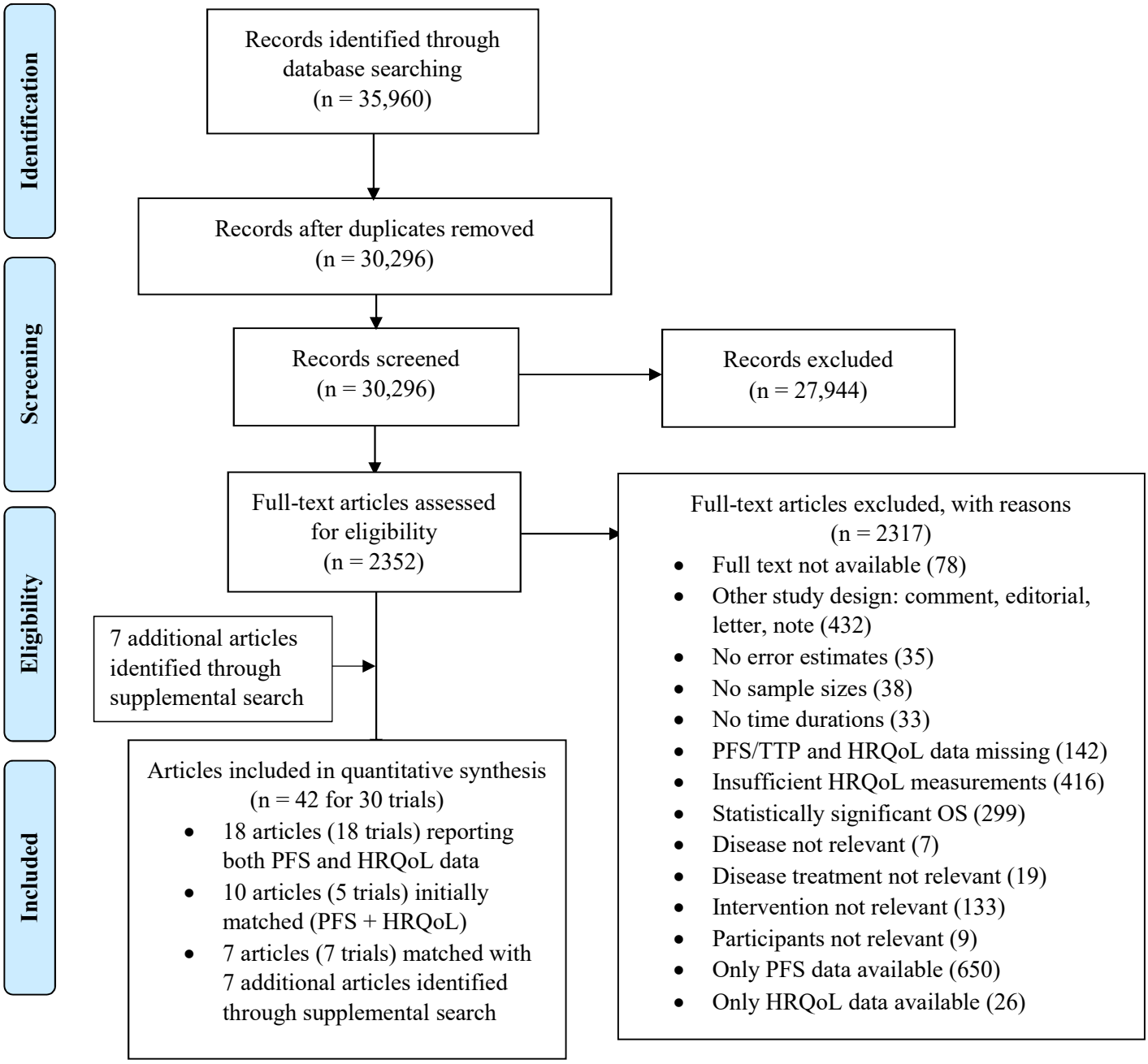
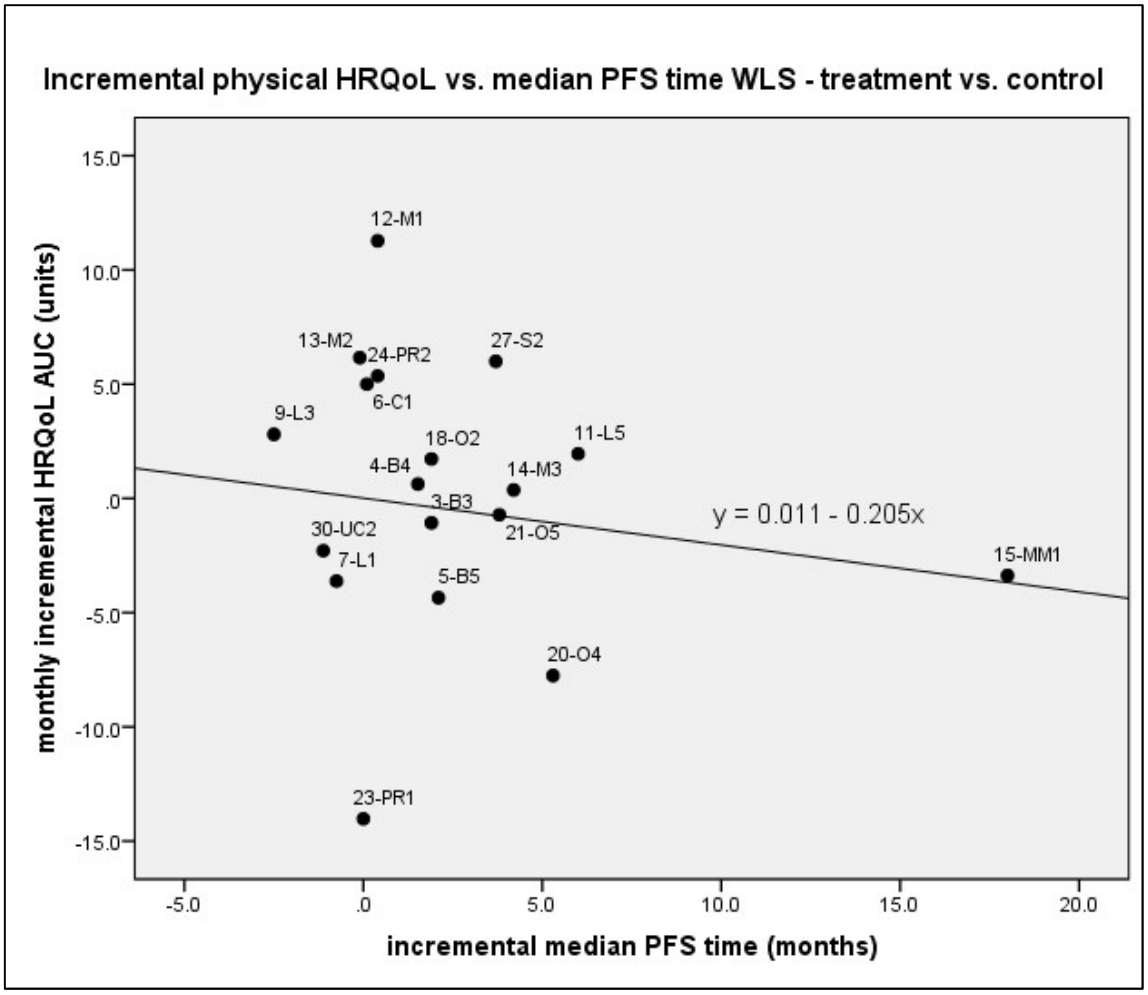
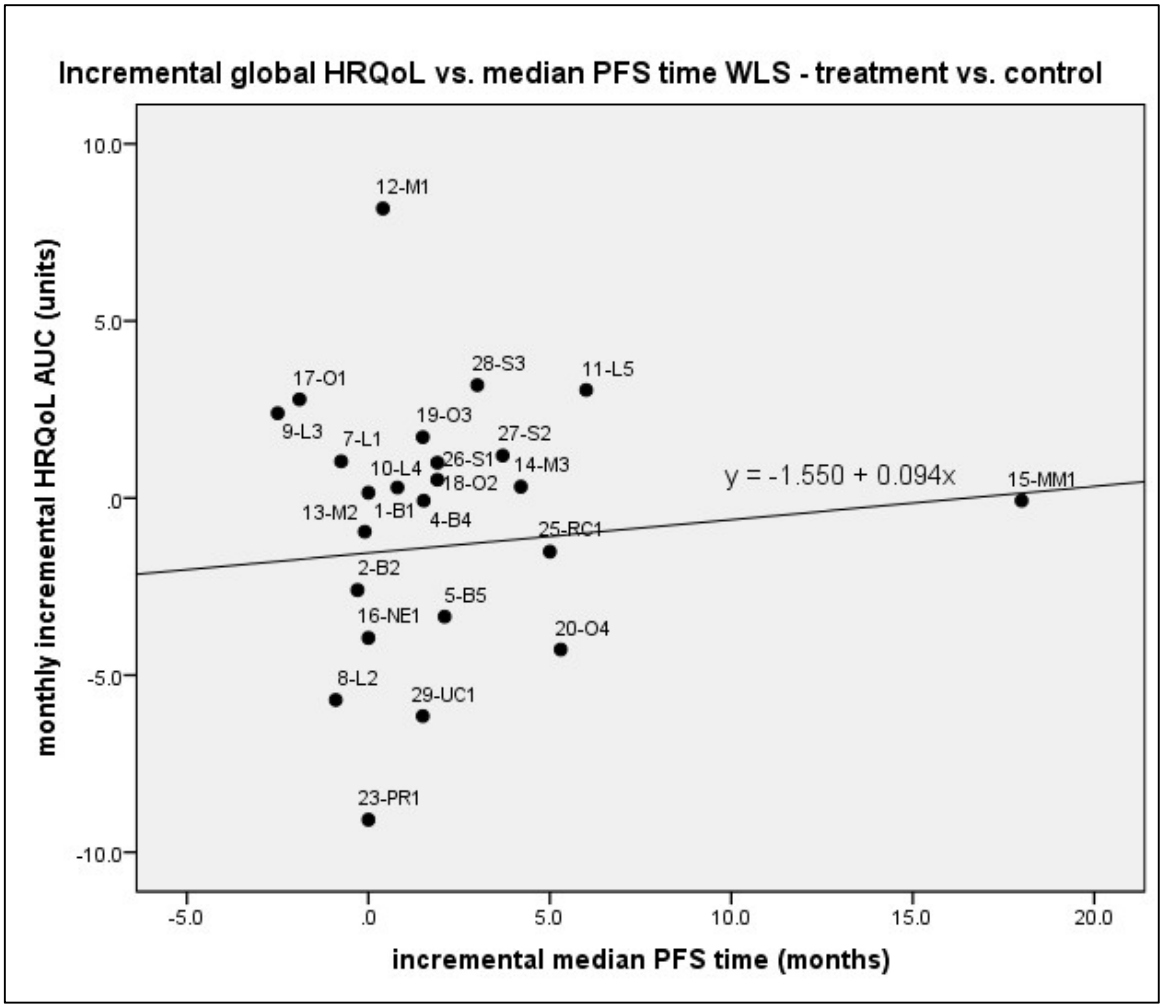


Figure 2: Physical HRQoL Regression Analysis (n=18)



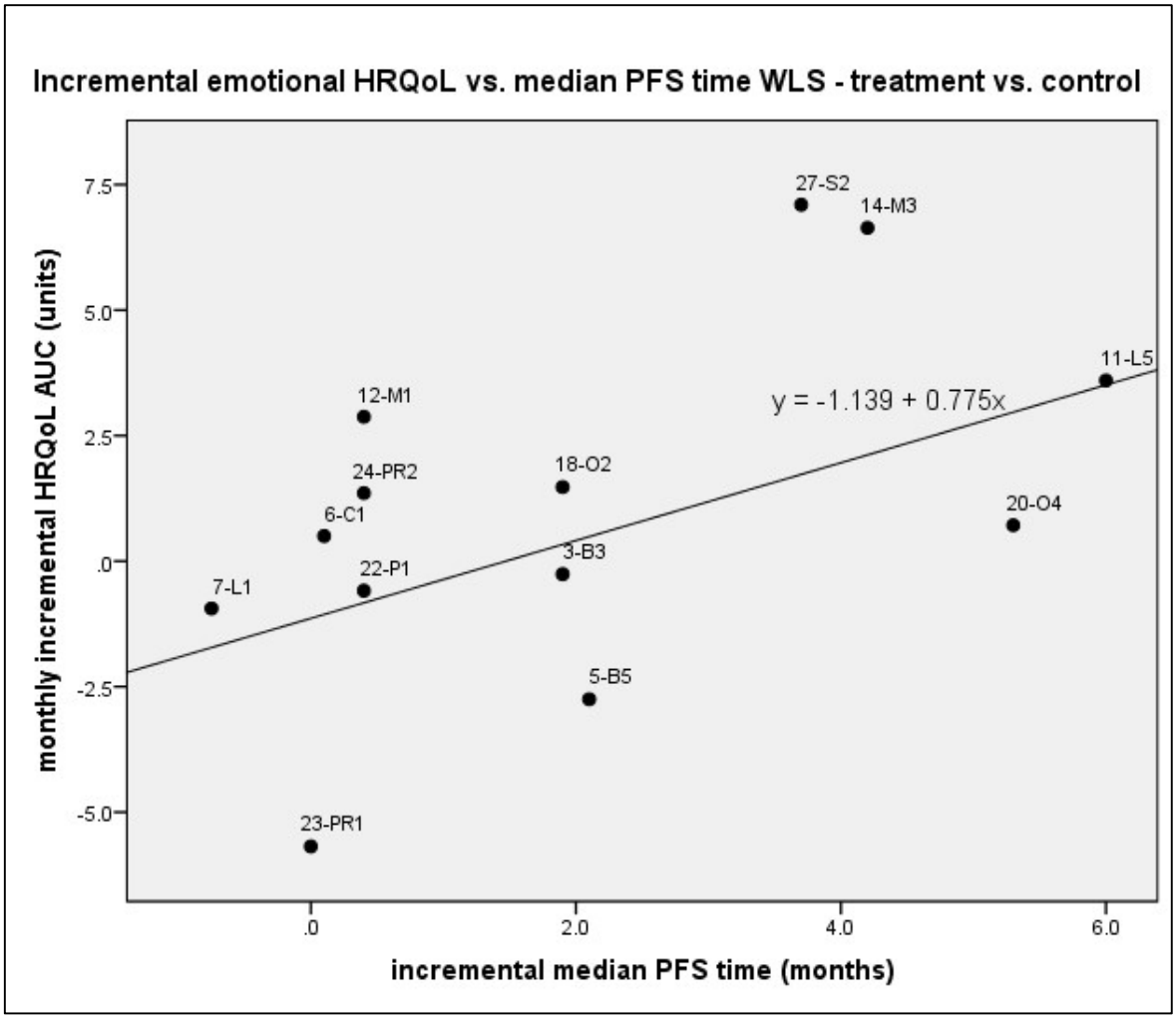
B = Breast, C = Colorectal, L = Lung, M = Melanoma, MM = Multiple Myeloma, NE = Neuroendocrine, O = Ovarian, P = Pancreas, PR = Prostate, RC = Renal Cell, S = Stomach, UC = Uterine Cervical, AUC = area under the curve, WLS = weighted least squares

Figure 3: Global HRQoL Regression Analysis (n=24)



B = Breast, C = Colorectal, L = Lung, M = Melanoma, MM = Multiple Myeloma, NE = Neuroendocrine, O = Ovarian, P = Pancreas, PR = Prostate, RC = Renal Cell, S = Stomach, UC = Uterine Cervical, AUC = area under the curve, WLS = weighted least squares

Figure 4: Emotional HRQoL Regression Analysis (n=13)



B = Breast, C = Colorectal, L = Lung, M = Melanoma, MM = Multiple Myeloma, NE = Neuroendocrine, O = Ovarian, P = Pancreas, PR = Prostate, RC = Renal Cell, S = Stomach, UC = Uterine Cervical, AUC = area under the curve, WLS = weighted least squares

Table 1: Summary of included studies

Trial ID	First Author, Publication Year	Age Range	Number of Patients	Cancer Type	Intervention (number of patients)	Comparator (number of patients)	Follow-up (months)	
							median PFS[A]	HRQoL
1-B1	Biganzoli L et al., 2002[22] Bottomley A et al., 2004[23]	28-70	273	Breast cancer	(AT) doxorubicin + paclitaxel (138)	(AC) doxorubicin + cyclophosphamide (135)	6	4.5
2-B2	Bottomley A et al., 2005[24]	26.1-79.8	448	Breast cancer	cyclophosphamide + epirubicin + filgrastim (224)	cyclophosphamide + epirubicin + fluorouracil (224)	33.7	34
3-B3	Cameron D et al., 2008[25] Zhou X et al., 2009[26]	26-83	528	Breast cancer	lapatinib + capecitabine (264)	Capecitabine (264)	6.2	6
4-B4	Di Leo A et al., 2008[27] Sherrill B et al., 2010[28]	23-87	579	Breast cancer	Lapatinib + paclitaxel (291)	placebo + Paclitaxel (288)	7.25	11.25
5-B5	Nuzzo F et al., 2011[29]	30-69	139	Breast cancer	(DOC) docetaxel (weekly) (70)	(DOC) docetaxel (3-weekly) (69)	15.2	1.5
6-C1	Comella P et al., 2008[30]	37-84	322	Colorectal cancer	(OXXEL) Oxaliplatin + capecitabine (158)	(OXAFAFU) Oxaliplatin + Xuorouracil or leucovorin (164)	6.6	6
7-L1	Wachters FM et al., 2003[31]	29-80	240	non-small-cell lung cancer (NSCLC)	Epirubicin + gemcitabine (121)	Cisplatin + gemcitabine (119)	5.75	5.25
8-L2	Lilenbaum RC et al., 2005[32]	42-86	165	non-small-cell lung cancer (NSCLC)	vinorelbine + gemcitabine (82)	carboplatin + paclitaxel (83)	3.9	3
9-L3	Maruyama R et al., 2008[33] Sekine I et al., 2009[34]	>=20	489	non-small-cell lung cancer (NSCLC)	Gefitinib (245)	Docetaxel (244)	11.5	3
10-L4	Han BH et al., 2011[35]	N/A	126	non-small-cell lung cancer (NSCLC)	Paclitaxel + Carbopolatin + Recombinant Human Endostatin (63)	(TC) Paclitaxel + Carboplatin (63)	7.1	2.25
11-L5	Sun JM et al., 2012[36]	30-78	135	non-small-cell lung cancer (NSCLC)	Gefitinib (68)	Pemetrexed (67)	9	1.5
12-M1	Middleton MR et al., 2000[37] Kiebert GM et al., 2003[38]	21-88	305	Melanoma	Temozolomide (156)	(DTIC) Dacarbazine (149)	1.9	6
13-M2	Avril MF et al., 2004[39]	18-79	229	Melanoma	Fotemustine (112)	(DTIC) Dacarbazine (117)	1.8	2
14-M3	Grob JJ et al., 2014[40]	>=18	250	Melanoma	Dabrafenib (187)	(DTIC) Dacarbazine (63)	6.9	3
15-MM1	Palumbo A et al., 2012[41] Dimopoulos MA et al., 2013[42]	65-91	306	Multiple Myeloma	(MPR-R) lenalidomide + melphalan + prednisone + maintenance lenalidomide (152)	(MP) melphalan + prednisone + maintenance placebo (154)	31	16
16-NE1	Arnold R et al., 2005[43]	18-77	109	Neuroendocrine foregut and midgut tumors	Octreotide + IFN-alpha (54)	Octreotide (55)	6	3
17-O1	du Bois A et al., 2003[44]	20.8-83.6	783	Ovarian cancer	(TC) Paclitaxel + Carboplatin (397)	(PT) Paclitaxel + Cisplatin (386)	17.2	2.25
18-O2	Pujade-Lauraine E et al., 2010[45] Brundage M et al., 2012[46]	24-82	976	Ovarian cancer	carboplatin + pegylated liposomal doxorubicin (PLD) (467)	carboplatin + paclitaxel (509)	11.3	6
19-O3	Monk BJ et al., 2010[47] Krasner CN et al., 2012[48]	>=18	672	Epithelial Ovarian, Fallopian Tube, or Primary Peritoneal carcinoma	Trabectedin + Pegylated liposomal doxorubicin (PLD) (337)	Pegylated liposomal doxorubicin (PLD) (335)	7.3	5
20-O4	Pokrzywinski R et al., 2011[49]	>=18	148	Ovarian cancer	Docetaxel + Carboplatin (74)	Docetaxel followed by Carboplatin (74)	13.7	5.25
21-O5	Burger RA et al., 2011[50] Monk BJ et al., 2013[51]	22-89	1248	Epithelial Ovarian, Primary Peritoneal, or Fallopian-tube cancer	chemotherapy + bevacizumab (623)	chemotherapy + placebo (625)	14.1	23

Trial ID	First Author, Publication Year	Age Range	Number of Patients	Cancer Type	Intervention (number of patients)	Comparator (number of patients)	Follow-up (months)	
							median PFS[A]	HRQoL
22-P1	Philip PA et al., 2010[52] Moinpour CM et al., 2010[53]	30-87	743	Pancreas adenocarcinoma	Gemcitabine + Cetuximab (372)	Gemcitabine (371)	3.4	4.25
23-PR1	Small EJ et al., 2002[54] Ahles TA et al., 2004[55]	40-85	256	Prostate cancer	suramin (high dose) (127)	suramin (low dose) (129)	3	5.5
24-PR2	Dawson N et al., 2010[56]	49-91	214	Prostate cancer	Zibotentan 10mg (107)	Placebo (107)	4	6
25-RC1	Cella D et al., 2012[57]	>=18	435	Renal Cell carcinoma	pazopanib (290)	placebo (145)	9.2	12
26-S1	Al-Batran SE et al., 2013[58]	>=65	143	Stomach or Oesophagogastric Junction adenocarcinoma	(FLOT) leucovorin + oxaliplatin + docetaxel (72)	(FLO) leucovorin + oxaliplatin (71)	9	6
27-S2	Bonnetain F et al., 2005[59]	37-76	90	Gastric adenocarcinoma	LV5FU2 + Irinotecan (45)	LV5FU2 (45)	6.9	6
28-S3	Jiang FS et al., 2009[60]	35-69	42	Stomach carcinoma	Recombinant human endostatin + XELOX (capecitabine + oxaliplatin) (20)	XELOX (capecitabine + oxaliplatin) (22)	6.9	1.5
29-UC1	Long III HJ et al., 2006[61]	N/A	123	Uterine cervix carcinoma	(MVAC) methotrexate + vinblastine + doxorubicin + cisplatin (63)	cisplatin (60)	4.4	9
30-UC2	Monk BJ et al., 2009[62] Cella D et al., 2010[63]	20-81	215	Cervical cancer	cisplatin + gemcitabine (112)	cisplatin + paclitaxel (103)	4.7	9.75

B = Breast, C = Colorectal, L = Lung, M = Melanoma, MM = Multiple Myeloma, NE = Neuroendocrine, O = Ovarian, P = Pancreas, PR = Prostate, RC = Renal Cell, S = Stomach, UC = Uterine Cervical, N/A = Not Available

[A]median PFS times provided for trial intervention arms used in quantitative analyses

Table 2: Regression Analyses Study Data

Trial ID	Incremental median PFS time in months	Incremental monthly HRQoL AUC units (monthly variance)		
		Physical HRQoL Domain (n=18)	Global HRQoL Domain (n=24)	Emotional HRQoL Domain (n=13)
1-B1	0.000	N/A	0.150 (0.046)	N/A
2-B2	-0.300	N/A	-2.599 (0.023)	N/A
3-B3	1.900	-1.071 (1.395)	N/A	-0.260 (1.141)
4-B4	1.525	0.627 (0.732)	-0.074 (0.560)	N/A
5-B5	2.100	-4.350 (10.827)	-3.350 (13.339)	-2.750 (17.052)
6-C1	0.100	5.000 (2.803)	N/A	0.500 (3.188)
7-L1	-0.750	-3.621 (10.091)	1.036 (6.212)	-0.943 (6.833)
8-L2	-0.900	N/A	-5.700 (3.131)	N/A
9-L3	-2.500	2.804 (1.989)	2.397 (1.206)	N/A
10-L4	0.800	N/A	0.297 (1.689)	N/A
11-L5	6.000	1.950 (10.168)	3.050 (12.022)	3.600 (11.190)
12-M1	0.400	11.275 (12.044)	8.175 (11.386)	2.875 (8.929)
13-M2	-0.100	6.160 (12.144)	-0.950 (11.174)	N/A
14-M3	4.200	0.370 (3.375)	0.320 (4.061)	6.638 (4.513)
15-MM1	18.000	-3.375 (1.394)	-0.075 (1.164)	N/A
16-NE1	0.000	N/A	-3.950 (29.415)	N/A
17-O1	-1.900	N/A	2.787 (1.198)	N/A
18-O2	1.900	1.725 (0.902)	1.000 (1.493)	1.475 (1.367)
19-O3	1.500	N/A	1.720 (3.341)	N/A
20-O4	5.300	-7.755 (2.579)	-4.276 (1.499)	0.714 (2.218)
21-O5	3.800	-0.718 (0.265)	N/A	N/A
22-P1	0.400	N/A	N/A	-0.588 (1.012)
23-PR1	0.000	-14.026 (2.498)	-9.083 (1.358)	-5.682 (2.314)
24-PR2	0.400	5.357 (4.442)	N/A	1.354 (4.902)
25-RC1	5.000	N/A	-1.513 (3.135)	N/A
26-S1	1.900	N/A	0.517 (7.951)	N/A
27-S2	3.700	6.000 (18.220)	1.200 (47.245)	7.100 (37.884)
28-S3	3.000	N/A	3.190 (2.612)	N/A
29-UC1	1.500	N/A	-6.155 (6.979)	N/A
30-UC2	-1.120	-2.289 (2.115)	N/A	N/A

B = Breast, C = Colorectal, L = Lung, M = Melanoma, MM = Multiple Myeloma, NE = Neuroendocrine, O = Ovarian, P = Pancreas, PR = Prostate, RC = Renal Cell, S = Stomach, UC = Uterine Cervical, N/A = Not Available, n = sample size

3.9 References

1. Peppercorn JM, Smith TJ, Helft PR, et al. American society of clinical oncology statement: toward individualized care for patients with advanced cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2011;29(6):755-60. doi: 10.1200/JCO.2010.33.1744

2. Booth CM, Eisenhauer EA. Progression-free survival: meaningful or simply measurable? *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2012;30(10):1030-3. doi: 10.1200/JCO.2011.38.7571

3. Pazdur R. Endpoints for assessing drug activity in clinical trials. *The oncologist* 2008;13 Suppl 2:19-21. doi: 10.1634/theoncologist.13-S2-19

4. Beckman M. More clinical cancer treatments judged by progression-free rather than overall survival. *J Natl Cancer Inst* 2007;99(14):1068-9. doi: 10.1093/jnci/djm073

5. Fallowfield LJ, Fleissig A. The value of progression-free survival to patients with advanced-stage cancer. *Nature reviews Clinical oncology* 2012;9(1):41-7. doi: 10.1038/nrclinonc.2011.156

6. Gutman SI, Piper M, Grant MD, et al. Progression-Free Survival: What Does It Mean for Psychological Well-Being or Quality of Life? Rockville (MD)2013.

7. Kim C, Prasad V. Strength of Validation for Surrogate End Points Used in the US Food and Drug Administration's Approval of Oncology Drugs. *Mayo Clin Proc* 2016 doi: 10.1016/j.mayocp.2016.02.012

8. Hotte SJ, Bjarnason GA, Heng DY, et al. Progression-free survival as a clinical trial endpoint in advanced renal cell carcinoma. *Curr Oncol* 2011;18 Suppl 2:S11-9.

9. Kovic B, Guyatt G, Brundage M, et al. Association between progression-free survival and health-related quality of life in oncology: a systematic review protocol. *BMJ Open* 2016;6(9):e012909. doi: 10.1136/bmjopen-2016-012909

10. Booth A, Clarke M, Dooley G, et al. The nuts and bolts of PROSPERO: an international prospective register of systematic reviews. *Syst Rev* 2012;1:2. doi: 10.1186/2046-4053-1-2

11. Booth A, Clarke M, Ghersi D, et al. An international registry of systematic-review protocols. *Lancet* 2011;377(9760):108-9. doi: 10.1016/S0140-6736(10)60903-8

12. Cochrane Handbook for Systematic Reviews of Interventions. In: Higgins JPT GSe, ed.: The Cochrane Collaboration, 2011.

13. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009;339:b2535. doi: 10.1136/bmj.b2535

14. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ* 2009;339:b2700. doi: 10.1136/bmj.b2700

15. Abridged Index Medicus (AIM or "Core Clinical") Journal Titles. 09 July 2003 ed: U.S. National Library of Medicine.

16. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33(1):159-74.

17. Lowry R. VassarStats - Kappa as a Measure of Concordance in Categorical Sorting; Kappa with Linear Weighting. 2016 [Available from: http://vassarstats.net/kappa.html accessed 31 August 2016.

18. Cella DF, Tulsky DS, Gray G, et al. The Functional Assessment of Cancer Therapy scale: development and validation of the general measure. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 1993;11(3):570-9. doi: 10.1200/JCO.1993.11.3.570

19. Oncology ASoC. Regarding: Draft Guidance for industry on "Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics" [Docket No. 2005D-0112] 2006.

20. Aaronson NK, Ahmedzai S, Bergman B, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst* 1993;85(5):365-76.

21. Luckett T, King MT, Butow PN, et al. Choosing between the EORTC QLQ-C30 and FACT-G for measuring health-related quality of life in cancer clinical research: issues, evidence and recommendations. *Ann Oncol* 2011;22(10):2179-90. doi: 10.1093/annonc/mdq721

22. Biganzoli L, Cufer T, Bruning P, et al. Doxorubicin and paclitaxel versus doxorubicin and cyclophosphamide as first-line chemotherapy in metastatic breast cancer: The European Organization for Research and Treatment of Cancer 10961 Multicenter Phase III Trial. *Journal of clinical oncology : official journal of the American*

Society of Clinical Oncology 2002;20(14):3114-21. doi: 10.1200/JCO.2002.11.005

23. Bottomley A, Biganzoli L, Cufer T, et al. Randomized, controlled trial investigating short-term health-related quality of life with doxorubicin and paclitaxel versus doxorubicin and cyclophosphamide as first-line chemotherapy in patients with metastatic breast cancer: European Organization for Research and Treatment of Cancer Breast Cancer Group, Investigational Drug Branch for Breast Cancer and the New Drug Development Group Study. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2004;22(13):2576-86. doi: 10.1200/JCO.2004.02.037
24. Bottomley A, Therasse P, Piccart M, et al. Health-related quality of life in survivors of locally advanced breast cancer: an international randomised controlled phase III trial. *Lancet Oncol* 2005;6(5):287-94. doi: 10.1016/S1470-2045(05)70100-5
25. Cameron D, Casey M, Press M, et al. A phase III randomized comparison of lapatinib plus capecitabine versus capecitabine alone in women with advanced breast cancer that has progressed on trastuzumab: updated efficacy and biomarker analyses. *Breast Cancer Res Treat* 2008;112(3):533-43. doi: 10.1007/s10549-007-9885-0
26. Zhou X, Cella D, Cameron D, et al. Lapatinib plus capecitabine versus capecitabine alone for HER2+ (ErbB2+) metastatic breast cancer: quality-of-life assessment. *Breast Cancer Res Treat* 2009;117(3):577-89. doi: 10.1007/s10549-009-0310-8

27. Di Leo A, Gomez HL, Aziz Z, et al. Phase III, double-blind, randomized study comparing lapatinib plus paclitaxel with placebo plus paclitaxel as first-line treatment for metastatic breast cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2008;26(34):5544-52. doi: 10.1200/JCO.2008.16.2578

28. Sherrill B, Di Leo A, Amonkar MM, et al. Quality-of-life and quality-adjusted survival (Q-TWiST) in patients receiving lapatinib in combination with paclitaxel as first-line treatment for metastatic breast cancer. *Curr Med Res Opin* 2010;26(4):767-75. doi: 10.1185/03007991003590860

29. Nuzzo F, Morabito A, Gravina A, et al. Effects on quality of life of weekly docetaxel-based chemotherapy in patients with locally advanced or metastatic breast cancer: results of a single-centre randomized phase 3 trial. *BMC Cancer* 2011;11:75. doi: 10.1186/1471-2407-11-75

30. Comella P, Massidda B, Filippelli G, et al. Randomised trial comparing biweekly oxaliplatin plus oral capecitabine versus oxaliplatin plus i.v. bolus fluorouracil/leucovorin in metastatic colorectal cancer patients: results of the Southern Italy Cooperative Oncology study 0401. *J Cancer Res Clin Oncol* 2009;135(2):217-26. doi: 10.1007/s00432-008-0454-7

31. Wachters FM, Van Putten JW, Kramer H, et al. First-line gemcitabine with cisplatin or epirubicin in advanced non-small-cell lung cancer: a phase III trial. *Br J Cancer* 2003;89(7):1192-9. doi: 10.1038/sj.bjc.6601283

32. Lilenbaum RC, Chen CS, Chidiac T, et al. Phase II randomized trial of vinorelbine and gemcitabine versus carboplatin and paclitaxel in advanced non-small-cell lung cancer. *Ann Oncol* 2005;16(1):97-101. doi: 10.1093/annonc/mdi009

33. Maruyama R, Nishiwaki Y, Tamura T, et al. Phase III study, V-15-32, of gefitinib versus docetaxel in previously treated Japanese patients with non-small-cell lung cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2008;26(26):4244-52. doi: 10.1200/JCO.2007.15.0185

34. Sekine I, Ichinose Y, Nishiwaki Y, et al. Quality of life and disease-related symptoms in previously treated Japanese patients with non-small-cell lung cancer: results of a randomized phase III study (V-15-32) of gefitinib versus docetaxel. *Ann Oncol* 2009;20(9):1483-8. doi: 10.1093/annonc/mdp031

35. Han BH, Xiu QY, Wang HM, et al. [A multicenter, randomized, double-blind, placebo-controlled safety study to evaluate the clinical effects and quality of life of paclitaxel-carboplatin (PC) alone or combined with endostar for advanced non-small cell lung cancer (NSCLC)]. *Zhonghua Zhong Liu Za Zhi* 2011;33(11):854-9.

36. Sun JM, Lee KH, Kim SW, et al. Gefitinib versus pemetrexed as second-line treatment in patients with nonsmall cell lung cancer previously treated with platinum-based chemotherapy (KCSG-LU08-01): an open-label, phase 3 trial. *Cancer* 2012;118(24):6234-42. doi: 10.1002/cncr.27630

37. Middleton MR, Grob JJ, Aaronson N, et al. Randomized phase III study of temozolomide versus dacarbazine in the treatment of patients with advanced metastatic malignant melanoma. *Journal of clinical oncology : official journal of*

the American Society of Clinical Oncology 2000;18(1):158-66. doi: 10.1200/JCO.2000.18.1.158

38. Kiebert GM, Jonas DL, Middleton MR. Health-related quality of life in patients with advanced metastatic melanoma: results of a randomized phase III study comparing temozolomide with dacarbazine. *Cancer Invest* 2003;21(6):821-9.
39. Avril MF, Aamdal S, Grob JJ, et al. Fotemustine compared with dacarbazine in patients with disseminated malignant melanoma: a phase III study. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2004;22(6):1118-25. doi: 10.1200/JCO.2004.04.165
40. Grob JJ, Amonkar MM, Martin-Algarra S, et al. Patient perception of the benefit of a BRAF inhibitor in metastatic melanoma: quality-of-life analyses of the BREAK-3 study comparing dabrafenib with dacarbazine. *Ann Oncol* 2014;25(7):1428-36. doi: 10.1093/annonc/mdu154
41. Palumbo A, Hajek R, Delforge M, et al. Continuous lenalidomide treatment for newly diagnosed multiple myeloma. *N Engl J Med* 2012;366(19):1759-69. doi: 10.1056/NEJMoa1112704
42. Dimopoulos MA, Delforge M, Hajek R, et al. Lenalidomide, melphalan, and prednisone, followed by lenalidomide maintenance, improves health-related quality of life in newly diagnosed multiple myeloma patients aged 65 years or older: results of a randomized phase III trial. *Haematologica* 2013;98(5):784-8. doi: 10.3324/haematol.2012.074534

43. Arnold R, Rinke A, Klose KJ, et al. Octreotide versus octreotide plus interferon-alpha in endocrine gastroenteropancreatic tumors: a randomized trial. *Clin Gastroenterol Hepatol* 2005;3(8):761-71.

44. du Bois A, Luck HJ, Meier W, et al. A randomized clinical trial of cisplatin/paclitaxel versus carboplatin/paclitaxel as first-line treatment of ovarian cancer. *J Natl Cancer Inst* 2003;95(17):1320-9.

45. Pujade-Lauraine E, Wagner U, Aavall-Lundqvist E, et al. Pegylated liposomal Doxorubicin and Carboplatin compared with Paclitaxel and Carboplatin for patients with platinum-sensitive ovarian cancer in late relapse. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2010;28(20):3323-9. doi: 10.1200/JCO.2009.25.7519

46. Brundage M, Gropp M, Mefti F, et al. Health-related quality of life in recurrent platinum-sensitive ovarian cancer--results from the CALYPSO trial. *Ann Oncol* 2012;23(8):2020-7. doi: 10.1093/annonc/mdr583

47. Monk BJ, Herzog TJ, Kaye SB, et al. Trabectedin plus pegylated liposomal Doxorubicin in recurrent ovarian cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2010;28(19):3107-14. doi: 10.1200/JCO.2009.25.4037

48. Krasner CN, Poveda A, Herzog TJ, et al. Patient-reported outcomes in relapsed ovarian cancer: results from a randomized Phase III study of trabectedin with pegylated liposomal doxorubicin (PLD) versus PLD alone. *Gynecol Oncol* 2012;127(1):161-7. doi: 10.1016/j.ygyno.2012.06.034

49. Pokrzywinski R, Secord AA, Havrilesky LJ, et al. Health-related quality of life outcomes of docetaxel/carboplatin combination therapy vs. sequential therapy with docetaxel then carboplatin in patients with relapsed, platinum-sensitive ovarian cancer: results from a randomized clinical trial. *Gynecol Oncol* 2011;123(3):505-10. doi: 10.1016/j.ygyno.2011.08.015

50. Burger RA, Brady MF, Bookman MA, et al. Incorporation of bevacizumab in the primary treatment of ovarian cancer. *N Engl J Med* 2011;365(26):2473-83. doi: 10.1056/NEJMoa1104390

51. Monk BJ, Huang HQ, Burger RA, et al. Patient reported outcomes of a randomized, placebo-controlled trial of bevacizumab in the front-line treatment of ovarian cancer: a Gynecologic Oncology Group Study. *Gynecol Oncol* 2013;128(3):573-8. doi: 10.1016/j.ygyno.2012.11.038

52. Philip PA, Benedetti J, Corless CL, et al. Phase III study comparing gemcitabine plus cetuximab versus gemcitabine in patients with advanced pancreatic adenocarcinoma: Southwest Oncology Group-directed intergroup trial S0205. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2010;28(22):3605-10. doi: 10.1200/JCO.2009.25.7550

53. Moinpour CM, Vaught NL, Goldman B, et al. Pain and emotional well-being outcomes in Southwest Oncology Group-directed intergroup trial S0205: a phase III study comparing gemcitabine plus cetuximab versus gemcitabine as first-line therapy in patients with advanced pancreas cancer. *Journal of clinical oncology :*

official journal of the American Society of Clinical Oncology 2010;28(22):3611-6. doi: 10.1200/JCO.2009.25.8285

54. Small EJ, Halabi S, Ratain MJ, et al. Randomized study of three different doses of suramin administered with a fixed dosing schedule in patients with advanced prostate cancer: results of intergroup 0159, cancer and leukemia group B 9480. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2002;20(16):3369-75. doi: 10.1200/JCO.2002.10.022
55. Ahles TA, Herndon JE, 2nd, Small EJ, et al. Quality of life impact of three different doses of suramin in patients with metastatic hormone-refractory prostate carcinoma: results of Intergroup O159/Cancer and Leukemia Group B 9480. *Cancer* 2004;101(10):2202-8. doi: 10.1002/cncr.20655
56. Dawson N, Payne H, Battersby C, et al. Health-related quality of life in pain-free or mildly symptomatic patients with metastatic hormone-resistant prostate cancer following treatment with the specific endothelin A receptor antagonist zibotentan (ZD4054). *J Cancer Res Clin Oncol* 2011;137(1):99-113. doi: 10.1007/s00432-010-0864-1
57. Cella D, Pickard AS, Duh MS, et al. Health-related quality of life in patients with advanced renal cell carcinoma receiving pazopanib or placebo in a randomised phase III trial. *Eur J Cancer* 2012;48(3):311-23. doi: 10.1016/j.ejca.2011.05.017
58. Al-Batran SE, Pauligk C, Homann N, et al. The feasibility of triple-drug chemotherapy combination in older adult patients with oesophagogastric cancer: a

randomised trial of the Arbeitsgemeinschaft Internistische Onkologie (FLOT65+). *Eur J Cancer* 2013;49(4):835-42. doi: 10.1016/j.ejca.2012.09.025

59. Bonnetain F, Bouche O, Conroy T, et al. Longitudinal quality of life study in patients with metastatic gastric cancer. Analysis modalities and clinical applicability of QoL in randomized phase II trial in a digestive oncology. *Gastroenterol Clin Biol* 2005;29(11):1113-24.

60. Jiang FS, Wang G, Sun YB, et al. Clinical observation on recombinant human vascular endostatin combined wih XELOX regimen in the treatment of advanced stomach carcinoma. [Chinese]. *Tumor* 2009;29(8):790-92. doi: http://dx.doi.org/10.3781/j.issn.1000-7431.2009.08.018

61. Long HJ, 3rd, Monk BJ, Huang HQ, et al. Clinical results and quality of life analysis for the MVAC combination (methotrexate, vinblastine, doxorubicin, and cisplatin) in carcinoma of the uterine cervix: A Gynecologic Oncology Group study. *Gynecol Oncol* 2006;100(3):537-43. doi: 10.1016/j.ygyno.2005.09.023

62. Monk BJ, Sill MW, McMeekin DS, et al. Phase III trial of four cisplatin-containing doublet combinations in stage IVB, recurrent, or persistent cervical carcinoma: a Gynecologic Oncology Group study. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2009;27(28):4649-55. doi: 10.1200/JCO.2009.21.8909

63. Cella D, Huang HQ, Monk BJ, et al. Health-related quality of life outcomes associated with four cisplatin-based doublet chemotherapy regimens for stage IVB

recurrent or persistent cervical cancer: a Gynecologic Oncology Group study. *Gynecol Oncol* 2010;119(3):531-7. doi: 10.1016/j.ygyno.2010.08.020

64. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *J Clin Epidemiol* 2009;62(10):1006-12. doi: 10.1016/j.jclinepi.2009.06.005

65. Guyatt G, Montori V, Devereaux PJ, et al. Patients at the center: in our practice, and in our use of language. *ACP J Club* 2004;140(1):A11-2.

66. Sprangers MA. Quality-of-life assessment in oncology. Achievements and challenges. *Acta Oncol* 2002;41(3):229-37.

67. Oliver A, Greenberg CC. Measuring outcomes in oncology treatment: the importance of patient-centered outcomes. *Surg Clin North Am* 2009;89(1):17-25, vii. doi: 10.1016/j.suc.2008.09.015

68. Bottomley A. The cancer patient and quality of life. *The oncologist* 2002;7(2):120-5.

69. Gondek K, Sagnier PP, Gilchrist K, et al. Current status of patient-reported outcomes in industry-sponsored oncology clinical trials and product labels. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2007;25(32):5087-93. doi: 10.1200/JCO.2007.11.3845

70. Bernhard J, Cella DF, Coates AS, et al. Missing quality of life data in cancer clinical trials: serious problems and challenges. *Stat Med* 1998;17(5-7):517-32.

71. Osoba D, Bezjak A, Brundage M, et al. 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. doi: 10.1111/j.1524-4733.2007.00278.x

3.10 Appendix 1

Properties of expectation and variance is represented by the formula $v(ax) = a^2v(x)$, where "a" equals the number used to standardize the scale out of 100, and "v(x)" equals the variance of the original unstandardized HRQoL measurement. This formula was used in the development of variance formulae as described in our published protocol.

Properties of expectation and variance of a random variable were also used to calculate within group area variance error estimates for each HRQoL treatment arm group, as per formula (4) [$V(\text{Area}_{\text{I or C}}) = d_1^2 \text{ X } (\sigma_1^2 + \sigma_2^2 + 2 \text{ X } \sigma_1 \text{ X } \sigma_2 \text{ X } p) + d_2^2 \text{ X } (\sigma_2^2 + \sigma_3^2 + 2 \text{ X } \sigma_2 \text{ X } \sigma_3 \text{ X } p) + \ldots$] in Appendix 2 of our protocol, by utilizing population variance for HRQoL measurements. Because all abstracted studies provided sample error estimates for each HRQoL measurement, rather than population error estimates, we corrected formula (4) by scaling sample variances (σ^2) by their respective sample sizes (i.e. dividing sample variance for each HRQoL measurement by the sample size of that measurement) so as to convert these to population variances for use in our formula.

We also adjusted incremental HRQoL benefit for all studies to per month values by utilizing the same properties of expectation and variance approach as in the instrument scale error estimate standardization, but in this case "a" equalled the adjustment factor used to adjust HRQoL benefit to per month values.

For a further description of the statistical methodology used in our analyses, please refer to Appendix 2 of our published protocol.

CHAPTER 4:

A descriptive survey and exploration of oncology trials reporting Health Related Quality of Life data.

Bruno Kovic,[1] Xuejing Jin,[1] Sean A. Kennedy,[2] Mathieu Hylands,[3] Michał Pędziwiatr,[4,5] Akira Kuriyama,[6] Huda Gomaa,[7] Yung Lee,[8] Morihiro Katsura,[9] Masafumi Tada,[10] Brian Y. Hong,[11] Sung Min Cho,[12] Patrick Jiho Hong,[11] Ashley Yu,[11] Yasmin Sivji,[13] Augustin Toma,[14] Li Xie,[15,16] Ludwig Tsoi,[17] Marcin Waligora,[18] Manya Prasad,[19] Neera Bhatnagar,[20] Lehana Thabane,[1] Michael Brundage,[21] Gordon Guyatt,[1,8] Feng Xie [1,22]

Author affiliations

[1]Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Ontario, Canada

[2]Diagnostic Radiology Residency Program, University of Toronto, Toronto, Ontario, Canada

[3]Department of Surgery, Université de Sherbrooke, Sherbrooke, Québec, Canada

[4]2nd Department of General Surgery, Jagiellonian University, Krakow, Poland

[5]Centre for Research, Training and Innovation in Surgery (CERTAIN Surgery). Krakow, Poland

[6]Department of General Medicine, Kurashiki Central Hospital, Kurashiki, Japan

[7]Department of pharmacy, drug information center, Tanta Chest Hospital, Gharbia, Egypt

[8]Faculty of Health Sciences, McMaster University, Hamilton, Ontario, Canada

[9]Department of Surgery, Okinawa Prefectural Nanbu Medical Center & Children's Medical Center, Okinawa, Japan

[10]Department of Health Promotion and Human Behavior, Kyoto University Graduate School of Medicine / School of Public Health, Kyoto, Japan

[11]Faculty of Medicine, University of Ottawa, Ottawa, Ontario, Canada

[12]Faculty of Medicine, University of Toronto, Toronto, Ontario, Canada

[13]Department of Medicine, McMaster University, Hamilton, Ontario, Canada

[14]Michael G. DeGroote School of Medicine, McMaster University, Hamilton, Ontario, Canada

[15]Department of Cancer Prevention, Fudan University Shanghai Cancer Center, Shanghai, China

[16]Department of Oncology, Shanghai Medical College, Shanghai, China

[17]A&E Department, Queen Mary Hospital, Hong Kong

[18]Department of Philosophy and Bioethics, Jagiellonian University Medical College, REMEDY, Research Ethics in Medicine Study Group, Krakow, Poland

[19]Department of Community Medicine, Postgraduate Institute of Medical Sciences, Rohtak, Haryana, India

[20]Health Sciences Library, McMaster University, Hamilton, Ontario, Canada

[21]Department of Oncology, Queen's University, Kingston, Ontario, Canada

[22]The Research Institute of St. Joseph's Healthcare, Hamilton, Ontario, Canada

Correspondence to Dr Feng Xie; fengxie@mcmaster.ca

Research Support None

4.1 Abstract

Background

We performed a systematic review of randomized controlled trials (RCTs) to examine the association between progression-free survival (PFS) and health-related quality of life (HRQoL) in oncology. In this article, we report on the observed variation in risk of bias across trials with the goal of guiding the future design and conduct of similar trials.

Methods

We searched MEDLINE, Embase, and Cochrane databases for randomized controlled trials (RCTs) published since 2000 addressing chemotherapy or biological treatments reporting PFS or HRQoL with no OS benefits. We extracted study characteristics and risk of bias data across 8 risk of bias domains. We documented study characteristics, including the type of progression assessment and type of HRQoL instrument used and risk of bias assessments for eligible trials considering source of funding, cancer types and stages, and ECOG status.

Results

Of 35,960 records identified, 42 articles reporting on 30 RCTs involving 10,731 patients across 12 cancer types proved eligible. Of the 30 trials, 26 (87%) had high risk of bias in at least one of four domains: loss to follow-up, failure to follow patients after progression, blinding of participants and personnel, and blinding of outcome assessment. Attrition bias proved the most common type of risk of bias across studies: 24 of 30 (80%) trials having

high risk of bias in at least one of two attrition domains. Blinding domains proved the second most frequent type of risk of bias: 10 of 30 (33%) trials had high risk of bias in at least one of two blinding domains. Trials that were not industry funded and addressed lung cancer were more likely to fail to blind interventions. Trials addressing breast and ovarian cancer, and trials with lower ECOG (i.e. less severe) patients had higher risk of attrition related risk of bias.

Conclusions

When designing and conducting oncology RCTs designed to measure HRQoL, investigators must if possible blind their studies, and should certainly institute strategies to minimize attrition and reduce missing data for patient reported questionnaires after disease progression in clinical trials. In deciding on the credibility of results, clinicians need to consider these risk of bias issues.

4.2 Introduction

The goal of patient-centred cancer care is to extend survival or improve health-related quality of life (HRQoL),[1-3] although regulatory authorities have also approved cancer treatments on the basis of surrogates such as progression-free survival (PFS).[3] We previously performed a systematic review of randomized controlled trials (RCTs) to examine the association between PFS and HRQoL in oncology. Results failed to find an association between PFS and HRQoL.

Limited evidence suggests that oncology RCTs often fail to measure HRQoL and, when they do, manifest limitations regarding methodology and reporting.[4 5 6] Problems in study design and implementation can bias estimates of effect on outcomes of PFS and HRQoL, and on cancer-specific and all-cause mortality. Detailed evaluation of the performance of trials that address both HRQoL and PFS may help to guide the design and development of the next generation of oncology trials, in particular trials examining the question of association between HRQoL and PFS.

The database of completed abstractions from our systematic review provides an opportunity to further explore study characteristics, and in particular the extent of bias in these RCTs. In this article, we characterize the oncology RCTs that proved eligible for our study, focusing on risk of bias issues, including exploration of variation in risk of bias for blinding and attrition domains by location of cancer, industry funding, Eastern Cooperative Oncology Group (ECOG) status, and cancer stage.

4.3 Methods

We have previously published our study protocol, detailing our systematic review design and analysis,[7] and also registered our protocol on PROSPERO (International Prospective Register of Ongoing Systematic Reviews)[8 9] with registration number CRD42016047162. Our systematic review adhered to the PRISMA Statement[10 11] guidelines to ensure transparent and complete reporting.

Briefly, we conducted a review of human cancer RCTs published from 2000 to May 04, 2016, utilizing standard methodology as described by the Cochrane Collaboration.[12] We included trials examining chemotherapy or biological type cancer therapies as the primary agent in treating cancer (i.e. primary investigational treatment), and that reported on PFS and HRQoL, but excluded the trials with statistically significant OS benefits.

Objectives

This exploration of the trials included in our systematic review, has two objectives, and is accordingly structured in two parts:

1. To report a full summary of study characteristics collected but undisclosed in our previous paper, including risk of bias assessments (as per chapter 8 of the Cochrane handbook for Systematic Reviews of Interventions).

2. To describe and explore the distribution of four critical domains[6] of risk of bias (i.e. blinding of participants and personnel, blinding of outcome assessment, loss to follow-up issues defined as >20% attrition, and failure to follow patients after progression), across different trial characteristics.

Objective 2 includes examining possible associations between the critical risk of bias domains across four trial characteristics: oncology type (differing body location), industry funding (yes, no), ECOG status (0, 1) and cancer stage (increasing severity).

Search and data collection

We used the OVID platform to search in MEDLINE, Embase, and Wiley's Cochrane databases, and used both MeSH terms as well as free text words to capture RCTs published in the *Abridged Index Medicus*.[13] Ten pairs of reviewers, trained through team meetings and email communications, used standardized and pilot tested forms constructed as per eligibility criteria in our protocol, as well as detailed written instructions to independently screen titles, abstracts, and full texts and to extract data. To ensure consistency, meetings included calibration exercises with sample data review exercises. Reviewer disagreement was resolved by either discussion or adjudication by an arbitrator (BK).

Reviewers extracted all data relevant to primary outcomes of PFS and HRQoL including median PFS time and/or PFS hazard ratios (HRs) for every intervention and control

group. Abstracted data included industry funding, cancer types and stages, ECOG status, and risk of bias. The main, or largest, group of patients enrolled in each trial were extracted for the characteristics of cancer stage and ECOG status.

Risk of bias appraisal

The Cochrane Risk of Bias tool was used to appraise the likelihood that the RCT results were affected by bias,[14] and assessments were made for individual risk of bias domains using our final 42 eligible articles representing 30 RCTs. Following training and calibration carried out as recommended in the handbook[12] with a pilot sample of three articles from our database of RCTs, pairs of reviewers assessed, independently and in duplicate, risk of bias.

The seven item tool was expanded to include an additional source of potential bias particularly relevant to studies examining HRQoL, that of following patients after progression.[14] Outcome specific evaluations focused on HRQoL and PFS outcomes.

Data analysis

We documented the performance of all eligible trials in each risk of bias domain, and whether there were any apparent differences in risk of bias in studies funded by industry, and those addressing patients with particular disease severity (i.e. through increased ECOG status & cancer stage).

Bar charts were constructed for the different subgroups within these characteristics at the trial level, and included any subgroup containing n>1 trials so proportions of low versus high risk of bias could be reported per characteristic. We assessed individual risk of bias domains by the frequency of trials in the low versus high risk of bias categories, excluding trials with unclear assessments.

We reported extracted information using both Microsoft Excel (version 15.0.4937.1000, Microsoft Corporation, Redmond, Washington, USA) and Review Manager (RevMan) [Computer program]. Version 5.3.5 Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014.

4.4 Results

Our search and review identified 42 eligible articles representing 30 RCTs (for article selection process see appendix 1 - PRISMA Flow Diagram[15]), with substantial inter-rater agreement (κ=0.75).[16]

Table 1 presents a summary of previously unreported study characteristics for all trials organized alphabetically by cancer type and publication year, with Trial ID format corresponding to number of the trial out of 30, followed by cancer type and trial number for the specific cancer type. The 30 eligible trials enrolled from 42 to 1248 adult patients, were published between 2000 and 2014, and involved 10,731 patients aged 18 to 91 across 12 cancer types. Cancer type trial distribution consisted of: 5 breast trials, 1 colorectal, 5 lung, 3 melanoma, 1 multiple myeloma, 1 neuroendocrine, 5 ovarian, 1 pancreas, 2 prostate, 1 renal cell, 3 stomach, and 2 uterine cervical type RCTs. Cancer progression was assessed by radiographic methods in 7 (23%) trials, by clinical methods in 2 (7%) trials, and by mixed methods in 10 (33%) trials. EORTC QLQ-C30 was the most common HRQoL instrument, used in 17 (57%) trials; FACT was used in 9 (30%) trials, and the remaining 4 (13%) trials used the Lung Cancer Symptom Scale, the eight-item linear analog self-assessment (LASA) questionnaire, and the clinician-reported Karnofsky score. Among the 30 eligible trials, 17 (57%) were industry funded, while 5 (17%) were not, and 8 (27%) did not report funding source. Five (17%) trials enrolled the largest group of patients in stage III, 18 (57%) in stage IV, and the other seven (23%) did

not report the largest patient stage group. Ten (33%) trials enrolled the largest group of patients with ECOG status 0, 8 (27%) trials with ECOG status 1, and 12 trials did not report the largest ECOG status patient group.

Appendix 2 presents risk of bias summary table assessments, showing review authors judgements about each risk of bias item for each included study. Figure 1 presents review authors' judgements about each risk of bias item as percentages. Of the 30 trials, 26 (87%) had high risk of bias in at least one of four critical domains (i.e. loss to follow-up, failure to follow patients after progression, blinding of participants and personnel, and blinding of outcome assessment). Attrition bias was the highest type of risk of bias across studies, with 24 of 30 (80%) trials having high risk of bias in one or both of the two attrition domains. The incomplete outcome data domain, defined as loss to follow-up (LTFU) greater than 20% or LTFU issues identified by study, had the greatest frequency of high risk of bias (19, 63% of trials), while failure to follow patients after progression occurred in 11 (37%) trials. 10 of these 11 trials (>90%) had high risk of bias due to low HRQoL measurement compliance rates, with 7 of these designed not to measure HRQoL past progression, and the remaining 3 either not performing or excluding HRQoL measurements past progression due to poor compliance rates of 27%-66%.

Blinding domains were the second most frequent type of risk of bias across studies, with 10 of 30 (33%) trials having high risk of bias in at least one of two blinding domains. The blinding of participants and personnel domain had high risk of bias in 10 (33%) trials,

while the blinding of outcome assessment had high risk of bias in 8 (27%) trials. The selective reporting domain also had 8 (27%) trials with high risk of bias, and was the only other domain with a large proportion of trials having high risk of bias. All other domains had 10% or less of trials with high risk of bias.

Figures 2-5 present risk of bias proportions, for trials reporting low or high risk of bias, for four trial characteristics of industry funding, cancer type, ECOG status, and disease stage across the four critical domains of: blinding of participants and personnel, blinding of outcome assessment, loss to follow-up issues defined as >20% attrition, and failure to follow patients after progression. Industry funded trials had lower proportions of high risk of bias for blinding domains, although attrition domains were not consistent with blinding domains or each other. The cancer type graph, organized by highest risk of bias per cancer type shows breast and ovarian cancer trials have the highest proportions of high risk of bias for attrition related domains, but the lowest proportions for blinding domains. Lung cancer trials are the opposite, having the highest proportions of high risk of bias for blinding domains, but the lowest proportions for attrition related domains. ECOG status had consistently greater proportions of high risk of bias for more functional (i.e. less severe) patients represented by ECOG status equals 0 as compared to ECOG status equals 1. Finally, there were no noticeable trends in blinding or attrition for disease stage.

4.5 Discussion

Our exploration of oncology RCTs that have reported results for both PFS and HRQoL found very frequent high risk of bias related to attrition, with considerably fewer, but still concerning, limitations in blinding. Non-industry funded trials showed a lower frequency of effective blinding practices compared to industry funded trials. Breast and ovarian cancer trials had the most attrition issues, and lung cancer trials had the worse blinding practices. Finally, trials with more functional patients generally had higher risk of bias.

Our study has several strengths. First, we conducted an exhaustive search with no language limitations. Second, we developed explicit eligibility criteria and conducted duplicate assessment of eligibility and data abstraction that included use of standardized and pilot-tested screening and data abstraction forms, review team meetings and communications, as well as calibration exercises to enhance the consistency between reviewers. The eligible trials proved to have a wide range of patients and trial characteristics, contributing to the generalizability of results. Finally, we used the widely accepted Cochrane criteria to assess risk of bias, including a domain (follow-up after recurrence) specific for cancer oncology trials, most particularly those measuring HRQoL.

Our study also has limitations. First, with only 30 total eligible RCTs, we could not address whether chance explained difference in findings across study characteristics

statistical analyses, and therefore were limited to only a descriptive review. Thus, there is the possibility that any differences or trends we observed could be explained by chance. Second, bar charts compared proportions of low versus high risk of bias per domain, and excluded trials with unclear assessments. Since unclear risk of bias assessments varied for each domain across trials, this led to a varying number of trials per domain, depending on the number of trials missing due to unclear assessments. Previous work has shown that, with respect to blinding, reviewers can make accurate inferences when trial reporting is not altogether explicit (obviating the need for "unclear" risk of bias ratings).[17] Inferences regarding the likely extent of blinding and other risk of bias issues on the basis of what was reported, or through contact with authors, may have been informative.

Thus far, only one previous study by the Agency for Healthcare Research and Quality (AHRQ) has addressed the question of PFS-HRQoL association in oncology.[6] Providing the comprehensive description of primary studies included in our review of this association could facilitate further exploration of relationships of relevance should additional relevant data become available in the future. The AHRQ study reported high risk of bias in all of their eligible trials. Our results provide a more nuanced examination of risk of bias issues and are consistent with other studies examining the design and conduct of oncology trials that that have reported frequent methodological limitations.[4 5]

There are two main implications of this study and its findings. The first is that Oncology RCTs measuring HRQoL, especially ones with the particular characteristics showing

higher proportions of high risk of bias, must be designed and conducted carefully to avoid the common attrition and blinding related methodological issues typically found in these types of trials. This is especially true with respect to better design to minimize attrition, and in particular making sure to follow patients after progression, given its relevance to RCTs examining HRQoL. Such steps could include, among others, obtaining information of several individuals not living with patients who are likely to know their whereabouts, checking contact information at the time initially obtained, careful monitoring to ensure measurement takes place at each patient visit, ensuring resources are available for tracking patients who prove hard to follow, and ensuring adequate logistical planning and institutional staff training.[18] These types of steps need to be taken into consideration from the design stage since it is quite common to have missing data for patient reported questionnaires, arising from patients skipping the entire assessment or specific questions within it, and given the typical practice of not assessing HRQoL after disease progression in oncology RCTs,[19] as confirmed by the large proportion of trials we found to be designed in this manner. The second implication is that future confirmatory research is required regarding whether non-industry funded trials, breast / ovarian / lung cancer trials, and trials with functionally better patients actually are associated to particular risk of bias concerns. Finally, we cannot consider the issue of the PFS-HRQoL association definitively closed; higher quality studies are necessary to provide that definitive answer.

4.6 Conclusion

The reported complete descriptive data from our review, and the findings of high proportions of high risk of bias in attrition and blinding domains of oncology RCTs must be addressed when designing, conducting, and reviewing oncology RCTs designed to measure HRQoL.

4.7 Figures and Tables

Figure 1: Overall risk of bias graph

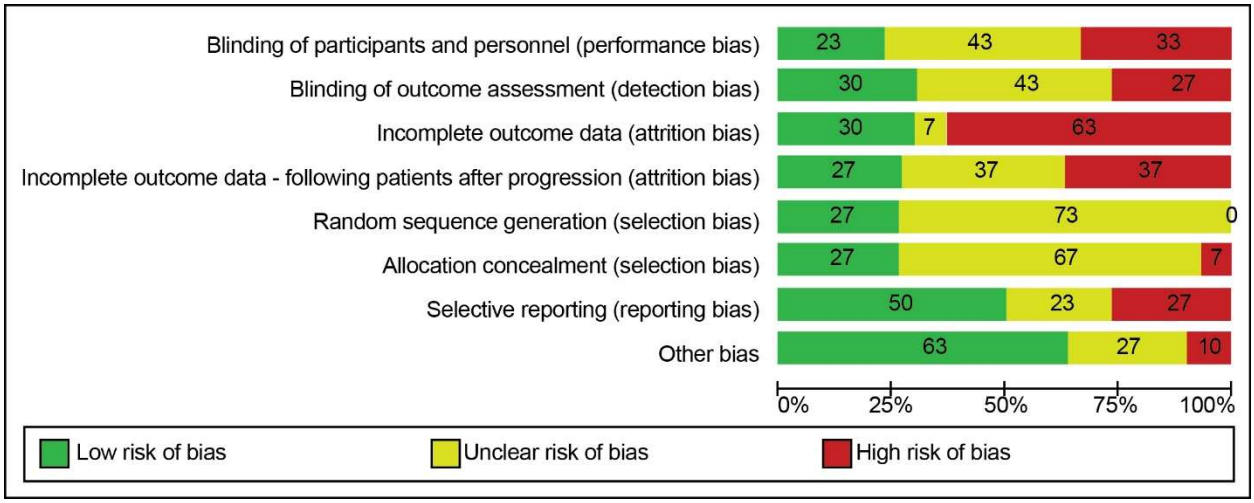
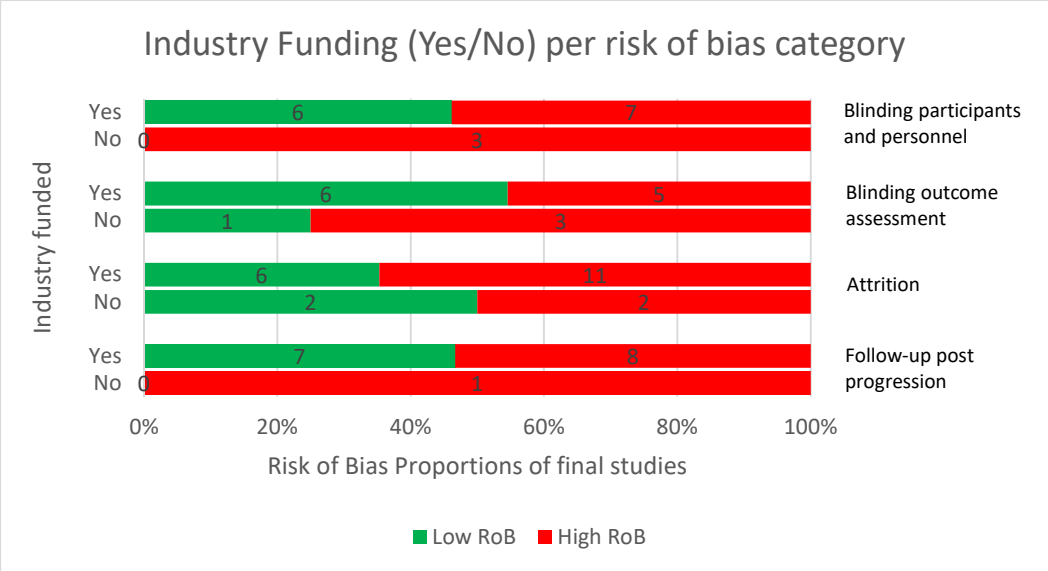
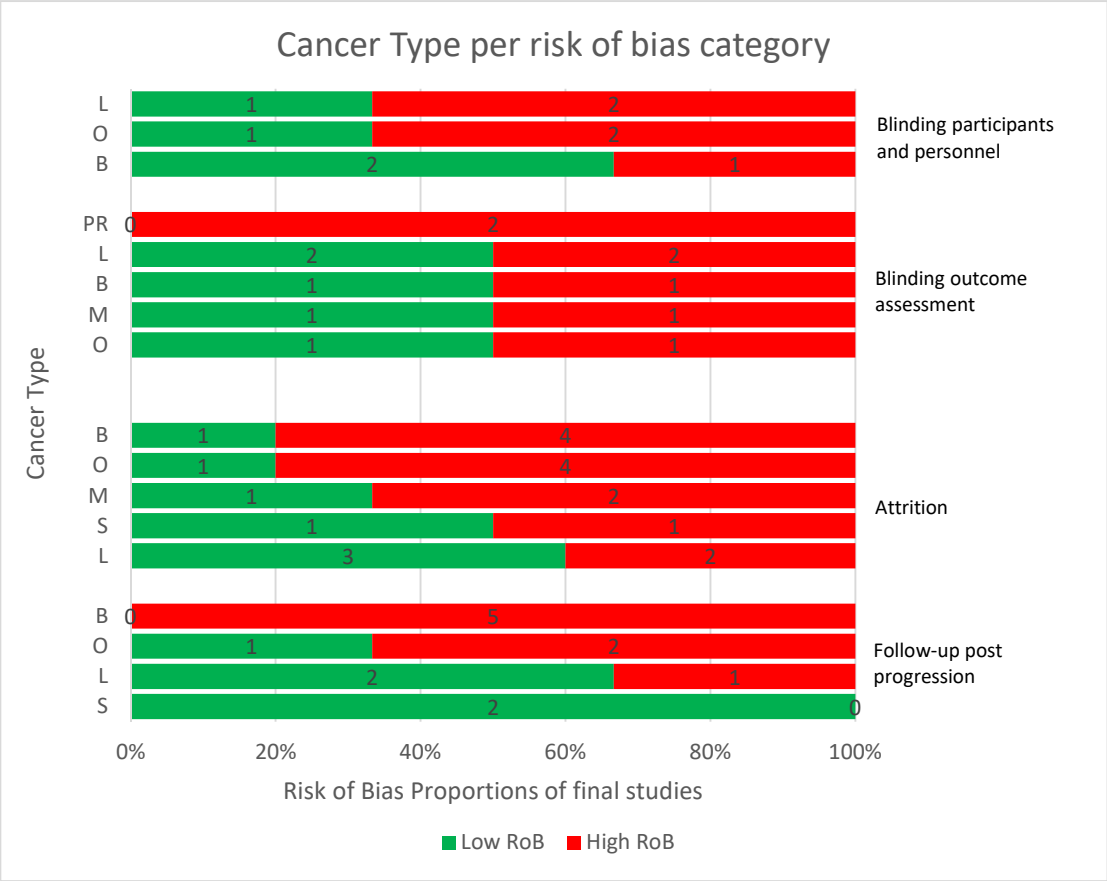


Figure 2: Industry funding risk of bias graph



Note: follow-up post progression bar represented by only one trial was included in this figure to show the complete domain assessment.

Figure 3: Cancer type risk of bias graph



B = Breast, L = Lung, M = Melanoma, O = Ovarian, PR = Prostate, S = Stomach.

Note: cancer types represented by only one trial were not included in this figure.

Figure 4: ECOG status risk of bias graph

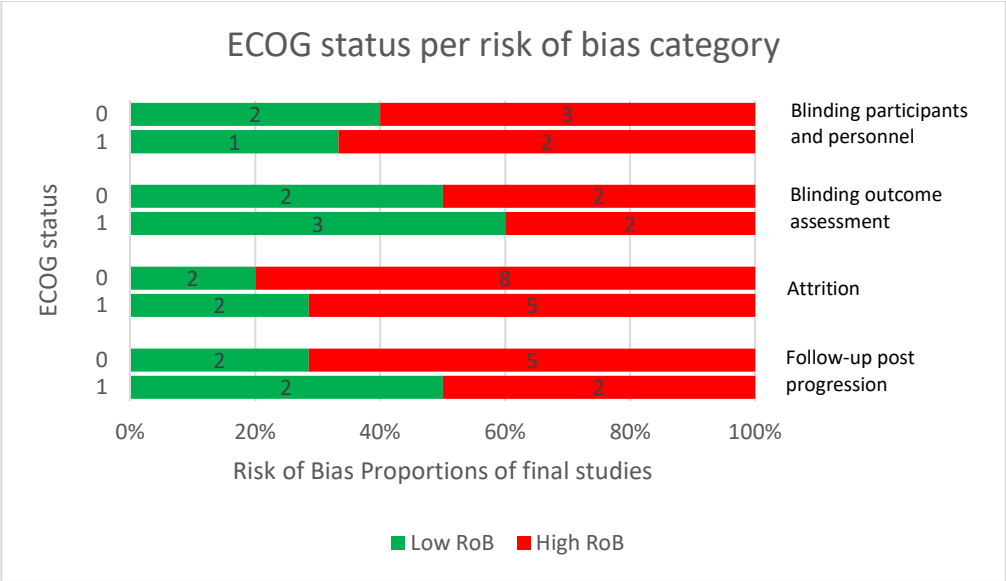


Figure 5: Disease stage risk of bias graph

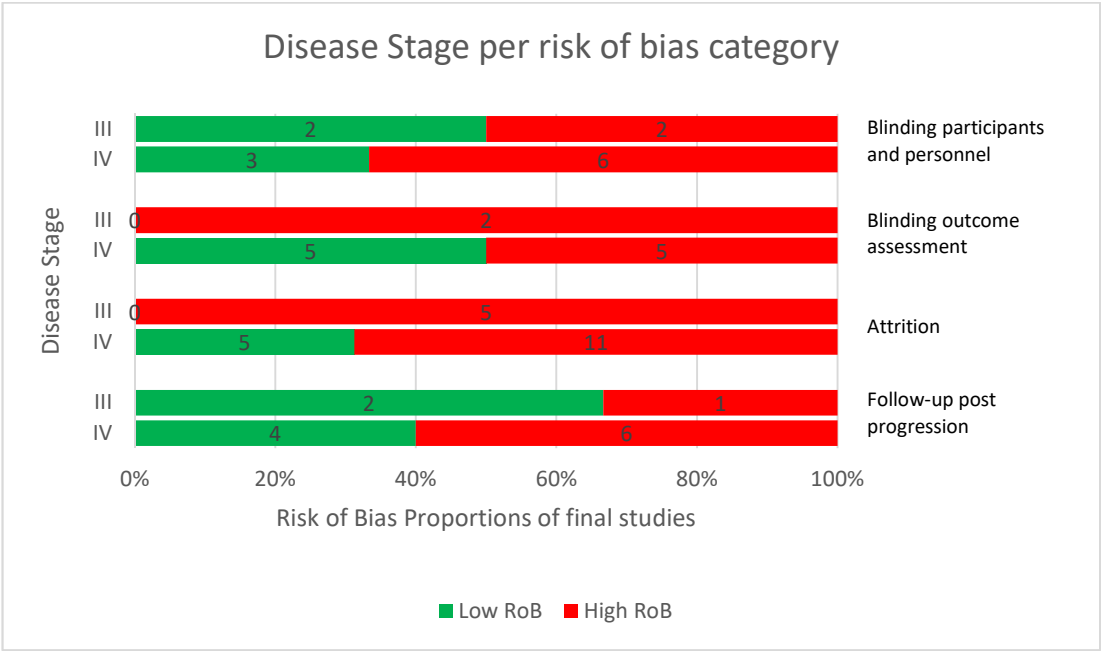


Table 1: Summary of study characteristics

Trial ID	First Author	Industry Funded	Cancer Type	Cancer Stage (Main*)	ECOG (Main*)	Progression Assessment	HRQoL Instruments used
1-B1	Biganzoli L et al., 2002[20] Bottomley A et al., 2004[21]	Yes	breast cancer	Stage IV	1	N/A	EORTC QLQ-C30
2-B2	Bottomley A et al., 2005[22]	No	Locally advanced breast cancer	Stage III	N/A	mix	EORTC QLQ-C30
3-B3	Cameron D et al., 2008[23] Zhou X et al., 2009[24]	Yes	breast cancer	Stage IV	0	mix	FACT-General (FACT-G)
4-B4	Di Leo A et al., 2008[25] Sherrill B et al., 2010[19]	Yes	metastatic Breast Cancer	Stage IV	0	radiographic	FACT-B, FACT-B TOI
5-B5	Nuzzo F et al., 2011[26]	N/A	breast cancer	Stage IV	0	clinical	EORTC QLQ-C30
6-C1	Comella P et al., 2008[27]	N/A	Colorectal Cancer	N/A	0	radiographic	EORTC QLQ-C30
7-L1	Wachters FM et al., 2003[28]	Yes	Non-small cell lung cancer (NSCLC)	Stage IV	1	mix	EORTC QLQ-C30 + EORTC QLQ-LC13
8-L2	Lilenbaum RC et al., 2005[29]	N/A	advanced non-small-cell lung cancer (NSCLC)	Stage IV	1	N/A	Patient reported Lung Cancer Symptom Scale (LCSS)
9-L3	Maruyama R et al., 2008[30] Sekine I et al., 2009[31]	Yes	Metastatic/advanced Non-small cell lung cancer (NSCLC)	Stage IV	N/A	N/A	FACT-L, FACT-L TOI
10-L4	Han BH et al., 2011[32]	N/A	lung cancer	N/A	N/A	mix	Patient reported Lung Cancer Symptom Scale (LCSS)
11-L5	Sun JM et al., 2012[33]	No	Non-small cell lung cancer (NSCLC)	Stage IV	1	radiographic	EORTC QLQ-C30
12-M1	Middleton MR et al., 2000[34] Kiebert GM et al., 2003[35]	N/A	Metastatic malignant melanoma	Stage III	0	N/A	EORTC QLQ-C30
13-M2	Avril MF et al., 2004[36]	No	Melanoma	Stage IV	N/A	radiographic	EORTC QLQ-C30
14-M3	Grob JJ et al., 2014[37]	Yes	BRAF V600E mutation-positive advanced and metastatic melanoma	N/A	N/A	N/A	EORTC QLQ-C30
15-MM1	Palumbo A et al., 2012[38] Dimopoulos MA et al., 2013[39]	Yes	Multiple myeloma	Stage III	N/A	N/A	EORTC QLQ-C30
16-NE1	Arnold R et al., 2005[40]	Yes	Progressive Metastatic Neuroendocrine foregut and midgut tumors	N/A	N/A	radiographic	EORTC QLQ-C30
17-O1	du Bois A et al., 2003[41]	Yes	Ovarian cancer	Stage IIIC	0	mix	EORTC QLQ-C30
18-O2	Pujade-Lauraine E et al., 2010[42] Brundage M et al., 2012[43]	Yes	Recurrent platinum-sensitive ovarian cancer	N/A	0	mix	EORTC QLQ-C30
19-O3	Monk BJ et al., 2010[44] Krasner CN et al., 2012[45]	Yes	epithelial ovarian, fallopian tube, or primary peritoneal carcinoma	Stage IV	0	mix	EORTC QLQ-C30
20-O4	Pokrzywinski R et al., 2011[46]	N/A	ovarian cancer	N/A	0	N/A	FACT-G, FACT-O
21-O5	Burger RA et al., 2011[47] Monk BJ et al., 2013[48]	Yes	epithelial ovarian, primary peritoneal, or fallopian-tube cancer	Stage III	N/A	mix	FACT-G

Trial ID	First Author	Industry Funded	Cancer Type	Cancer Stage (Main*)	ECOG (Main*)	Progression Assessment	HRQoL Instruments used
22-P1	Philip PA et al., 2010[49] Moinpour CM et al., 2010[50]	Yes	adenocarcinoma of the pancreas	Stage IVb	1	N/A	eight-item linear analog self-assessment (LASA) questionnaire
23-PR1	Small EJ et al., 2002[51] Ahles TA et al., 2004[52]	N/A	prostate cancer	N/A	1	N/A	FACT-G
24-PR2	Dawson N et al., 2010[53]	Yes	Hormone-resistant prostate cancer	Stage IV	N/A	mix	FACT-G
25-RC1	Cella D et al., 2012[54]	Yes	Renal cell carcinoma (RCC)	Stage IV	1	radiographic	EORTC QLQ-C30
26-S1	Al-Batran SE et al., 2013[55]	Yes	adenocarcinoma of the stomach or oesophagogastric junction	Stage IV	1	radiographic	EORTC QLQ-C30
27-S2	Bonnetain F et al., 2005[56]	Yes	Gastric adenocarcinoma	Stage IV	N/A	N/A	EORTC QLQ-C30
28-S3	Jiang FS et al., 2009[57]	No	stomach carcinoma	Stage IVb	N/A	mix	Karnofsky score
29-UC1	Long III HJ et al., 2006[58]	N/A	carcinoma of the uterine cervix	Stage IVb	N/A	clinical	FACT-Cervix (FACT-Cx)
30-UC2	Monk BJ et al., 2009[59] Cella D et al., 2010[60]	No	Cervical cancer	Stage IVb	0	N/A	FACT-Cx, FACT-B TOI

N/A = Not Available, B = Breast, C = Colorectal, L = Lung, M = Melanoma, MM = Multiple Myeloma, NE = Neuroendocrine, O = Ovarian, P = Pancreas, PR = Prostate, RC = Renal Cell, S = Stomach, UC = Uterine Cervical, EORTC QLQ-C30 = European Organization for Research and Treatment of Cancer quality of life questionnaire, FACT = Functional Assessment of Cancer Therapy, TOI = Trial Outcome Index

*Main defined as largest group of patients in each trial

4.8 References

1. Peppercorn JM, Smith TJ, Helft PR, et al. American society of clinical oncology statement: toward individualized care for patients with advanced cancer. *J Clin Oncol* 2011;29(6):755-60. doi: 10.1200/JCO.2010.33.1744

2. Booth CM, Eisenhauer EA. Progression-free survival: meaningful or simply measurable? *J Clin Oncol* 2012;30(10):1030-3. doi: 10.1200/JCO.2011.38.7571

3. Pazdur R. Endpoints for assessing drug activity in clinical trials. *Oncologist* 2008;13 Suppl 2:19-21. doi: 10.1634/theoncologist.13-S2-19

4. Bernhard J, Cella DF, Coates AS, et al. Missing quality of life data in cancer clinical trials: serious problems and challenges. *Stat Med* 1998;17(5-7):517-32.

5. Bouca-Machado R, Rosario M, Alarcao J, et al. Clinical trials in palliative care: a systematic review of their methodological characteristics and of the quality of their reporting. *BMC Palliat Care* 2017;16(1):10. doi: 10.1186/s12904-016-0181-9

6. Gutman SI, Piper M, Grant MD, et al. Progression-Free Survival: What Does It Mean for Psychological Well-Being or Quality of Life? Rockville (MD)2013.

7. Kovic B, Guyatt G, Brundage M, et al. Association between progression-free survival and health-related quality of life in oncology: a systematic review protocol. *BMJ Open* 2016;6(9):e012909. doi: 10.1136/bmjopen-2016-012909

8. Booth A, Clarke M, Dooley G, et al. The nuts and bolts of PROSPERO: an international prospective register of systematic reviews. *Syst Rev* 2012;1:2. doi: 10.1186/2046-4053-1-2

9. Booth A, Clarke M, Ghersi D, et al. An international registry of systematic-review protocols. *Lancet* 2011;377(9760):108-9. doi: 10.1016/S0140-6736(10)60903-8

10. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009;339:b2535. doi: 10.1136/bmj.b2535

11. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ* 2009;339:b2700. doi: 10.1136/bmj.b2700

12. Cochrane Handbook for Systematic Reviews of Interventions. In: Higgins JPT GSe, ed.: The Cochrane Collaboration, 2011.

13. Abridged Index Medicus (AIM or "Core Clinical") Journal Titles. 09 July 2003 ed: U.S. National Library of Medicine.

14. Higgins JP, Altman DG, Gotzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011;343:d5928. doi: 10.1136/bmj.d5928

15. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *J Clin Epidemiol* 2009;62(10):1006-12. doi: 10.1016/j.jclinepi.2009.06.005

16. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33(1):159-74.

17. Akl EA, Sun X, Busse JW, et al. Specific instructions for estimating unclearly reported blinding status in randomized trials were reliable and valid. *J Clin Epidemiol* 2012;65(3):262-7. doi: 10.1016/j.jclinepi.2011.04.015

18. Osoba D, Bezjak A, Brundage M, et al. 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. doi: 10.1111/j.1524-4733.2007.00278.x

19. Sherrill B, Di Leo A, Amonkar MM, et al. Quality-of-life and quality-adjusted survival (Q-TWiST) in patients receiving lapatinib in combination with paclitaxel as first-line treatment for metastatic breast cancer. *Curr Med Res Opin* 2010;26(4):767-75. doi: 10.1185/03007991003590860

20. Biganzoli L, Cufer T, Bruning P, et al. Doxorubicin and paclitaxel versus doxorubicin and cyclophosphamide as first-line chemotherapy in metastatic breast cancer: The European Organization for Research and Treatment of Cancer 10961 Multicenter Phase III Trial. *J Clin Oncol* 2002;20(14):3114-21. doi: 10.1200/JCO.2002.11.005

21. Bottomley A, Biganzoli L, Cufer T, et al. Randomized, controlled trial investigating short-term health-related quality of life with doxorubicin and paclitaxel versus doxorubicin and cyclophosphamide as first-line chemotherapy in patients with metastatic breast cancer: European Organization for Research and Treatment of Cancer Breast Cancer Group, Investigational Drug Branch for Breast Cancer and

the New Drug Development Group Study. *J Clin Oncol* 2004;22(13):2576-86. doi: 10.1200/JCO.2004.02.037

22. Bottomley A, Therasse P, Piccart M, et al. Health-related quality of life in survivors of locally advanced breast cancer: an international randomised controlled phase III trial. *Lancet Oncol* 2005;6(5):287-94. doi: 10.1016/S1470-2045(05)70100-5
23. Cameron D, Casey M, Press M, et al. A phase III randomized comparison of lapatinib plus capecitabine versus capecitabine alone in women with advanced breast cancer that has progressed on trastuzumab: updated efficacy and biomarker analyses. *Breast Cancer Res Treat* 2008;112(3):533-43. doi: 10.1007/s10549-007-9885-0
24. Zhou X, Cella D, Cameron D, et al. Lapatinib plus capecitabine versus capecitabine alone for HER2+ (ErbB2+) metastatic breast cancer: quality-of-life assessment. *Breast Cancer Res Treat* 2009;117(3):577-89. doi: 10.1007/s10549-009-0310-8
25. Di Leo A, Gomez HL, Aziz Z, et al. Phase III, double-blind, randomized study comparing lapatinib plus paclitaxel with placebo plus paclitaxel as first-line treatment for metastatic breast cancer. *J Clin Oncol* 2008;26(34):5544-52. doi: 10.1200/JCO.2008.16.2578
26. Nuzzo F, Morabito A, Gravina A, et al. Effects on quality of life of weekly docetaxel-based chemotherapy in patients with locally advanced or metastatic breast cancer: results of a single-centre randomized phase 3 trial. *BMC Cancer* 2011;11:75. doi: 10.1186/1471-2407-11-75

27. Comella P, Massidda B, Filippelli G, et al. Randomised trial comparing biweekly oxaliplatin plus oral capecitabine versus oxaliplatin plus i.v. bolus fluorouracil/leucovorin in metastatic colorectal cancer patients: results of the Southern Italy Cooperative Oncology study 0401. *J Cancer Res Clin Oncol* 2009;135(2):217-26. doi: 10.1007/s00432-008-0454-7

28. Wachters FM, Van Putten JW, Kramer H, et al. First-line gemcitabine with cisplatin or epirubicin in advanced non-small-cell lung cancer: a phase III trial. *Br J Cancer* 2003;89(7):1192-9. doi: 10.1038/sj.bjc.6601283

29. Lilenbaum RC, Chen CS, Chidiac T, et al. Phase II randomized trial of vinorelbine and gemcitabine versus carboplatin and paclitaxel in advanced non-small-cell lung cancer. *Ann Oncol* 2005;16(1):97-101. doi: 10.1093/annonc/mdi009

30. Maruyama R, Nishiwaki Y, Tamura T, et al. Phase III study, V-15-32, of gefitinib versus docetaxel in previously treated Japanese patients with non-small-cell lung cancer. *J Clin Oncol* 2008;26(26):4244-52. doi: 10.1200/JCO.2007.15.0185

31. Sekine I, Ichinose Y, Nishiwaki Y, et al. Quality of life and disease-related symptoms in previously treated Japanese patients with non-small-cell lung cancer: results of a randomized phase III study (V-15-32) of gefitinib versus docetaxel. *Ann Oncol* 2009;20(9):1483-8. doi: 10.1093/annonc/mdp031

32. Han BH, Xiu QY, Wang HM, et al. [A multicenter, randomized, double-blind, placebo-controlled safety study to evaluate the clinical effects and quality of life of paclitaxel-carboplatin (PC) alone or combined with endostar for advanced non-small cell lung cancer (NSCLC)]. *Zhonghua Zhong Liu Za Zhi* 2011;33(11):854-9.

33. Sun JM, Lee KH, Kim SW, et al. Gefitinib versus pemetrexed as second-line treatment in patients with nonsmall cell lung cancer previously treated with platinum-based chemotherapy (KCSG-LU08-01): an open-label, phase 3 trial. *Cancer* 2012;118(24):6234-42. doi: 10.1002/cncr.27630

34. Middleton MR, Grob JJ, Aaronson N, et al. Randomized phase III study of temozolomide versus dacarbazine in the treatment of patients with advanced metastatic malignant melanoma. *J Clin Oncol* 2000;18(1):158-66. doi: 10.1200/JCO.2000.18.1.158

35. Kiebert GM, Jonas DL, Middleton MR. Health-related quality of life in patients with advanced metastatic melanoma: results of a randomized phase III study comparing temozolomide with dacarbazine. *Cancer Invest* 2003;21(6):821-9.

36. Avril MF, Aamdal S, Grob JJ, et al. Fotemustine compared with dacarbazine in patients with disseminated malignant melanoma: a phase III study. *J Clin Oncol* 2004;22(6):1118-25. doi: 10.1200/JCO.2004.04.165

37. Grob JJ, Amonkar MM, Martin-Algarra S, et al. Patient perception of the benefit of a BRAF inhibitor in metastatic melanoma: quality-of-life analyses of the BREAK-3 study comparing dabrafenib with dacarbazine. *Ann Oncol* 2014;25(7):1428-36. doi: 10.1093/annonc/mdu154

38. Palumbo A, Hajek R, Delforge M, et al. Continuous lenalidomide treatment for newly diagnosed multiple myeloma. *N Engl J Med* 2012;366(19):1759-69. doi: 10.1056/NEJMoa1112704

39. Dimopoulos MA, Delforge M, Hajek R, et al. Lenalidomide, melphalan, and prednisone, followed by lenalidomide maintenance, improves health-related quality of life in newly diagnosed multiple myeloma patients aged 65 years or older: results of a randomized phase III trial. *Haematologica* 2013;98(5):784-8. doi: 10.3324/haematol.2012.074534

40. Arnold R, Rinke A, Klose KJ, et al. Octreotide versus octreotide plus interferon-alpha in endocrine gastroenteropancreatic tumors: a randomized trial. *Clin Gastroenterol Hepatol* 2005;3(8):761-71.

41. du Bois A, Luck HJ, Meier W, et al. A randomized clinical trial of cisplatin/paclitaxel versus carboplatin/paclitaxel as first-line treatment of ovarian cancer. *J Natl Cancer Inst* 2003;95(17):1320-9.

42. Pujade-Lauraine E, Wagner U, Aavall-Lundqvist E, et al. Pegylated liposomal Doxorubicin and Carboplatin compared with Paclitaxel and Carboplatin for patients with platinum-sensitive ovarian cancer in late relapse. *J Clin Oncol* 2010;28(20):3323-9. doi: 10.1200/JCO.2009.25.7519

43. Brundage M, Gropp M, Mefti F, et al. Health-related quality of life in recurrent platinum-sensitive ovarian cancer--results from the CALYPSO trial. *Ann Oncol* 2012;23(8):2020-7. doi: 10.1093/annonc/mdr583

44. Monk BJ, Herzog TJ, Kaye SB, et al. Trabectedin plus pegylated liposomal Doxorubicin in recurrent ovarian cancer. *J Clin Oncol* 2010;28(19):3107-14. doi: 10.1200/JCO.2009.25.4037

45. Krasner CN, Poveda A, Herzog TJ, et al. Patient-reported outcomes in relapsed ovarian cancer: results from a randomized Phase III study of trabectedin with pegylated liposomal doxorubicin (PLD) versus PLD alone. *Gynecol Oncol* 2012;127(1):161-7. doi: 10.1016/j.ygyno.2012.06.034

46. Pokrzywinski R, Secord AA, Havrilesky LJ, et al. Health-related quality of life outcomes of docetaxel/carboplatin combination therapy vs. sequential therapy with docetaxel then carboplatin in patients with relapsed, platinum-sensitive ovarian cancer: results from a randomized clinical trial. *Gynecol Oncol* 2011;123(3):505-10. doi: 10.1016/j.ygyno.2011.08.015

47. Burger RA, Brady MF, Bookman MA, et al. Incorporation of bevacizumab in the primary treatment of ovarian cancer. *N Engl J Med* 2011;365(26):2473-83. doi: 10.1056/NEJMoa1104390

48. Monk BJ, Huang HQ, Burger RA, et al. Patient reported outcomes of a randomized, placebo-controlled trial of bevacizumab in the front-line treatment of ovarian cancer: a Gynecologic Oncology Group Study. *Gynecol Oncol* 2013;128(3):573-8. doi: 10.1016/j.ygyno.2012.11.038

49. Philip PA, Benedetti J, Corless CL, et al. Phase III study comparing gemcitabine plus cetuximab versus gemcitabine in patients with advanced pancreatic adenocarcinoma: Southwest Oncology Group-directed intergroup trial S0205. *J Clin Oncol* 2010;28(22):3605-10. doi: 10.1200/JCO.2009.25.7550

50. Moinpour CM, Vaught NL, Goldman B, et al. Pain and emotional well-being outcomes in Southwest Oncology Group-directed intergroup trial S0205: a phase

III study comparing gemcitabine plus cetuximab versus gemcitabine as first-line therapy in patients with advanced pancreas cancer. *J Clin Oncol* 2010;28(22):3611-6. doi: 10.1200/JCO.2009.25.8285

51. Small EJ, Halabi S, Ratain MJ, et al. Randomized study of three different doses of suramin administered with a fixed dosing schedule in patients with advanced prostate cancer: results of intergroup 0159, cancer and leukemia group B 9480. *J Clin Oncol* 2002;20(16):3369-75. doi: 10.1200/JCO.2002.10.022
52. Ahles TA, Herndon JE, 2nd, Small EJ, et al. Quality of life impact of three different doses of suramin in patients with metastatic hormone-refractory prostate carcinoma: results of Intergroup O159/Cancer and Leukemia Group B 9480. *Cancer* 2004;101(10):2202-8. doi: 10.1002/cncr.20655
53. Dawson N, Payne H, Battersby C, et al. Health-related quality of life in pain-free or mildly symptomatic patients with metastatic hormone-resistant prostate cancer following treatment with the specific endothelin A receptor antagonist zibotentan (ZD4054). *J Cancer Res Clin Oncol* 2011;137(1):99-113. doi: 10.1007/s00432-010-0864-1
54. Cella D, Pickard AS, Duh MS, et al. Health-related quality of life in patients with advanced renal cell carcinoma receiving pazopanib or placebo in a randomised phase III trial. *Eur J Cancer* 2012;48(3):311-23. doi: 10.1016/j.ejca.2011.05.017
55. Al-Batran SE, Pauligk C, Homann N, et al. The feasibility of triple-drug chemotherapy combination in older adult patients with oesophagogastric cancer: a

randomised trial of the Arbeitsgemeinschaft Internistische Onkologie (FLOT65+). *Eur J Cancer* 2013;49(4):835-42. doi: 10.1016/j.ejca.2012.09.025

56. Bonnetain F, Bouche O, Conroy T, et al. Longitudinal quality of life study in patients with metastatic gastric cancer. Analysis modalities and clinical applicability of QoL in randomized phase II trial in a digestive oncology. *Gastroenterol Clin Biol* 2005;29(11):1113-24.

57. Jiang FS, Wang G, Sun YB, et al. Clinical observation on recombinant human vascular endostatin combined wih XELOX regimen in the treatment of advanced stomach carcinoma. [Chinese]. *Tumor* 2009;29(8):790-92. doi: http://dx.doi.org/10.3781/j.issn.1000-7431.2009.08.018

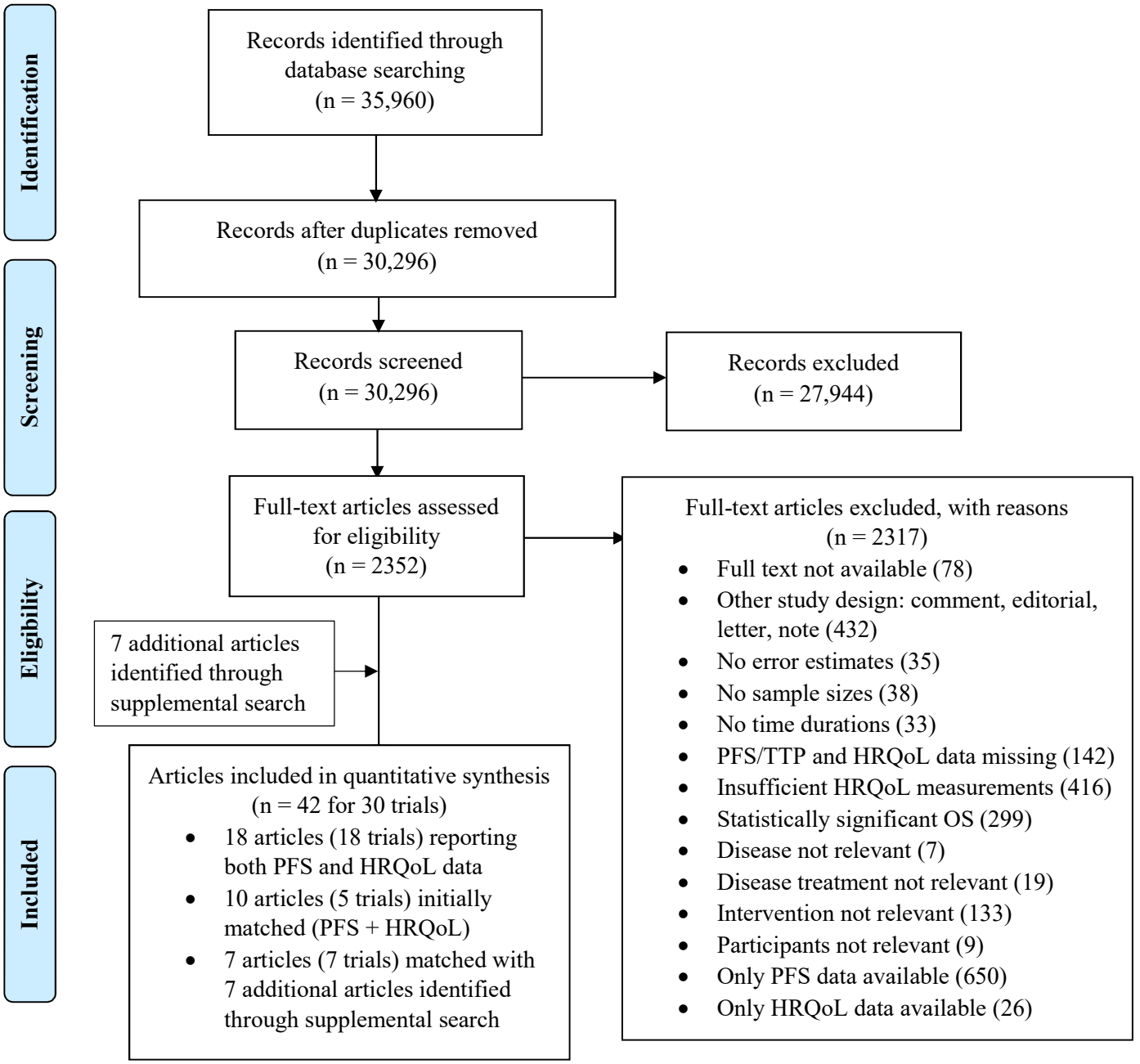
58. Long HJ, 3rd, Monk BJ, Huang HQ, et al. Clinical results and quality of life analysis for the MVAC combination (methotrexate, vinblastine, doxorubicin, and cisplatin) in carcinoma of the uterine cervix: A Gynecologic Oncology Group study. *Gynecol Oncol* 2006;100(3):537-43. doi: 10.1016/j.ygyno.2005.09.023

59. Monk BJ, Sill MW, McMeekin DS, et al. Phase III trial of four cisplatin-containing doublet combinations in stage IVB, recurrent, or persistent cervical carcinoma: a Gynecologic Oncology Group study. *J Clin Oncol* 2009;27(28):4649-55. doi: 10.1200/JCO.2009.21.8909

60. Cella D, Huang HQ, Monk BJ, et al. Health-related quality of life outcomes associated with four cisplatin-based doublet chemotherapy regimens for stage IVB recurrent or persistent cervical cancer: a Gynecologic Oncology Group study. *Gynecol Oncol* 2010;119(3):531-7. doi: 10.1016/j.ygyno.2010.08.020

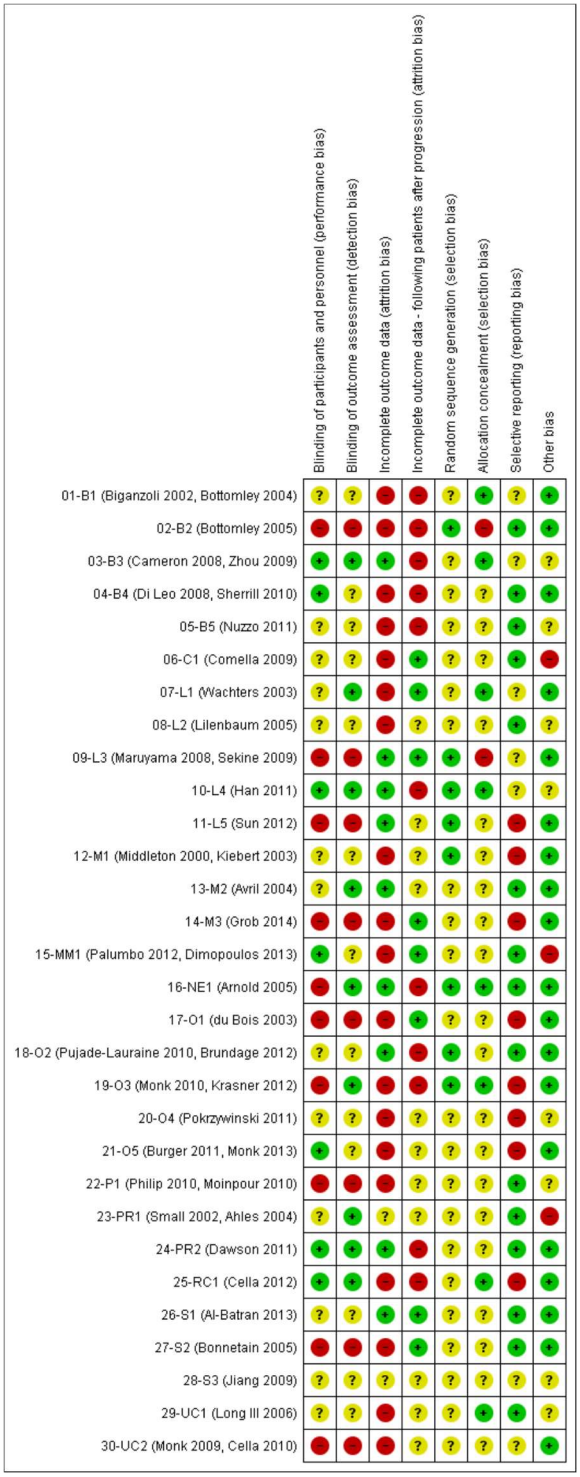
4.9 Appendix 1

PRISMA Flow Diagram



4.10 Appendix 2

Overall risk of bias summary table



CHAPTER 5:

CONCLUSION

Overview

This thesis focused on issues related to the association between PFS and HRQoL in oncology. This chapter summarizes the key findings, limitations, and future directions arising from the work that contributed to this thesis.

Methodological Development

This thesis begins in **Chapter 2**, with a protocol to support the conduct of a systematic review and quantitative analysis. In developing this protocol we worked with an experienced research librarian to develop a highly comprehensive search strategy designed to capture all relevant literature in our review. This preparatory work to maximize the captured relevant evidence was crucially important, since the only other similar previous work had been limited by a lack of evidence from the literature. The protocol also outlined the design and conduct of our systematic review.

Since no previous quantitative analysis between HRQoL and PFS had ever been attempted, we also outlined the development of a new analytical approach to optimally explore the PFS-HRQoL association from published RCTs. Given that HRQoL is measured differently across trials, we used an AUC approach to measure the difference in HRQoL units across treatments for each study. Using regression, these incremental

HRQoL units could then be compared across trials with the incremental PFS from these same trials in order to find an overall association of HRQoL with PFS (i.e. incremental HRQoL versus incremental PFS). Furthermore, we also designed our statistical approach to account for variance in each trial, for both within and between treatment groups, using the properties of expectation and variance. Lastly, in order to more precisely account for the relative contribution of each trial to the overall association by taking trial size into account, we further expanded the statistical methodology by performing a weighted regression. Our new sophisticated analytic approach allowed us to optimally explore the PFS-HRQoL association by allowing for the use of the maximum amount of available published information.

PFS-HRQoL Association

Our main analysis of the PFS-HRQoL association takes place in **Chapter 3**, where we conducted a systematic review of RCTs that reported on both outcomes of PFS and HRQoL. In this study we convened an international team of 20 reviewers to tackle the nearly 40,000 citations from our comprehensive review, and provide reviewer language capabilities to reduce language bias in our review. 30 eligible trials enrolling 10,731 patients across 12 cancer types were used in the quantitative analysis that examined the PFS-HRQoL association across the three physical, global, and emotional HRQoL domains.

Our results failed to find an association between PFS and HRQoL in the absence of OS in oncology across all HRQoL domains. These results, showing that longer PFS was not significantly associated with better HRQoL, raises questions regarding the assumption that interventions prolonging PFS also improve quality of life in patients with cancer. These results are cause for concern in cases where oncology drugs and biologicals are being used by patients based solely on PFS benefit, which is an increasingly common reality. The implications of our review are profound in that trials must either be adequately powered for OS, or designed to ensure rigorous and trustworthy measurement of HRQoL. Another interesting finding of our review was the proportionally low number of RCTs measuring and reporting HRQoL information, as compared to PFS. Given the importance that HRQoL holds for patients and patient-centred cancer care, it is important for oncology RCTs to measure and report HRQoL with more frequency so as to be able to properly evaluate the full impact that a new treatment has on patients.

Though we used rigorous systematic review and statistical analyses methodologies, attempted to control for potential biases, and had a large sample of patients with optimally generalizable results, several limitations exist in this study. First, with only 30 eligible RCTs, we could not perform some planned sensitivity and subgroup analyses and had a lack of statistical power that could have contributed to not finding a PFS-HRQoL association. Second, a large proportion of eligible trials were found to have relatively short follow-up times, which could have contributed to an underestimation of the association results. Nevertheless, our study provides a critically important answer to a

previously unstudied and unresolved question of great importance to cancer patients and cancer care.

Methodological issues in HRQoL oncology RCTs

Using the same database of eligible studies from our review, we further examined relevant methodological issues in **Chapter 4**, such as risk of bias, for oncology RCTs that measure and report on HRQoL. For our descriptive survey we extracted study characteristics and risk of bias data across 8 risk of bias domains, using the Cochrane tool to explore the overall distribution of risk of bias domains. Critical domains of particular concern to the PFS-HRQoL association in oncology, such as those related to attrition and blinding, were further explored for their possible association to four different trial characteristics, through the construction of bar graphs showing frequency of high versus low risk of bias across these characteristics. This study was conducted to evaluate the design and conduct of oncology trials containing data relevant to the PFS-HRQoL association, in order to provide methodological guidance to the future design and conduct of these types of RCTs.

The critical domains related to attrition and blinding had the highest overall proportions of high risk of bias. Attrition bias, however, proved the most common type of high risk of bias across studies, and was of most concern. Among the explored characteristics, trials that were not industry funded and addressed lung cancer were more likely to fail to blind interventions, and trials addressing breast and ovarian cancer, as well as trials with lower

ECOG (i.e. less severe) patients had higher risk of attrition related risk of bias. These results highlight areas for improvement in the design and conduct of oncology RCTs designed to measure HRQoL. This is especially important for attrition related risk of bias, including following patients after progression, and clinicians need to consider these risk of bias issues when deciding on the credibility of these particular types of RCTs. Given the high risk of bias found in our database of eligible studies used to examine the PFS-HRQoL association, higher quality studies are needed to provide a definitive answer on the issue of the PFS-HRQoL association.

Although our study utilized data from a rigorously conducted systematic review, there were a couple of limitations in the study. First, it is possible that any differences or trends observed could be explained by chance, since we could not address this possibility through further statistical analyses due to the low number of eligible trials in our review. Second, trials with unclear assessments were excluded from descriptive review and these may have been informative had inferences on these been made.

Concluding Remarks

This thesis failed to find an association between PFS and HRQoL in the absence of OS in oncology. These results are cause for concern in the oncology field, especially when patients are receiving treatments based solely on PFS benefit. Trials must therefore always either be adequately powered for OS, or designed to ensure rigorous and trustworthy measurement of HRQoL. However, given the lack of oncology RCTs measuring and reporting HRQoL that inform on the association, and the lack of proper design and conduct of available evidence, we cannot consider the issue of the PFS-HRQoL association definitively closed. Additional higher quality future RCTs are needed to confirm our findings.