BMJ Open

COMPARING THE ACCURACY OF QUANTITATIVE VERSUS QUALITATIVE ANALYSES OF INTERIM PET TO PROGNOSTICATE HODGKIN LYMPHOMA: A SYSTEMATIC REVIEW PROTOCOL OF DIAGNOSTIC TEST ACCURACY

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-011729
Article Type:	Protocol
Date Submitted by the Author:	29-Feb-2016
Complete List of Authors:	Procházka, Vít; Faculty of Medicine and Dentistry, Palacky University, Department of Hemato-Oncology Klugar, Miloslav; Palacký University Olomouc Faculty of Medicine and Dentistry, Department of Social Medicine and Public Health Bachanova, Veronika; Oncology and Transplantation University of Minnesota, Division of Hematology Klugarova, Jitka; Faculty of Medicine and Dentistry, Palacký University in Olomouc, Department of Social Medicine and Public Health Tuckova, Dagmar; Faculty of Medicine and Dentistry, Department of Social Medicine and Public Health Papajik, Tomas; Faculty of Medicine and Dentistry, Palacky University, Olomouc, Department of Hemato-Oncology
Primary Subject Heading:	Oncology
Secondary Subject Heading:	Evidence based practice, Diagnostics, Haematology (incl blood transfusion)
Keywords:	Head & neck tumours < ONCOLOGY, Lymphoma < ONCOLOGY, hodgkin lymphoma, PET, TLG, SUV

COMPARING THE ACCURACY OF QUANTITATIVE VERSUS QUALITATIVE ANALYSES OF INTERIM PET TO PROGNOSTICATE HODGKIN LYMPHOMA: A SYSTEMATIC REVIEW PROTOCOL OF DIAGNOSTIC TEST ACCURACY

Authors

1 Vít Procházka, Department of Hemato-Oncology, Faculty of Medicine and Dentistry, Palacky University, Olomouc, I. P. Pavlova 185/6, 77900, Olomouc, Czech Republic

2 Miloslav Klugar, Department of Social Medicine and Public Health, Faculty of Medicine and Dentistry, Palacký University in Olomouc, Hněvotínská 3, 779 00, Olomouc, Czech Republic

The Czech Republic (Middle European) Centre for Evidence-Based Health Care: An affiliated Centre of the Joanna Briggs Institute, Faculty of Medicine and Dentistry, Palacký University in Olomouc

3 Veronika Bachanova, Division of Hematology, Oncology and Transplantation University of Minnesota, Twin Cities, Medicine Hematology Office, MMC 480 Mayo 8480A, 420 Delaware St SE, Minneapolis, MN 55455, United States

4 Jitka Klugarová, Department of Social Medicine and Public Health, Faculty of Medicine and Dentistry, Palacký University in Olomouc, Hněvotínská 3, 779 00, Olomouc, Czech Republic

The Czech Republic (Middle European) Centre for Evidence-Based Health Care: An affiliated Centre of the Joanna Briggs Institute, Faculty of Medicine and Dentistry, Palacký University in Olomouc

5 Dagmar Tučková, Department of Social Medicine and Public Health, Faculty of Medicine and Dentistry, Palacký University in Olomouc, Hněvotínská 3, 779 00, Olomouc, Czech Republic

The Czech Republic (Middle European) Centre for Evidence-Based Health Care: An affiliated Centre of the Joanna Briggs Institute, Faculty of Medicine and Dentistry, Palacký University in Olomouc

6 Tomáš Papajík, Department of Hemato-Oncology, Faculty of Medicine and Dentistry, Palacky University, Olomouc, I. P. Pavlova 185/6, 77900, Olomouc, Czech Republic

CORRESPONDING AUTHOR:

Miloslav Klugar, Ph.D.

Email: miloslav.klugar@upol.cz

Amendments

The current protocol was not an amendment to an existing protocol. To our knowledge, no other protocol intends to systematically review diagnostic test accuracy of quantitative versus qualitative interim PET in prognostication of Hodgkin lymphoma. If necessary, this protocol will be amended in the future with descriptions and rationale provided for any and all alterations.

Support

Grant funding for this manuscript was provided by: Faculty of Medicine and Dentistry, Palacký University in Olomouc, Czech Republic (IGA_LF_2016_001 and RVO: 61989592) and Takeda Pharmaceuticals International AG (IISR-2015-101289).

The final review will disclose all financial and non-financial sources of support and will not be sponsored by private sources (i.e., pharmaceutical companies); financial support will be drawn only from the publicly-funded sources mentioned above.

Key words: Hodgkin lymphoma, PET, qualitative PET, TLG, SUV

Word count: 2422

ABSTRACT

Introduction

Hodgkin lymphoma is an effectively treated malignancy, yet 20% of patients relapse or are refractory to front-line treatments with potentially fatal outcomes. Early detection of poor treatment responders is crucial for appropriate application of tailored treatment strategies. Tumour metabolic imaging of Hodgkin lymphoma using visual (qualitative) 18-fluorodeoxyglucose- positron emission tomography (FDG-PET) is a gold standard for staging and final outcome assessment, but results gathered during the interim period are less accurate.

Analysis of continuous metabolic-morphologic data (quantitative) FDG-PET may enhance the robustness of interim disease monitoring, and help to improve treatment decision-making processes. The objective of this review is to compare diagnostic test accuracy of quantitative versus qualitative interim FDG-PET in the prognostication of Hodgkin lymphoma patients.

Methods

Literature on this topic will be reviewed in a three-step strategy that follows methods described by the Joanna Briggs Institute (JBI). First, MEDLINE and EMBASE databases will be searched. Second, listed databases for published literature (MEDLINE, Tripdatabase, Pedro, EMBASE, the Cochrane Central Register of Controlled Trials, and WoS) and unpublished literature (Open Grey, Current Controlled Trials, MedNar, ClinicalTrials.gov, Cos Conference Papers Index, and International Clinical Trials Registry Platform of the WHO) will be queried. Third, two independent reviewers will analyse titles, abstracts and full texts, and then perform critical appraisal and data extraction from selected studies using the DATARI tool (JBI). If possible, a statistical meta-analysis will be performed on pooled sensitivity and specificity data gathered from the selected studies. Statistical heterogeneity will be assessed. Funnel plots, Begg's rank correlations and Egger's regression tests will be used to detect and/or correct publication bias.

Ethics and dissemination

The results will be disseminated by publishing in a peer-reviewed journal. Ethical assessment will not be needed; only existing sources of literature will be searched.

Systematic review registration number PROSPERO: CRD42016027953

INTRODUCTION

Background

Classical Hodgkin lymphoma (cHL) is the most common lymphoid malignancy affecting patients below the age of 30. Incidence rates of cHL in the US and Central Europe are comparable, with 2.7 new cases per 100,000 men and women per year, and rates trending upward (1,2). Despite high cure rates and effective treatments for cHL, 20% of patients relapse or are refractory to front-line therapies. About 15% of these patients die within five years of diagnosis (3). Overall outcomes are unsatisfactory for patients with relapsed/refractory Hodgkin

lymphoma who proceed to high dose therapies and autologous stem cell transplants (SCT). About 40-50% of SCT recipients relapse and require additional treatments (4). Given our entry into the era of novel "targeted" drugs and immune modulators, identification of poor front-line treatment responders is a growing concern (5).

Implementations of modern imaging methods such as positron emission tomography (PET)/computed tomography (CT) have provided the capability to precisely assess tumour metabolic activity concurrent with an exact measurement of tumour burden. Hodgkin lymphoma has been described as ubiquitously 18-fluorodeoxyglucose (^{18}FDG)-avid. Revised response criteria for malignant lymphoma have therefore included tumour metabolic activity as a key parameter for determining remission status. Historically, complete metabolic responses have been assessed visually using either binary (positive/negative) or semi-quantitative (Deauville) scales (6, 7).

FGG-PET is an inherently quantitative method that generates large amounts of metabolic and morphologic data. Visual binary and semi-quantitative PET analyses do not include quantitative and volumetric parameters [e.g., total metabolic volume (TMV), total lesion glycolysis (TLG) or maximal standardized uptake volume (SUVmax)], and may be observer-biased (8). Recent studies have encouraged quantitative FDG-PET analyses to serve as novel biomarkers for both staging and assessment of early (referred to as interim) and final malignant lymphoma tumour responses to treatments (9, 10). FDG-PET-based tumour metabolic activities at diagnoses were demonstrated to predict survival in both Hodgkin lymphoma and non-Hodgkin lymphoma (NHL) cases. Quantitative metabolic parameters have shown superiority when compared to semi-quantitative assessments in untreated HL and primary diffuse large B-cell lymphoma cases (11-13). Given that personalized medicine has strongly emphasized individualized treatment approaches for all patients, evaluation of chemo-sensitivity is needed during oncology treatment. For cHL, those at risk of treatment failure may be identified by quantitative FDG-PET after a few cycles of therapy (referred to as "interim PET").

Early (interim) visual assessment of cHL tumour metabolism has shown superiority when compared to standard prognostic scoring methods (14). Meta-analysis of these studies showed that interim FDG-PET had high prognostic value for identifying treatment failure (15). Unfortunately, interim PET has not been implemented in routine clinical practice due to the moderate quality of previous evidence and inter-study heterogeneity. One way to circumvent

these barriers is to analyse quantitative FDG-PET results as a method of improving interim PET diagnostic accuracy and reproducibility. Several previous studies have investigated quantitative FDG-PET parameters during the interim period. For example, Rossi and colleagues demonstrated that interim PET after 2 cycles of anthracycline-based chemotherapy captured SUVmax [ΔSUVmax] reductions as large as 71% below baseline. This technique identified positive responders with greater precision than visual assessment, alone (16). Quantitative ΔSUVmax achieved 85% diagnostic accuracy compared to just 76% from the visual method. Furthermore, positive predictive value increased by 24% (from 46% to 70%) when the ΔSUVmax method was used in lieu of visual inspection. Additionally, Tseng and colleagues analysed thirty cHL patients who were scanned at diagnosis and again during treatment. In this study, TMV, SUVmax and TLG were calculated together to determine cumulative changes during treatment regimens. Quantitative interim PET predicted both progression-free and overall survival rates (17).

To our knowledge, a systematic review of the role of quantitative interim PET in cHL patients has yet to be established. We hypothesize that measurements of quantitative tumour characteristics will improve diagnostic and predictive accuracy of interim PET. Thus, more successful candidates will be identified by interim PET for novel treatment approaches. The systematic review protocol described here has an extensive search strategy. It seeks to clarify the role of quantitative interim PET in cHL prognostication and influence practice by informing physician recommendations. Preliminary searches as of January, 2016, were conducted using the MEDLINE, Prospero, JBI Library and Cochrane databases to establish whether previous systematic reviews on this topic were publically available. No systematic reviews or guidelines related to this issue were discovered.

Objective

The objective of this review will be to compare diagnostic test accuracies between quantitative and qualitative interim PET methods with the aim of improving cHL prognostication.

METHODS AND ANALYSIS

Methods

This systematic review protocol was developed according to: 1) the Preferred Reporting Items for Systematic Reviews and Meta-Analysis Protocols (PRISMA-P) (18), and 2) the Joanna

Briggs Institute methodology for systematic reviews of diagnostic test accuracy (19). This protocol has been enrolled with the PROSPERO prospective register of systematic reviews: CRD42016027953.

Study eligibility

Types of participants

The systematic review will consider all studies that investigated adult cHL (determined with WHO diagnostic criteria) (20). Studies that included adolescents (≤ 18 years old) will be excluded.

Index test

The systematic review will consider all studies that measure one or more of the following as an index test: quantitative FDG-PET (QT PET), quantitative evaluation of interim FDG-PET by Metabolic Tumour Volume (MTV), Total Tumour Glycolysis (TLG), or Maximal Standardized Uptake Value (SUVmax).

Reference test

The systematic review will consider studies that perform qualitative FDG-PET (QL PET) or visual evaluation of interim PET as a reference test.

Diagnosis of interest

The systematic review will consider studies that evaluate prognostic accuracy of QT PET in cHL patients as calculated by changes in negative- and positive-predictive values when compared to QL PET.

Types of studies

The systematic review will only include diagnostic cross-sectional study designs.

Search strategy

A search strategy will be developed using medical subject headings (e.g., MeSH for MedLine) and then adopted to query each database. Keywords related to the overarching topic will also be identified. The search strategy seeks to identify and include both published and unpublished work, and will therefore use a three-step search strategy. First, limited searches of MEDLINE

and EMBASE will be undertaken followed by analyses of keywords contained in the title, abstract, and the index terms used to describe an article. Second, all identified keywords and index terms will be searched across all relevant databases. Third, reference lists from the newly identified reports and articles will be searched for additional studies. All studies with title and abstract in English will be considered for inclusion, regardless of the language used in the body of the manuscript. Studies published with no time restriction will also be considered for inclusion.

The databases to be searched include:

MedLine@Ovid, MEDLINE(R), Tripdatabase, Pedro, EMBASE, Cochrane Central Register of Controlled Trials, Cinahl, and Web of Science.

Searches for unpublished studies will be performed using:

Open Grey, Current Controlled Trials, MedNar, ClinicalTrials.gov, Cos Conference Papers Index, and International Clinical Trials Registry Platform of the World Health Organization.

Example search strategy (MedLine@Ovid interface):

1) Hodgkin*

2) Quantitative PET OR Metabolic Tumour Volume OR Total Tumour Glycolysis OR Standardized Uptake Value

3) Qualitative PET OR Visual evaluation PET OR Visual analysis PET

4) Diag* OR sensitivity OR specificity OR predictive

5) 1 AND 2 AND 3 AND 4

Study Records

Literature search results will be compiled and shared by the authorship team using EndNote X7, enabling collaborative study selection. Two reviewers (VP and JK) will independently screen and select studies for possible inclusion in two phases. First, titles and abstracts will be assessed. Second, all relevant full texts will be analysed. Any disagreements will be resolved by discussion and consultation of a third reviewer (MK), as necessary.

Risk of bias in individual studies

Papers selected for retrieval will be assessed by two independent reviewers (VP and DT) for methodological quality prior to inclusion in the systematic review. Assessments will use standardised critical appraisal instruments from the JBI Diagnostic Accuracy Test Assessment and Review Instrument (JBI-DATARI; QUADAS 2; Appendix I) (21). Any disagreements will be resolved by discussion and consultation of a third reviewer (MK), as necessary.

Data collection process

Data will be independently extracted by two reviewers (VP and MK) from studies included in the review using standardised data extraction tools from JBI-DATARI (Appendix II) (21). Extracted data will include: characteristics of the populations, index tests, reference tests, and the diagnoses relevant to the systematic review objectives. Disagreements will be resolved during team discussions, as necessary.

Data items/dealing with missing data

Both generic and trade names of the index tests will be extracted. Diagnostic accuracy of index versus reference tests will be compared using sensitivity, specificity, and receiver operating characteristics (ROC) readouts, as well as patient characteristics (e.g., age, gender, given disease). Study authors will be contacted, as necessary, to provide relevant information for comparative assessments.

Outcomes and prioritisation

The primary outcome of this systematic review will be to compare diagnostic and prognostic accuracy of quantitative and qualitative PET results in cHL patients.

We will seek data answering the following specific questions:

1) What was the rate of five-year progression-free survival (followed from enrolment through the end of the study period)?

2) What is the predicted rate of treatment failure?

Data synthesis

All available diagnostic data will be pooled into a statistical meta-analysis using JBI-DATARI. Results from the included studies will be subjected to double data entry. Meta-analysis results will be presented with two graphical techniques. First, forest plots will illustrate sensitivity and

specificity of each selected primary study by graphing the means and confidence intervals. Means and confidence intervals will also be in numeric form. Additionally, true positive, false positive, true negative, and false negative values will be listed. Second, summary ROC curves will be created. The Bivariate Model for performing meta-analyses will be used.

Assessment of heterogeneity

Initially, clinical heterogeneity will be assessed by determining whether study inclusion criteria are sufficiently similar to the pooled results. If they are clinically homogeneous, statistical heterogeneity will be assessed using standard Chi^2 tests (alpha level: 0.1). If heterogeneity is found, characteristics of the differing studies will be carefully investigated. If it seems that heterogeneity is due to the existence of specific risks of bias in some studies, then the meta-analysis will be restricted to studies that do not contain those risks. To ensure sensitivity analysis, we will exclude all studies that are appraised as having a high risk of bias.

Subgroup analysis

Subgroup analysis will be used for different age and gender characteristics. Another subgroup analysis will be used for cHL and different comorbidities according to their type and severity.

Meta-bias assessment

To show potential reporting bias, we will use funnel plots if more than ten studies are available. Begg's rank correlation and Egger's regression tests will be used for detecting and correcting publication bias.

Confidence in cumulative evidence

Based on the results and quality of evidence, the 'Grading of Recommendation Assessment, Development and Evaluation' (GRADE) tool will be used (22). Quality of evidence will be assessed across the domains of: risk of bias, consistency, directness, precision, and publication bias. Quality will be assessed as: high (further research is very unlikely to alter confidence in the accuracy estimate), moderate (further research will likely impact confidence in the accuracy estimate, and may change the estimate), low (further research is very likely to impact confidence in the accuracy estimate, and will likely change the estimate), or very low (the accuracy estimate is very uncertain).

ETHICS AND DISSEMINATION

This systematic review protocol was crafted in February, 2016. Next, the systematic review development team will begin performing the protocol described herein. Dissemination of results will be targeted at patients and oncology practitioners through publication in a peer-reviewed journal. Ethical assessment is unnecessary as only existing sources of literature will be queried and evaluated.

Acknowledgements

We would like to thank the Charlesworth Group for professional editing of this manuscript. Grant funding for this work was provided by the Faculty of Medicine and Dentistry, Palacký University in Olomouc, Czech Republic (IGA_LF_2016_001 and RVO: 61989592) and by Takeda Pharmaceuticals International AG (IISR-2015-101289).

REFERENCES:

1. Shenoy P, Maggioncalda A, Malik N, *et al.* Incidence Patterns and Outcomes for Hodgkin Lymphoma Patients in the United States. *Adv Hematology*2010;**2011**:725219. doi:10.1155/2011/725219
2. Dusek, Pavlík, Májek, *et al.* Estimating cancer incidence, prevalence, and the number of cancer patients treated with antitumor therapy in 2015 and 2020 - analysis of the Czech National Cancer Registry. *Klinická Onkologie Casopis Ceské Slovenské Onkologické Spolecnosti*2015;**28**:30–43. doi:10.14735/amko201530
3. Canellos G, Rosenberg S, Friedberg J, *et al.*Treatment of Hodgkin Lymphoma: A 50-Year Perspective. *J Clin Oncol* 2014;**32**:163–168. doi:10.1200/JCO.2013.53.1194
4. Josting, Rudolph, Mapara, *et al.* Cologne high-dose sequential chemotherapy in relapsed and refractory Hodgkin lymphoma: results of a large multicenter study of the German Hodgkin Lymphoma Study Group (GHSG). *Ann Oncol* 2005;**16**:116–123. doi:10.1093/annonc/mdi003
5. Moskowitz CH, Matasar MJ, Zelenetz AD, *et al.*Normalization of pre-ASCT, FDG-PET imaging with second-line, non-cross-resistant, chemotherapy programs improves event-

free survival in patients with Hodgkin lymphoma. *Blood* 2012;**119**:1665–70. doi:10.1182/blood-2011-10-388058

6. Barrington S, Mikhaeel G, Kostakoglu L, *et al.* Role of Imaging in the Staging and Response Assessment of Lymphoma: Consensus of the International Conference on Malignant Lymphomas Imaging Working Group. *J Clin Oncol* 2014;**32**:3048–3058. doi:10.1200/JCO.2013.53.5229
7. Meignan M, Barrington S, Itti E, *et al.* Report on the 4th International Workshop on Positron Emission Tomography in Lymphoma held in Menton, France, 3–5 October 2012. *Leukemia Lymphoma* 2014;**55**:31–37. doi:10.3109/10428194.2013.802784
8. Carlier T, Bailly C. State-Of-The-Art and Recent Advances in Quantification for Therapeutic Follow-Up in Oncology Using PET. *Front Medicine* 2015;**2**. doi:10.3389/fmed.2015.00018
9. Meignan M, Sasanelli M, Casasnovas R, *et al.*Metabolic tumour volumes measured at staging in lymphoma: methodological evaluation on phantom experiments and patients. *Eur J Nucl Med Mol I*2014;**41**:1113–1122. doi:10.1007/s00259-014-2705-y
10. Meignan M. VI. FDG-PET as a biomarker in lymphoma: from qualitative to quantitative analysis.*Hematol Oncol* 2015;**33**:38–41. doi:10.1002/hon.2214
11. Kobe C, Kuhnert G, Kahraman D, *et al.* Assessment of Tumor Size Reduction Improves Outcome Prediction of Positron Emission Tomography/Computed Tomography After Chemotherapy in Advanced-Stage Hodgkin Lymphoma. *J Clin Oncol* 2014;**32**:1776–1781. doi:10.1200/JCO.2013.53.2507
12. Kanoun S, Rossi C, Berriolo-Riedinger A, *et al.*Baseline metabolic tumour volume is an independent prognostic factor in Hodgkin lymphoma. *Eur J Nucl Med Mol I* 2014;**41**:1735–1743. doi:10.1007/s00259-014-2783-x
13. Ceriani L, Martelli M, Zinzani PL, *et al.* Utility of baseline 18FDG-PET/CT functional parameters in defining prognosis of primary mediastinal (thymic) large B-cell lymphoma. *Blood* 2015;**126**:950–6. doi:10.1182/blood-2014-12-616474
14. Gallamini A, Hutchings M, Rigacci L, *et al.* Early Interim 2-[18F]Fluoro-2-Deoxy-D-Glucose Positron Emission Tomography Is Prognostically Superior to International

Prognostic Score in Advanced-Stage Hodgkin's Lymphoma: A Report From a Joint Italian-Danish Study. *J Clin Oncol* 2007;**25**:3746–3752. doi:10.1200/JCO.2007.11.6525

15. Adams H, Nievelstein R, Kwee T. Prognostic value of interim FDG-PET in Hodgkin lymphoma: systematic review and meta-analysis. *Brit J Haematol*2015;**170**:356–366.
16. Rossi, Kanoun, Berriolo-Riedinger, *et al.* Interim 18F-FDG PET SUVmax Reduction Is Superior to Visual Analysis in Predicting Outcome Early in Hodgkin Lymphoma Patients. *J Nucl Med*2014;**55**:569–573. doi:10.2967/jnumed.113.130609
17. Tseng D, Rachakonda L, Su Z, *et al.* Interim-treatment quantitative PET parameters predict progression and death among patients with hodgkin's disease. *Radiat Oncol* 2012;**7**:5. doi:10.1186/1748-717X-7-5
18. Moher D, Shamseer L, Clarke M, *et al.* Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev 2015*;**4**:1.
19. Campbell J, Klugar M, Ding S, et al. *The Systematic Review of studies of diagnostic test accuracy*. In: JBI, ed. Joanna Briggs Institute Reviewers' Manual: 2015 edition/supplement. Adelaide, South Australia: The University of Adelaide, 2015:1-46.
20. Swerdlow S, Campo E, Harris N. WHO Classification of Tumours, Volume 2; IARC WHO Classification of Tumours, No 2: Lyon, France, IARC Press, 2008.
21. Campbell JM, Klugar M, Ding S, *et al.* Diagnostic test accuracy: methods for systematic review and meta-analysis. *Int J Evid Based Healthc 2015*;**13**(3):154-62.
22. Munn Z, Porritt K, Aromataris E, et al. Summary of Findings Tables for Joanna Briggs Institute Systematic Reviews. The University of Adelaide, South Australia: The Joanna Briggs Institute 2014.

AUTHOR CONTRIBUTIONS

VP and MK conceptualized and designed the study. All authors contributed to selection criteria development, risk of bias assessment strategy, and data extraction. MK, JK and DT were methodologists. VP, VB and TP were Hodgkin lymphoma content experts. All authors (VP, MK, VB, TP) read, provided feedback, and approved the final manuscript.

COMPETING INTERESTS

None to declare.

Appendix I: Critical appraisal instrument

Appendix I: Critical appraisal checklist

JBI Critical Appraisal Checklist for Diagnostic Test Accuracy Studies

Reviewer_____ Date_____

Author_____ Year_____ Record Number_____

	Yes	No	Unclear	Not applicable
1. Was a consecutive or random sample of patients enrolled?	☐	☐	☐	☐
2. Was a case-control design avoided?	☐	☐	☐	☐
3. Did the study avoid inappropriate exclusions?	☐	☐	☐	☐
4. Were the index test results interpreted without knowledge of the results of the reference standard?	☐	☐	☐	☐
5. If a threshold was used, was it pre-specified?	☐	☐	☐	☐
6. Is the reference standard likely to correctly classify the target condition?	☐	☐	☐	☐
7. Were the reference standard results interpreted without knowledge of the results of the index test?	☐	☐	☐	☐
8. Was there an appropriate interval between the index test and the reference standard?	☐	☐	☐	☐
9. Did all patients receive the same reference standard?	☐	☐	☐	☐
10. Were all patients included in the analysis?	☐	☐	☐	☐

Overall appraisal: Include ☐ Exclude ☐ Seek further info ☐

Appendix II: Data extraction instrument

Author/Date	
Inclusion/exclusion criteria: i.e. presenting symptoms, results from previous tests	Inclusion: Exclusion:
Sample size	
Participant demographics (i.e. age, sex, spectrum of presenting symptoms, comorbidity, current treatments, recruitment centres)	
Study methodology (consecutive or random; retrospective or prospective)	
Period that study was carried out (beginning and end date)	
Index test description (including criteria for positive test)	
Reference test description (including criteria for positive test)	
Geographical location of data collection	
Setting of data collection	
Persons executing and interpreting index tests (numbers, training, and expertise)	
Persons executing and interpreting reference test	
Index/reference time interval (and treatments carried out in between)	
Distribution of severity of disease in those with target condition	
Other diagnoses in those without target condition	
Adverse events from index test	
Adverse events from reference test	

Index test results Threshold=	**Condition positive**	**Condition negative**	**Total**
Index test positive (T+)			
Index test negative (T-)			
Total			

Appendix I: Critical appraisal instrument

Appendix I: Critical appraisal checklist

JBI Critical Appraisal Checklist for Diagnostic Test Accuracy Studies

Reviewer_____Date_____

Author_____Year_____Record Number_____

	Yes	No	Unclear	Not applicable
1. Was a consecutive or random sample of patients enrolled?	☐	☐	☐	☐
2. Was a case-control design avoided?	☐	☐	☐	☐
3. Did the study avoid inappropriate exclusions?	☐	☐	☐	☐
4. Were the index test results interpreted without knowledge of the results of the reference standard?	☐	☐	☐	☐
5. If a threshold was used, was it pre-specified?	☐	☐	☐	☐
6. Is the reference standard likely to correctly classify the target condition?	☐	☐	☐	☐
7. Were the reference standard results interpreted without knowledge of the results of the index test?	☐	☐	☐	☐
8. Was there an appropriate interval between the index test and the reference standard?	☐	☐	☐	☐
9. Did all patients receive the same reference standard?	☐	☐	☐	☐
10. Were all patients included in the analysis?	☐	☐	☐	☐

Overall appraisal: Include ☐ Exclude ☐ Seek further info ☐

Appendix II: Data extraction instrument

Author/Date	
Inclusion/exclusion criteria: i.e. presenting symptoms, results from previous tests	Inclusion: Exclusion:
Sample size	
Participant demographics (i.e. age, sex, spectrum of presenting symptoms, comorbidity, current treatments, recruitment centres)	
Study methodology (consecutive or random; retrospective or prospective)	
Period that study was carried out (beginning and end date)	
Index test description (including criteria for positive test)	
Reference test description (including criteria for positive test)	
Geographical location of data collection	
Setting of data collection	
Persons executing and interpreting index tests (numbers, training, and expertise)	
Persons executing and interpreting reference test	
Index/reference time interval (and treatments carried out in between)	
Distribution of severity of disease in those with target condition	
Other diagnoses in those without target condition	
Adverse events from index test	
Adverse events from reference test	

Index test results Threshold=	**Condition positive**	**Condition negative**	**Total**
Index test positive (T+)			
Index test negative (T-)			
Total			

PRISMA-P (Preferred Reporting Items for Systematic review and Meta-Analysis Protocols) 2015 checklist: recommended items to address in a systematic review protocol*

Section and topic	Item No	Checklist item
ADMINISTRATIVE INFORMATION		
Title:		
Identification	1a	Identify the report as a protocol of a systematic review Identified in tittle: **The accuracy of Quantitative interim PET compared to Qualitative interim PET in prognosis of Hodgkin lymphoma: a systematic review protocol of diagnostic test accuracy**
Update	1b	If the protocol is for an update of a previous systematic review, identify as such – N/A
Registration	2	If registered, provide the name of the registry (such as PROSPERO) and registration number - **Systematic review registration number** PROSPERO; CRD42016027953 First paragraph in Methods section
Authors:		
Contact	3a	Provide name, institutional affiliation, e-mail address of all protocol authors; provide physical mailing address of corresponding author – First page – all institutional affiliation, e-mail address just to corresponding author, as it is usual for BMJ in other protocols
Contributions	3b	Describe contributions of protocol authors and identify the guarantor of the review – Contributions identified on the first page: **Contributions** Vít Procházka and Miloslav Klugar were responsible for the study conception and design. All authors contributed to the development of the selection criteria, the risk of bias assessment strategy, and data extraction. Miloslav Klugar, Jitka Klugarová and Dagmar Tučková are methodologists. Vít Procházka, Veronika Bachanova and Tomáš Papajík are the content experts for Hodgkin lymphoma. All authors (Vít Procházka, Miloslav Klugar, Veronika Bachanova, Tomáš Papajík) read, provided feedback and approved the final manuscript.
Amendments	4	If the protocol represents an amendment of a previously completed or published protocol, identify as such and list changes; otherwise, state plan for documenting important protocol amendments – Identified on the first page: **Amendments** This protocol is not an amendment to another existing protocol for the systematic review of diagnostic test accuracy for the accuracy of Quantitative interim PET compared to Qualitative interim PET in prognosis of Hodgkin lymphoma. If necessary, this protocol will be accompanied in the future by amendments indicating a description of the change(s) made and the rationale for making it (them).
Support:		

Sources	5a	Indicate sources of financial or other support for the review – Indicated on the second page: **Support** This paper was supported by grants provided by the Faculty of Medicine and Dentistry, Palacký University in Olomouc, Czech Republic (IGA_LF_2016_001 and RVO: 61989592) and by Takeda Pharmaceuticals International AG (IISR-2015-101289). This review will disclose all financial and non-financial sources of support. Further this review will not be sponsored by any pharmaceutical companies and any financial sources will be drawn only from the above mentioned grant.
Sponsor	5b	Provide name for the review funder and/or sponsor - N/A
Role of sponsor or funder	5c	Describe roles of funder(s), sponsor(s), and/or institution(s), if any, in developing the protocol - N/A
INTRODUCTION		
Rationale	6	Describe the rationale for the review in the context of what is already known – Described in whole background part, mainly in this paragraph: A systematic review of the role of the qualitative interim PET in the HL patients has not been done yet. We presume, that analysis of qualitative tumor parameters will improve diagnostic (predictive) accuracy of interim FDG-PET and will help better identify a candidates for novel treatment approaches. This systematic review with its extensive search strategy may clarify this issue and influence practice by informing recommendations aimed at physicians and patients with Hodgkin lymphoma.
Objectives	7	Provide an explicit statement of the question(s) the review will address with reference to participants, interventions, comparators, and outcomes (PICO) – Provided on the 6th page: **Objectives** The objective of this review is to determine by comparison of Quantitative interim FDG-PET parameters with Qualitative interim FDG-PET parameters diagnostic test accuracy in Hodgkin lymphoma patients prognosis.
METHODS		
Eligibility criteria	8	Specify the study characteristics (such as PICO, study design, setting, time frame) and report characteristics (such as years considered, language, publication status) to be used as criteria for eligibility for the review – Specified on the 6/7th pages Study eligibility Types of participants This review will consider studies that include the adult Hodgkin lymphoma (as diagnosed using WHO diagnostic criteria). Excluded will be studies that include adolescents (under 18 years of age). Index test This review will consider studies that include as index test Quantitative PET (QT PET) Quantitative evaluation of interim PET by Metabolic Tumor Volume (MTV), Total Tumor Glycolysis (TLG) and Standardized Uptake Value (SUV). Reference test This review will consider studies that include as reference test qualitative PET (QL PET) visual evaluation of interim PET

		Diagnosis of interest This review will consider studies that evaluate accuracy of prognosis for Hodgkin Lymphoma patients. Change in negative-predictive value and positive-predictive value compared to QL-PET. Types of studies This review will consider only diagnostic cross-sectional study design for inclusion.
Information sources	9	Describe all intended information sources (such as electronic databases, contact with study authors, trial registers or other grey literature sources) with planned dates of coverage – Described on the page 7 in Search strategy
Search strategy	10	Present draft of search strategy to be used for at least one electronic database, including planned limits, such that it could be repeated – Described and example given in the part Search strategy, page 8
Study records:		
Data management	11a	Describe the mechanism(s) that will be used to manage records and data throughout the review – Described on the page 8 in part Study records
Selection process	11b	State the process that will be used for selecting studies (such as two independent reviewers) through each phase of the review (that is, screening, eligibility and inclusion in meta-analysis) – Described on the page 8 in parts Study records and Risk of bias in individual studies
Data collection process	11c	Describe planned method of extracting data from reports (such as piloting forms, done independently, in duplicate), any processes for obtaining and confirming data from investigators – Described on the page 8 in part Data collection process
Data items	12	List and define all variables for which data will be sought (such as PICO items, funding sources), any pre-planned data assumptions and simplifications – Described on the page 9 in part Data items/dealing with missing data
Outcomes and prioritization	13	List and define all outcomes for which data will be sought, including prioritization of main and additional outcomes, with rationale – Described on the page 9 in part Outcomes and prioritization
Risk of bias in individual studies	14	Describe anticipated methods for assessing risk of bias of individual studies, including whether this will be done at the outcome or study level, or both; state how this information will be used in data synthesis – Described on the page 8 in part Risk of bias in individual studies
Data synthesis	15a	Describe criteria under which study data will be quantitatively synthesised – Described on the page 9 in part Data synthesis
	15b	If data are appropriate for quantitative synthesis, describe planned summary measures, methods of handling data and methods of combining data from studies, including any planned exploration of consistency (such as I^2, Kendall's τ) – Described on the page 10 in parts Data synthesis and Assessment of heterogeneity
	15c	Describe any proposed additional analyses (such as sensitivity or subgroup analyses, meta-regression) – Described on the page 9 in part Data synthesis
	15d	If quantitative synthesis is not appropriate, describe the type of summary planned – Described on the page 9 in part Data synthesis
Meta-bias(es)	16	Specify any planned assessment of meta-bias(es) (such as publication bias across studies, selective reporting within studies) – Described on the page 10 in part Meta-bias assessment

Confidence in cumulative evidence	17	Describe how the strength of the body of evidence will be assessed (such as GRADE) – Described on the page 10 in part Confidence in cumulative evidence

*** It is strongly recommended that this checklist be read in conjunction with the PRISMA-P Explanation and Elaboration (cite when available) for important clarification on the items. Amendments to a review protocol should be tracked and dated. The copyright for PRISMA-P (including checklist) is held by the PRISMA-P Group and is distributed under a Creative Commons Attribution Licence 4.0.**

From: Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart L, PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ. 2015 Jan 2;349(jan02 1):g7647.

BMJ Open

COMPARING THE ACCURACY OF QUANTITATIVE VERSUS QUALITATIVE ANALYSES OF INTERIM PET TO PROGNOSTICATE HODGKIN LYMPHOMA: A SYSTEMATIC REVIEW PROTOCOL OF DIAGNOSTIC TEST ACCURACY

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-011729.R1
Article Type:	Protocol
Date Submitted by the Author:	27-Jun-2016
Complete List of Authors:	Procházka, Vít; Faculty of Medicine and Dentistry, Palacky University, Department of Hemato-Oncology Klugar, Miloslav; Palacký University Olomouc Faculty of Medicine and Dentistry, Department of Social Medicine and Public Health Bachanova, Veronika; Oncology and Transplantation University of Minnesota, Division of Hematology Klugarova, Jitka; Faculty of Medicine and Dentistry, Palacký University in Olomouc, Department of Social Medicine and Public Health Tuckova, Dagmar; Faculty of Medicine and Dentistry, Department of Social Medicine and Public Health Papajik, Tomas; Faculty of Medicine and Dentistry, Palacky University, Olomouc, Department of Hemato-Oncology
Primary Subject Heading:	Oncology
Secondary Subject Heading:	Evidence based practice, Diagnostics, Haematology (incl blood transfusion)
Keywords:	Head & neck tumours < ONCOLOGY, Lymphoma < ONCOLOGY, hodgkin lymphoma, PET, TLG, SUV

COMPARING THE ACCURACY OF QUANTITATIVE VERSUS QUALITATIVE ANALYSES OF INTERIM PET TO PROGNOSTICATE HODGKIN LYMPHOMA: A SYSTEMATIC REVIEW PROTOCOL OF DIAGNOSTIC TEST ACCURACY

Authors

1 Vít Procházka, Department of Hemato-Oncology, Faculty of Medicine and Dentistry, Palacky University, Olomouc, I. P. Pavlova 185/6, 77900, Olomouc, Czech Republic

2 Miloslav Klugar, Department of Social Medicine and Public Health, Faculty of Medicine and Dentistry, Palacký University in Olomouc, Hněvotínská 3, 779 00, Olomouc, Czech Republic

The Czech Republic (Middle European) Centre for Evidence-Based Health Care: An affiliated Centre of the Joanna Briggs Institute, Faculty of Medicine and Dentistry, Palacký University in Olomouc

3 Veronika Bachanova, Division of Hematology, Oncology and Transplantation University of Minnesota, Twin Cities, Medicine Hematology Office, MMC 480 Mayo 8480A, 420 Delaware St SE, Minneapolis, MN 55455, United States

4 Jitka Klugarová, Department of Social Medicine and Public Health, Faculty of Medicine and Dentistry, Palacký University in Olomouc, Hněvotínská 3, 779 00, Olomouc, Czech Republic

The Czech Republic (Middle European) Centre for Evidence-Based Health Care: An affiliated Centre of the Joanna Briggs Institute, Faculty of Medicine and Dentistry, Palacký University in Olomouc

5 Dagmar Tučková, Department of Social Medicine and Public Health, Faculty of Medicine and Dentistry, Palacký University in Olomouc, Hněvotínská 3, 779 00, Olomouc, Czech Republic

The Czech Republic (Middle European) Centre for Evidence-Based Health Care: An affiliated Centre of the Joanna Briggs Institute, Faculty of Medicine and Dentistry, Palacký University in Olomouc

6 Tomáš Papajík, Department of Hemato-Oncology, Faculty of Medicine and Dentistry, Palacky University, Olomouc, I. P. Pavlova 185/6, 77900, Olomouc, Czech Republic

CORRESPONDING AUTHOR:

Miloslav Klugar, Ph.D.

Email: miloslav.klugar@upol.cz

Amendments

The current protocol was not an amendment to an existing protocol. To our knowledge, no other protocol intends to systematically review diagnostic test accuracy of quantitative versus qualitative interim PET in prognostication of Hodgkin lymphoma. If necessary, this protocol will be amended in the future with descriptions and rationale provided for any and all alterations.

Support

Grant funding for this manuscript was provided by: Faculty of Medicine and Dentistry, Palacký University in Olomouc, Czech Republic (IGA_LF_2016_001 and RVO: 61989592) and Takeda Pharmaceuticals International AG (IISR-2015-101289).

The final review will disclose all financial and non-financial sources of support and will not be sponsored by private sources (i.e., pharmaceutical companies); financial support will be drawn only from the publicly-funded sources mentioned above.

Key words: Hodgkin lymphoma, PET, qualitative PET, TLG, SUV

Word count: 2424

ABSTRACT

Introduction

Hodgkin lymphoma is an effectively treated malignancy, yet 20% of patients relapse or are refractory to front-line treatments with potentially fatal outcomes. Early detection of poor treatment responders is crucial for appropriate application of tailored treatment strategies. Tumour metabolic imaging of Hodgkin lymphoma using visual (qualitative) 18-fluorodeoxyglucose- positron emission tomography (FDG-PET) is a gold standard for staging and final outcome assessment, but results gathered during the interim period are less accurate.

Analysis of continuous metabolic-morphologic data (quantitative) FDG-PET may enhance the robustness of interim disease monitoring, and help to improve treatment decision-making processes. The objective of this review is to compare diagnostic test accuracy of quantitative versus qualitative interim FDG-PET in the prognostication of Hodgkin lymphoma patients.

Methods

Literature on this topic will be reviewed in a three-step strategy that follows methods described by the Joanna Briggs Institute (JBI). First, MEDLINE and EMBASE databases will be searched. Second, listed databases for published literature (MEDLINE, Tripdatabase, Pedro, EMBASE, the Cochrane Central Register of Controlled Trials, and WoS) and unpublished literature (Open Grey, Current Controlled Trials, MedNar, ClinicalTrials.gov, Cos Conference Papers Index, and International Clinical Trials Registry Platform of the WHO) will be queried. Third, two independent reviewers will analyse titles, abstracts and full texts, and then perform critical appraisal and data extraction from selected studies using the DATARI tool (JBI). If possible, a statistical meta-analysis will be performed on pooled sensitivity and specificity data gathered from the selected studies. Statistical heterogeneity will be assessed. Funnel plots, Begg's rank correlations and Egger's regression tests will be used to detect and/or correct publication bias.

Ethics and dissemination

The results will be disseminated by publishing in a peer-reviewed journal. Ethical assessment will not be needed; only existing sources of literature will be searched.

Systematic review registration number PROSPERO: CRD42016027953

INTRODUCTION

Background

Classical Hodgkin lymphoma (cHL) is the most common lymphoid malignancy affecting patients below the age of 30. Incidence rates of cHL in the US and Central Europe are comparable, with 2.7 new cases per 100,000 men and women per year, and rates trending upward (1,2). Despite high cure rates and effective treatments for cHL, 20% of patients relapse or are refractory to front-line therapies. About 15% of these patients die within five years of diagnosis (3). Overall outcomes are unsatisfactory for patients with relapsed/refractory Hodgkin

lymphoma who proceed to high dose therapies and autologous stem cell transplants (SCT). About 40-50% of SCT recipients relapse and require additional treatments (4). Given our entry into the era of novel "targeted" drugs and immune modulators, identification of poor front-line treatment responders is a growing concern (5).

Implementations of modern imaging methods such as positron emission tomography (PET)/computed tomography (CT) have provided the capability to precisely assess tumour metabolic activity concurrent with an exact measurement of tumour burden. Hodgkin lymphoma has been described as ubiquitously 18-fluorodeoxyglucose (^{18}FDG)-avid. Revised response criteria for malignant lymphoma have therefore included tumour metabolic activity as a key parameter for determining remission status. Historically, complete metabolic responses have been assessed visually using either binary (positive/negative) or semi-quantitative (Deauville) scales (6, 7).

FGG-PET is an inherently quantitative method that generates large amounts of metabolic and morphologic data. Visual binary and semi-quantitative PET analyses do not include quantitative and volumetric parameters [e.g., total metabolic volume (TMV), total lesion glycolysis (TLG) or maximal standardized uptake volume (SUVmax)], and may be observer-biased (8). Recent studies have encouraged quantitative FDG-PET analyses to serve as novel biomarkers for both staging and assessment of early (referred to as interim) and final malignant lymphoma tumour responses to treatments (9, 10). FDG-PET-based tumour metabolic activities at diagnoses were demonstrated to predict survival in both Hodgkin lymphoma and non-Hodgkin lymphoma (NHL) cases. Quantitative metabolic parameters have shown superiority when compared to semi-quantitative assessments in untreated HL and primary diffuse large B-cell lymphoma cases (11-13). Given that personalized medicine has strongly emphasized individualized treatment approaches for all patients, evaluation of chemo-sensitivity is needed during oncology treatment. For cHL, those at risk of treatment failure may be identified by quantitative FDG-PET after a few cycles of therapy (referred to as "interim PET").

Early (interim) visual assessment of cHL tumour metabolism has shown superiority when compared to standard prognostic scoring methods (14). Meta-analysis of these studies showed that interim FDG-PET had high prognostic value for identifying treatment failure. Unfortunately, interim PET has not been implemented in routine clinical practice due to the moderate quality of previous evidence and inter-study heterogeneity (15). Moreover, interim FDG-PET could not be

used as a tool for tailored therapy as shown by results of two systematic reviews published by Sickinger and colleagues (16, 17) One way to circumvent these barriers is to analyse quantitative FDG-PET results as a method of improving interim PET diagnostic accuracy and reproducibility. Several previous studies have investigated quantitative FDG-PET parameters during the interim period. For example, Rossi and colleagues demonstrated that interim PET after 2 cycles of anthracycline-based chemotherapy captured SUVmax [ΔSUVmax] reductions as large as 71% below baseline. This technique identified positive responders with greater precision than visual assessment, alone (18). Quantitative ΔSUVmax achieved 85% diagnostic accuracy compared to just 76% from the visual method. Furthermore, positive predictive value increased by 24% (from 46% to 70%) when the ΔSUVmax method was used in lieu of visual inspection. Additionally, Tseng and colleagues analysed thirty cHL patients who were scanned at diagnosis and again during treatment. In this study, TMV, SUVmax and TLG were calculated together to determine cumulative changes during treatment regimens. Quantitative interim PET predicted both progression-free and overall survival rates (19).

To our knowledge, a systematic review of the role of quantitative interim PET in cHL patients has yet to be established. We hypothesize that measurements of quantitative tumour characteristics will improve diagnostic and predictive accuracy of interim PET. Thus, more successful candidates will be identified by interim PET for novel treatment approaches. The systematic review protocol described here has an extensive search strategy. It seeks to clarify the role of quantitative interim PET in cHL prognostication and influence practice by informing physician recommendations. Preliminary searches as of January, 2016, were conducted using the MEDLINE, Prospero, JBI Library and Cochrane databases to establish whether previous systematic reviews on this topic were publically available. No systematic reviews or guidelines related to this issue were discovered.

Objective

The objective of this review will be to compare diagnostic test accuracies between quantitative and qualitative interim PET methods with the aim of improving cHL prognostication.

METHODS AND ANALYSIS

Methods

This systematic review protocol was developed according to: 1) the Preferred Reporting Items for Systematic Reviews and Meta-Analysis Protocols (PRISMA-P) (20), and 2) the Joanna Briggs Institute methodology for systematic reviews of diagnostic test accuracy (21). This protocol has been enrolled with the PROSPERO prospective register of systematic reviews: CRD42016027953.

Study eligibility

Types of participants

The systematic review will consider all studies that investigated adult cHL (determined with WHO diagnostic criteria) (22), who were treated according to the current international guidelines (23, 24). Studies that included adolescents (≤ 18 years old) will be excluded.

Index test

The systematic review will consider all studies that measure one or more of the following as an index test: quantitative FDG-PET (QT PET), quantitative evaluation of interim FDG-PET by Metabolic Tumour Volume (MTV), Total Tumour Glycolysis (TLG), or Maximal Standardized Uptake Value (SUVmax). Only studies which used standardized international criteria for interim FDG-PET interpretation will be analysed (6,7).

Reference test

The systematic review will consider studies that perform qualitative FDG-PET (QL PET) or visual evaluation of interim PET as a reference test.

Diagnosis of interest

The systematic review will consider studies that evaluate prognostic accuracy of QT PET in cHL patients as calculated by changes in negative- and positive-predictive values when compared to QL PET.

Types of studies

The systematic review will only include diagnostic cross-sectional study designs.

Search strategy

A search strategy will be developed using medical subject headings (e.g., MeSH for MedLine)

and then adopted to query each database. Keywords related to the overarching topic will also be identified. The search strategy seeks to identify and include both published and unpublished work, and will therefore use a three-step search strategy. First, limited searches of MEDLINE and EMBASE will be undertaken followed by analyses of keywords contained in the title, abstract, and the index terms used to describe an article. Second, all identified keywords and index terms will be searched across all relevant databases. Third, reference lists from the newly identified reports and articles will be searched for additional studies. All studies with title and abstract in English will be considered for inclusion, regardless of the language used in the body of the manuscript. Studies published with no time restriction will also be considered for inclusion.

The databases to be searched include:

MedLine@Ovid, MEDLINE(R), Tripdatabase, Pedro, EMBASE, Cochrane Central Register of Controlled Trials, Cinahl, and Web of Science.

Searches for unpublished studies will be performed using:

Open Grey, Current Controlled Trials, MedNar, ClinicalTrials.gov, Cos Conference Papers Index, and International Clinical Trials Registry Platform of the World Health Organization.

Example search strategy (MedLine@Ovid interface):

1) Hodgkin*

2) Quantitative PET OR Metabolic Tumour Volume OR Total Tumour Glycolysis OR Standardized Uptake Value

3) Qualitative PET OR Visual evaluation PET OR Visual analysis PET

4) Diag* OR sensitivity OR specificity OR predictive

5) 1 AND 2 AND 3 AND 4

Study Records

Literature search results will be compiled and shared by the authorship team using EndNote X7, enabling collaborative study selection. Two reviewers (VP and JK) will independently screen and select studies for possible inclusion in two phases. First, titles and abstracts will be assessed. Second, all relevant full texts will be analysed. Any disagreements will be resolved by discussion

and consultation of a third reviewer (MK), as necessary.

Risk of bias in individual studies

Papers selected for retrieval will be assessed by two independent reviewers (VP and DT) for methodological quality prior to inclusion in the systematic review. Assessments will use standardised critical appraisal instruments from the JBI Diagnostic Accuracy Test Assessment and Review Instrument (JBI-DATARI; QUADAS 2; Appendix I) (25). Any disagreements will be resolved by discussion and consultation of a third reviewer (MK), as necessary.

Data collection process

Data will be independently extracted by two reviewers (VP and MK) from studies included in the review using standardised data extraction tools from JBI-DATARI (Appendix II) (25). Extracted data will include: characteristics of the populations, index tests, reference tests, and the diagnoses relevant to the systematic review objectives. Disagreements will be resolved during team discussions, as necessary.

Data items/dealing with missing data

Both generic and trade names of the index tests will be extracted. Diagnostic accuracy of index versus reference tests will be compared using sensitivity, specificity, and receiver operating characteristics (ROC) readouts, as well as patient characteristics (e.g., age, gender, given disease). Study authors will be contacted, as necessary, to provide relevant information for comparative assessments.

Outcomes and prioritisation

The primary outcome of this systematic review will be to compare diagnostic and prognostic accuracy of quantitative and qualitative PET results in cHL patients.

We will seek data answering the following specific questions:

1) What was the rate of five-year progression-free survival (followed from enrolment through the end of the study period)?

2) What is the predicted rate of treatment failure?

Data synthesis

All available diagnostic data will be pooled into a statistical meta-analysis using JBI-DATARI. Results from the included studies will be subjected to double data entry. Meta-analysis results will be presented with two graphical techniques. First, forest plots will illustrate sensitivity and specificity of each selected primary study by graphing the means and confidence intervals. Means and confidence intervals will also be in numeric form. Additionally, true positive, false positive, true negative, and false negative values will be listed. Second, summary ROC curves will be created. The Bivariate Model for performing meta-analyses will be used.

Assessment of heterogeneity

Initially, clinical heterogeneity will be assessed by determining whether study inclusion criteria are sufficiently similar to the pooled results. If heterogeneity is found, characteristics of the differing studies will be carefully investigated. If it seems that heterogeneity is due to the existence of specific risks of bias in some studies, then the meta-analysis will be restricted to studies that do not contain those risks. To ensure sensitivity analysis, we will exclude all studies that are appraised as having a high risk of bias.

Subgroup analysis

Subgroup analysis will be used for different age and gender characteristics. Another subgroup analysis will be used for cHL and different comorbidities according to their type and severity. Another subgroup analysis will be used for initial disease stage and type of chemotherapy given. If the data are available in primary studies, we will perform subgroup analysis according to; Progression-free survival (PFS); Standardized PET using Body Phantom experiments.

Meta-bias assessment

To show potential reporting bias, we will use funnel plots if more than ten studies are available. Begg's rank correlation and Egger's regression tests will be used for detecting and correcting publication bias.

Confidence in cumulative evidence

Based on the results and quality of evidence, the 'Grading of Recommendation Assessment, Development and Evaluation' (GRADE) tool will be used (26). Quality of evidence will be assessed across the domains of: risk of bias, consistency, directness, precision, and publication bias. Quality will be assessed as: high (further research is very unlikely to alter confidence in the

accuracy estimate), moderate (further research will likely impact confidence in the accuracy estimate, and may change the estimate), low (further research is very likely to impact confidence in the accuracy estimate, and will likely change the estimate), or very low (the accuracy estimate is very uncertain).

ETHICS AND DISSEMINATION

This systematic review protocol was crafted in February, 2016. Next, the systematic review development team will begin performing the protocol described herein. Dissemination of results will be targeted at patients and oncology practitioners through publication in a peer-reviewed journal. Ethical assessment is unnecessary as only existing sources of literature will be queried and evaluated.

Acknowledgements

We would like to thank the Charlesworth Group for professional editing of this manuscript. Grant funding for this work was provided by the Faculty of Medicine and Dentistry, Palacký University in Olomouc, Czech Republic (IGA_LF_2016_001 and RVO: 61989592) and by Takeda Pharmaceuticals International AG (IISR-2015-101289).

AUTHOR CONTRIBUTIONS

VP and MK conceptualized and designed the study. All authors contributed to selection criteria development, risk of bias assessment strategy, and data extraction. MK, JK and DT were methodologists. VP, VB and TP were Hodgkin lymphoma content experts. All authors (VP, MK, VB, TP) read, provided feedback, and approved the final manuscript.

COMPETING INTERESTS

None to declare.

REFERENCES:

1. Shenoy P, Maggioncalda A, Malik N, *et al.* Incidence Patterns and Outcomes for Hodgkin Lymphoma Patients in the United States. *Adv Hematology*2010;**2011**:725219. doi:10.1155/2011/725219

2. Dusek, Pavlík, Májek, *et al.* Estimating cancer incidence, prevalence, and the number of cancer patients treated with antitumor therapy in 2015 and 2020 - analysis of the Czech National Cancer Registry. *Klinická Onkologie Casopis Ceské Slovenské Onkologické Spolecnosti*2015;**28**:30–43. doi:10.14735/amko201530
3. Canellos G, Rosenberg S, Friedberg J, *et al.*Treatment of Hodgkin Lymphoma: A 50-Year Perspective. *J Clin Oncol* 2014;**32**:163–168. doi:10.1200/JCO.2013.53.1194
4. Josting, Rudolph, Mapara, *et al.* Cologne high-dose sequential chemotherapy in relapsed and refractory Hodgkin lymphoma: results of a large multicenter study of the German Hodgkin Lymphoma Study Group (GHSG). *Ann Oncol* 2005;**16**:116–123. doi:10.1093/annonc/mdi003
5. Moskowitz CH, Matasar MJ, Zelenetz AD, *et al.*Normalization of pre-ASCT, FDG-PET imaging with second-line, non-cross-resistant, chemotherapy programs improves event-free survival in patients with Hodgkin lymphoma. *Blood* 2012;**119**:1665–70. doi:10.1182/blood-2011-10-388058
6. Barrington S, Mikhaeel G, Kostakoglu L, *et al.* Role of Imaging in the Staging and Response Assessment of Lymphoma: Consensus of the International Conference on Malignant Lymphomas Imaging Working Group. *J Clin Oncol* 2014;**32**:3048–3058. doi:10.1200/JCO.2013.53.5229
7. Meignan M, Barrington S, Itti E, *et al.* Report on the 4th International Workshop on Positron Emission Tomography in Lymphoma held in Menton, France, 3–5 October 2012. *Leukemia Lymphoma* 2014;**55**:31–37. doi:10.3109/10428194.2013.802784
8. Carlier T, Bailly C. State-Of-The-Art and Recent Advances in Quantification for Therapeutic Follow-Up in Oncology Using PET. *Front Medicine* 2015;**2**. doi:10.3389/fmed.2015.00018
9. Meignan M, Sasanelli M, Casasnovas R, *et al.*Metabolic tumour volumes measured at staging in lymphoma: methodological evaluation on phantom experiments and patients. *Eur J Nucl Med Mol I*2014;**41**:1113–1122. doi:10.1007/s00259-014-2705-y
10. Meignan M. VI. FDG-PET as a biomarker in lymphoma: from qualitative to quantitative analysis.*Hematol Oncol* 2015;**33**:38–41. doi:10.1002/hon.2214

11. Kobe C, Kuhnert G, Kahraman D, *et al.* Assessment of Tumor Size Reduction Improves Outcome Prediction of Positron Emission Tomography/Computed Tomography After Chemotherapy in Advanced-Stage Hodgkin Lymphoma. *J Clin Oncol* 2014;**32**:1776–1781. doi:10.1200/JCO.2013.53.2507
12. Kanoun S, Rossi C, Berriolo-Riedinger A, *et al.*Baseline metabolic tumour volume is an independent prognostic factor in Hodgkin lymphoma. *Eur J Nucl Med Mol I* 2014;**41**:1735–1743. doi:10.1007/s00259-014-2783-x
13. Ceriani L, Martelli M, Zinzani PL, *et al.* Utility of baseline 18FDG-PET/CT functional parameters in defining prognosis of primary mediastinal (thymic) large B-cell lymphoma. *Blood* 2015;**126**:950–6. doi:10.1182/blood-2014-12-616474
14. Gallamini A, Hutchings M, Rigacci L, *et al.* Early Interim 2-[18F]Fluoro-2-Deoxy-D-Glucose Positron Emission Tomography Is Prognostically Superior to International Prognostic Score in Advanced-Stage Hodgkin's Lymphoma: A Report From a Joint Italian-Danish Study. *J Clin Oncol* 2007;**25**:3746–3752. doi:10.1200/JCO.2007.11.6525
15. Adams H, Nievelstein R, Kwee T. Prognostic value of interim FDG-PET in Hodgkin lymphoma: systematic review and meta-analysis. *Brit J Haematol*2015;**170**:356–366.
16. Sickinger MT, von Tresckow B, Kobe C, et al. Positron emission tomography-adapted therapy for first-line treatment in individuals with Hodgkin lymphoma. Cochrane database of systematic reviews 2015;**1**:Cd010533.
17. Sickinger MT, von Tresckow B, Kobe C, et al. PET-adapted omission of radiotherapy in early stage Hodgkin lymphoma-a systematic review and meta-analysis. Critical reviews in oncology/hematology 2016;**101**:86-92.
18. Rossi, Kanoun, Berriolo-Riedinger, *et al.* Interim 18F-FDG PET SUVmax Reduction Is Superior to Visual Analysis in Predicting Outcome Early in Hodgkin Lymphoma Patients. *J Nucl Med*2014;**55**:569–573. doi:10.2967/jnumed.113.130609
19. Tseng D, Rachakonda L, Su Z, *et al.* Interim-treatment quantitative PET parameters predict progression and death among patients with hodgkin's disease. *Radiat Oncol* 2012;**7**:5. doi:10.1186/1748-717X-7-5
20. Moher D, Shamseer L, Clarke M, *et al.* Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev 2015*;**4**:1.

21. Campbell J, Klugar M, Ding S, et al. *The Systematic Review of studies of diagnostic test accuracy*. In: JBI, ed. Joanna Briggs Institute Reviewers' Manual: 2015 edition/supplement. Adelaide, South Australia: The University of Adelaide, 2015:1-46.
22. Swerdlow S, Campo E, Harris N. WHO Classification of Tumours, Volume 2; IARC WHO Classification of Tumours, No 2: Lyon, France, IARC Press, 2008.
23. Eichenauer DA, Engert A, Andre M, et al. Hodgkin's lymphoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Annals of oncology : official journal of the European Society for Medical Oncology / ESMO 2014;**25 Suppl 3**:iii70-5.
24. Hoppe RT, Advani RH, Ai WZ, et al. Hodgkin lymphoma, version 2.2012 featured updates to the NCCN guidelines. Journal of the National Comprehensive Cancer Network : JNCCN 2012;**10**(5):589-97
25. Campbell JM, Klugar M, Ding S, *et al.* Diagnostic test accuracy: methods for systematic review and meta-analysis. *Int J Evid Based Healthc 2015*;**13**(3):154-62.
26. Munn Z, Porritt K, Aromataris E, et al. Summary of Findings Tables for Joanna Briggs Institute Systematic Reviews. The University of Adelaide, South Australia: The Joanna Briggs Institute 2014.

Appendix I: Critical appraisal instrument

Appendix I: Critical appraisal checklist

JBI Critical Appraisal Checklist for Diagnostic Test Accuracy Studies

Reviewer_____Date_____

Author_____Year_____Record Number_____

	Yes	No	Unclear	Not applicable
1. Was a consecutive or random sample of patients enrolled?	☐	☐	☐	☐
2. Was a case-control design avoided?	☐	☐	☐	☐
3. Did the study avoid inappropriate exclusions?	☐	☐	☐	☐
4. Were the index test results interpreted without knowledge of the results of the reference standard?	☐	☐	☐	☐
5. If a threshold was used, was it pre-specified?	☐	☐	☐	☐
6. Is the reference standard likely to correctly classify the target condition?	☐	☐	☐	☐
7. Were the reference standard results interpreted without knowledge of the results of the index test?	☐	☐	☐	☐
8. Was there an appropriate interval between the index test and the reference standard?	☐	☐	☐	☐
9. Did all patients receive the same reference standard?	☐	☐	☐	☐
10. Were all patients included in the analysis?	☐	☐	☐	☐

Overall appraisal: Include ☐ Exclude ☐ Seek further info ☐

Appendix II: Data extraction instrument

Author/Date	
Inclusion/exclusion criteria: i.e. presenting symptoms, results from previous tests	Inclusion: Exclusion:
Sample size	
Participant demographics (i.e. age, sex, spectrum of presenting symptoms, comorbidity, current treatments, recruitment centres)	
Study methodology (consecutive or random; retrospective or prospective)	
Period that study was carried out (beginning and end date)	
Index test description (including criteria for positive test)	
Reference test description (including criteria for positive test)	
Geographical location of data collection	
Setting of data collection	
Persons executing and interpreting index tests (numbers, training, and expertise)	
Persons executing and interpreting reference test	
Index/reference time interval (and treatments carried out in between)	
Distribution of severity of disease in those with target condition	
Other diagnoses in those without target condition	
Adverse events from index test	
Adverse events from reference test	

Index test results Threshold=	**Condition positive**	**Condition negative**	**Total**
Index test positive (T+)			
Index test negative (T-)			
Total			

PRISMA-P (Preferred Reporting Items for Systematic review and Meta-Analysis Protocols) 2015 checklist: recommended items to address in a systematic review protocol*

Section and topic	Item No	Checklist item
ADMINISTRATIVE INFORMATION		
Title:		
Identification	1a	Identify the report as a protocol of a systematic review Identified in tittle: **The accuracy of Quantitative interim PET compared to Qualitative interim PET in prognosis of Hodgkin lymphoma: a systematic review protocol of diagnostic test accuracy**
Update	1b	If the protocol is for an update of a previous systematic review, identify as such – N/A
Registration	2	If registered, provide the name of the registry (such as PROSPERO) and registration number **- Systematic review registration number** PROSPERO; CRD42016027953 First paragraph in Methods section
Authors:		
Contact	3a	Provide name, institutional affiliation, e-mail address of all protocol authors; provide physical mailing address of corresponding author – First page – all institutional affiliation, e-mail address just to corresponding author, as it is usual for BMJ in other protocols
Contributions	3b	Describe contributions of protocol authors and identify the guarantor of the review – Contributions identified on the first page: **Contributions** Vít Procházka and Miloslav Klugar were responsible for the study conception and design. All authors contributed to the development of the selection criteria, the risk of bias assessment strategy, and data extraction. Miloslav Klugar, Jitka Klugarová and Dagmar Tučková are methodologists. Vít Procházka, Veronika Bachanova and Tomáš Papajík are the content experts for Hodgkin lymphoma. All authors (Vít Procházka, Miloslav Klugar, Veronika Bachanova, Tomáš Papajík) read, provided feedback and approved the final manuscript.
Amendments	4	If the protocol represents an amendment of a previously completed or published protocol, identify as such and list changes; otherwise, state plan for documenting important protocol amendments – Identified on the first page: **Amendments** This protocol is not an amendment to another existing protocol for the systematic review of diagnostic test accuracy for the accuracy of Quantitative interim PET compared to Qualitative interim PET in prognosis of Hodgkin lymphoma. If necessary, this protocol will be accompanied in the future by amendments indicating a description of the change(s) made and the rationale for making it (them).
Support:		

Sources	5a	Indicate sources of financial or other support for the review – Indicated on the second page: **Support** This paper was supported by grants provided by the Faculty of Medicine and Dentistry, Palacký University in Olomouc, Czech Republic (IGA_LF_2016_001 and RVO: 61989592) and by Takeda Pharmaceuticals International AG (IISR-2015-101289). This review will disclose all financial and non-financial sources of support. Further this review will not be sponsored by any pharmaceutical companies and any financial sources will be drawn only from the above mentioned grant.
Sponsor	5b	Provide name for the review funder and/or sponsor - N/A
Role of sponsor or funder	5c	Describe roles of funder(s), sponsor(s), and/or institution(s), if any, in developing the protocol - N/A
INTRODUCTION		
Rationale	6	Describe the rationale for the review in the context of what is already known – Described in whole background part, mainly in this paragraph: A systematic review of the role of the qualitative interim PET in the HL patients has not been done yet. We presume, that analysis of qualitative tumor parameters will improve diagnostic (predictive) accuracy of interim FDG-PET and will help better identify a candidates for novel treatment approaches. This systematic review with its extensive search strategy may clarify this issue and influence practice by informing recommendations aimed at physicians and patients with Hodgkin lymphoma.
Objectives	7	Provide an explicit statement of the question(s) the review will address with reference to participants, interventions, comparators, and outcomes (PICO) – Provided on the 6th page: **Objectives** The objective of this review is to determine by comparison of Quantitative interim FDG-PET parameters with Qualitative interim FDG-PET parameters diagnostic test accuracy in Hodgkin lymphoma patients prognosis.
METHODS		
Eligibility criteria	8	Specify the study characteristics (such as PICO, study design, setting, time frame) and report characteristics (such as years considered, language, publication status) to be used as criteria for eligibility for the review – Specified on the 6/7th pages Study eligibility Types of participants This review will consider studies that include the adult Hodgkin lymphoma (as diagnosed using WHO diagnostic criteria). Excluded will be studies that include adolescents (under 18 years of age). Index test This review will consider studies that include as index test Quantitative PET (QT PET) Quantitative evaluation of interim PET by Metabolic Tumor Volume (MTV), Total Tumor Glycolysis (TLG) and Standardized Uptake Value (SUV). Reference test This review will consider studies that include as reference test qualitative PET (QL PET) visual evaluation of interim PET

		Diagnosis of interest This review will consider studies that evaluate accuracy of prognosis for Hodgkin Lymphoma patients. Change in negative-predictive value and positive-predictive value compared to QL-PET. Types of studies This review will consider only diagnostic cross-sectional study design for inclusion.
Information sources	9	Describe all intended information sources (such as electronic databases, contact with study authors, trial registers or other grey literature sources) with planned dates of coverage – Described on the page 7 in Search strategy
Search strategy	10	Present draft of search strategy to be used for at least one electronic database, including planned limits, such that it could be repeated – Described and example given in the part Search strategy, page 8
Study records:		
Data management	11a	Describe the mechanism(s) that will be used to manage records and data throughout the review – Described on the page 8 in part Study records
Selection process	11b	State the process that will be used for selecting studies (such as two independent reviewers) through each phase of the review (that is, screening, eligibility and inclusion in meta-analysis) – Described on the page 8 in parts Study records and Risk of bias in individual studies
Data collection process	11c	Describe planned method of extracting data from reports (such as piloting forms, done independently, in duplicate), any processes for obtaining and confirming data from investigators – Described on the page 8 in part Data collection process
Data items	12	List and define all variables for which data will be sought (such as PICO items, funding sources), any pre-planned data assumptions and simplifications – Described on the page 9 in part Data items/dealing with missing data
Outcomes and prioritization	13	List and define all outcomes for which data will be sought, including prioritization of main and additional outcomes, with rationale – Described on the page 9 in part Outcomes and prioritization
Risk of bias in individual studies	14	Describe anticipated methods for assessing risk of bias of individual studies, including whether this will be done at the outcome or study level, or both; state how this information will be used in data synthesis – Described on the page 8 in part Risk of bias in individual studies
Data synthesis	15a	Describe criteria under which study data will be quantitatively synthesised – Described on the page 9 in part Data synthesis
	15b	If data are appropriate for quantitative synthesis, describe planned summary measures, methods of handling data and methods of combining data from studies, including any planned exploration of consistency (such as I^2, Kendall's τ) – Described on the page 10 in parts Data synthesis and Assessment of heterogeneity
	15c	Describe any proposed additional analyses (such as sensitivity or subgroup analyses, meta-regression) – Described on the page 9 in part Data synthesis
	15d	If quantitative synthesis is not appropriate, describe the type of summary planned – Described on the page 9 in part Data synthesis
Meta-bias(es)	16	Specify any planned assessment of meta-bias(es) (such as publication bias across studies, selective reporting within studies) – Described on the page 10 in part Meta-bias assessment

Confidence in cumulative evidence	17	Describe how the strength of the body of evidence will be assessed (such as GRADE) – Described on the page 10 in part Confidence in cumulative evidence

*** It is strongly recommended that this checklist be read in conjunction with the PRISMA-P Explanation and Elaboration (cite when available) for important clarification on the items. Amendments to a review protocol should be tracked and dated. The copyright for PRISMA-P (including checklist) is held by the PRISMA-P Group and is distributed under a Creative Commons Attribution Licence 4.0.**

From: Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart L, PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ. 2015 Jan 2;349(jan02 1):g7647.