BMJ Open

Evaluating methodological quality of Prognostic models Including Patient-reported HeAlth outcomes iN oncologY (EPIPHANY): a systematic review protocol

Nina Deliu,[1] Francesco Cottone,[1] Gary S Collins,[2] Amélie Anota,[3] Fabio Efficace[1]

[1]Data Center and Health Outcomes Research Unit, Italian Group for Adult Hematologic Diseases (GIMEMA), Rome, Italy
[2]Centre for Statistics in Medicine, University of Oxford, Oxford, UK
[3]Methodology and Quality of Life in Oncology Unit (INSERM UMR 1098), University Hospital of Besançon, Besançon, France

Correspondence to
Dr Fabio Efficace;
f.efficace@gimema.it

ABSTRACT

Introduction While there is mounting evidence of the independent prognostic value of patient-reported outcomes (PROs) for overall survival (OS) in patients with cancer, it is known that the conduct of these studies may hold a number of methodological challenges. The aim of this systematic review is to evaluate the quality of published studies in this research area, in order to identify methodological and statistical issues deserving special attention and to also possibly provide evidence-based recommendations.

Methods and analysis An electronic search strategy will be performed in PubMed to identify studies developing or validating a prognostic model which includes PROs as predictors. Two reviewers will independently be involved in data collection using a predefined and standardised data extraction form including information related to study characteristics, PROs measures used and multivariable prognostic models. Studies selection will be reported following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines, with data extraction form using fields from the Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies (CHARMS) checklist for multivariable models. Methodological quality assessment will also be performed and will be based on prespecified domains of the CHARMS checklist. As a substantial heterogeneity of included studies is expected, a narrative evidence synthesis will also be provided.

Ethics and dissemination Given that this systematic review will use only published data, ethical permissions will not be required. Findings from this review will be published in peer-reviewed scientific journals and presented at major international conferences. We anticipate that this review will contribute to identify key areas of improvement for conducting and reporting prognostic factor analyses with PROs in oncology and will lay the groundwork for developing future evidence-based recommendations in this area of research.

Prospero registration number CRD42018099160.

INTRODUCTION

Patient-reported outcomes (PROs) play a key role in cancer research in addition to traditional

Strengths and limitations of this study

► This systematic review will critically evaluate the methodological quality of recent studies in oncology in order to examine the robustness of patient-reported outcomes (PROs) as independent prognostic factors for overall survival (OS). Compared with the existing literature, it will extensively consider and evaluate recently published studies in all cancer sites and all types and measures of PROs.

► The main strength of our review is the amount of relevant information, related to both clinical outcomes and methodological issues, collected through the designed data-extraction form.

► Another key strength of our review is that it will contribute to identify key areas of potential improvement for conducting methodologically sound prognostic modelling analyses, and would contribute to healthcare delivery; for example, by improving management of patients in clinical practice and by informing a more accurate stratification of patients in clinical trials.

► The main limitation of this systematic review might be the heterogeneity of selected studies in terms of study design and study population, therefore limiting our ability to provide strong recommendations for a specific cancer population.

► An additional limitation of this study is that we will only search one database.

clinical outcomes, such as overall survival (OS) or tumour response.[1 2]

A PRO is an umbrella term covering a measurement of any aspect of a patient's health status directly reported from the patient, and it is now included by the US, Food and Drug Administration (FDA) among the four types of Clinical Outcomes Assessment (COA) measures that can be used to determine whether drugs provide a treatment benefit.[3] The value of including PROs as endpoints in comparative studies is that of generating data

that can be used to more robustly inform patient care. The increased use of PROs in clinical trials[4] also reflects a growing recognition by major international health policy and regulatory authorities, of the importance of PROs in providing additional information about benefits, risks and cost of therapy, thus supporting claims in approved medical product labelling.[5–7]

While PROs are frequently included in clinical trial settings for treatment outcome comparisons, a more recent line of research has indicated that PROs, including functional aspects and symptomatic dimensions, can also provide additional prognostic information for survival outcomes. Furthermore, in some specific cancer populations, it has been shown that prognostic accuracy of well-established and validated prognostic models, based on laboratory and clinical data, can be improved by adding PROs.[8]

One of the most comprehensive critical reviews of studies investigating the prognostic value of baseline PROs on patients' survival was reported by Gotay et al.[9] Notably, out of the 39 clinical trials reviewed and published from 1989 to 2006, 36 studies found at least one PRO associated with OS in the multivariable analysis. PROs most frequently associated with OS are global quality of life (QoL) and physical functioning (PF). Further evidence that PROs provide additional prognostic value beyond traditional clinical measures in patients with cancer has been reported in several secondary pooled analyses of EORTC clinical trials.[10–13] For example, a large meta-analysis published in 2009 by Quinten et al,[11] comprising 30 randomised controlled trials (RCTs), including data for >7000 patients with a variety of cancers, found that including both traditional predictors and PROs in multivariable analyses increased the ability to predict OS by 6% compared with traditional information alone. These findings have been replicated across other cancer populations, further strengthening the scientific grounds of this association.[14–23] It is worth highlighting that diverse PRO questionnaires are often used which has frequently hindered clear comparisons across studies.[9]

RATIONALE

The completeness of reporting of studies describing the development or validation of prediction models in medicine, in general, is known to be suboptimal and major efforts have been made in recent years to raise standards and transparency of such studies. The Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) Statement has been recently developed to guide authors of reports describing prediction models.[24]

While there is a substantial body of literature documenting the association between PROs and survival outcomes, there is also an indication that the design, modelling techniques and reporting of many of these studies are frequently plagued by a number of methodological shortcomings.[25] The inclusion of PROs into studies of prognosis holds several major challenges. As PROs questionnaires are actually developed to measure the patient's perspective of

disease or treatment burden, they typically include multidomains which often measure aspects of a patient's health status that might be, more or less, correlated with each other. Indeed, one of the key methodological challenges is the possible *harmful multicollinearity* among questionnaire subscales entering the analysis, which may negatively affect the stability of the final prognostic model.[26]

One of the most rigorous reviews examining the methodological quality of studies evaluating the prognostic value of PROs was published by Mauer et al.[25] In this systematic review (which included studies published between 1989 and 2006), the authors found that most analyses were undertaken retrospectively, and on an exploratory basis, often not preplanned in the original study protocol. Also, the selection process for the prognostic factors was rarely described, and only in one study,[27] for example, the total number of included predictors relative to the total number of deaths, as recommended by Harrell et al,[28] was checked. Correlations among all PRO domains included in the model were rarely investigated, therefore limiting a critical appraisal of the potential impact of multicollinearity in final model selection.[26 29] Finally, most studies lacked an assessment of the added prognostic value of individual PROs as compared with traditional clinical factors alone, and none of the publications reviewed by Mauer et al[25] reported an external validation of the multivariable model (ie, an evaluation of model performance on a new dataset of patients).

While the independent prognostic value of PROs has been documented across many cancer disease sites and in several studies, the lack of methodological and analytic rigour makes current evidence far from being conclusive, likely undermining the translation of findings into realworld practice applications.[25] Therefore, a critical evaluation of the methodological quality of most recent studies is needed in order to examine the robustness of PROs included in prognostic models for OS.

OBJECTIVES

The aim of this systematic review is to critically investigate and report the quality of studies assessing the prognostic value of PROs for OS to identify methodological areas in need of particular attention and to provide evidencebased recommendations.

The specific aims of this systematic review are:
- ▶ To investigate the methodological quality of cancer prognostic modelling studies incorporating PRO data.
- ▶ To examine whether the methodological quality of these studies has improved over time.
- ▶ To describe the prognostic significance of PROs across different cancer types, stages and treatments.

METHODS AND ANALYSIS

This systematic review protocol adheres to the preferred reporting process and checklist outlined within the Preferred Reporting Items for Systematic Reviews and Meta-Analyses for Protocols 2015 (PRISMA-P) guidelines[30] (online supplementary file 1). In accordance with

these guidelines, this systematic review protocol has been registered with the International Prospective Register of Systematic Reviews (PROSPERO) on 28 June 2018 under registration number CRD42018099160. At the time of PROSPERO submission, formal screening of search results against eligibility criteria was ongoing. Any amendments to the study protocol will be documented contemporaneously on the PROSPERO database site.

ELIGIBILITY CRITERIA
Types of studies
According to the checklist for critical appraisal and data extraction for systematic reviews of prediction modelling studies (CHARMS) checklist[31] the proposed review will include:

► Prognostic model development studies without external validation on independent data.
► Prognostic model development studies with external validation on independent data.
► External model validation studies.

As a comprehensive review on the prognostic role of PROs for survival has already been published outcomes in 2008,[25] this systematic review will be based on the most recent literature, that is, previous 5 years (January 2013 to December 2017). All studies' designs, thus also secondary data analyses, will be considered for inclusion. This specific time window for the identification of studies published within the previous 5 years was selected in order to provide state of the art evidence in this area and to also possibly compare our recent findings with previously published work, reporting on older studies.[25] Only English language articles will be considered. Grey literature such as reports or working papers, conference presentations, opinions or commentaries and literature reviews will be excluded.

Participants
We will include studies on patients with a confirmed diagnosis of any cancer type, whatever stage and treatment(s) received. Both male and female adult patients (aged ≥18 years) will be considered.

Types of prognostic models
We will assess multivariable prognostic models and validation studies developed for patients with cancer with the following characteristics:

► OS as the outcome to be predicted, where OS is defined as the time from a prespecified time point (eg, date of randomisation, date of diagnosis and date of start of treatment) until death from any cause. Models predicting survival outcomes other than OS, for example, disease-free survival, progression-free survival or survival rate, will not be considered for inclusion.
► At least one baseline PRO as a predictor, defined as a patient-reported indicator of well-being and measured with a multidimensional instrument. A baseline

PRO might be a PRO collected at any fixed predefined time point, for example, before the start of intervention, before randomisation or, in general, at study entry. Studies reporting only changes in PRO scores will be excluded, as such measures may reflect rather than predict changes in survival time related to disease status.

Studies examining the association of baseline PROs with OS in a univariate or unadjusted fashion analysis only, without controlling for sociodemographic or clinical parameters, will not be included in this systematic review.

Search strategy
An electronic search strategy will be performed in PubMed to identify studies that match our eligibility criteria. As specific search filters are not available neither for PROs nor for prognostic studies, available published filters will be adapted and combined for a highly sensitive search strategy.[32] We will use medical subject heading (MeSH) and free-text words related to PROs and broad statistical strategies to identify prognostic modelling studies incorporating at least one PRO parameter and limited to adult (≥18 years old) cancer populations. Details on the key searching strategy used are reported in the online supplementary file 2 (PubMed key search).

Recognising potential limitations of electronic search strategies, we will supplement our search to identify potentially relevant studies from other sources, including reference lists of included studies, index-related articles on PubMed and existing relevant reviews.

Study selection and data extraction
Articles identified through database searching will be exported from PubMed into Microsoft Excel and duplicate records will be removed electronically. Record screening will occur in the exported Excel file to ensure that all retrieved references are fully tracked. One reviewer will compare the titles and abstracts yielded by the search against the inclusion criteria and will select articles for a full-text review. A second reviewer will confirm articles identified as eligible for inclusion, and reasons for exclusion will be documented for all excluded articles. The selection process, including search results and the reason for exclusion, will be reported in a flow diagram, as recommended in the PRISMA statement.[33]

At least two trained reviewers will independently be involved in the data extraction of studies meeting the eligibility criteria, which will be uploaded in a Google Drive folder for complete review. Data collection will be performed electronically in the non-proprietary Research Electronic Data Capture (REDCap, Vanderbilt University, USA; https://www.project-redcap.org/) platform,[34] which is a secure web application developed for building and managing online surveys and databases. Each reviewer will be provided with a personal password to access the study website and complete a standardised electronic data extraction form (eDEF).

The eDEF will collect information related to study characteristics (eg, study design/source of data, study objectives and study population) and PROs (eg, type of questionnaire used, handling of PROs in multivariate analysis, compliance and their association with OS). It will also include fields relevant to multivariable models, based on CHARMS[31] and TRIPOD[24] checklists, including modelling method, handling of predictors, predictors selection process, model performance and evaluation.

Items included in the eDEF were piloted in a sample of similar previously published (ie, prior January 2013) articles to assess their clarity and applicability. The full list of items included in the eDEF (see box 1) along with a comprehensive guideline on how to evaluate them is available to reviewers. This guideline is meant to minimise the number of discrepancies arising from the reviewing process. Should disagreements arise regarding extracted data, the reviewers will discuss and revisit the paper to reconcile any differences until consensus will be achieved. The reconciliation process will be facilitated by consulting a third senior reviewer or contacting study authors, if necessary. The final eDEF will then be validated by the project manager and locked for a final analysis.

We confirm that we have already started articles screening (January 2018) and we are currently validating the studies resulted eligible for inclusion. We plan to complete data extraction by the end of 2018.

Methodological quality assessment

Critical appraisal of methodological quality and risk of bias of included papers will be undertaken independently by two or more reviewers according to an adapted version of the CHARMS checklist for critical appraisal of prognostic model studies.[31] Following these guidelines, the reviewers will extract from individual studies relevant items belonging to the following domains: source of data, participants, outcome to be predicted, candidate predictors, sample size, missing data, model development, model evaluation and results (see box 1).

For each criterion, judgement will be made using one of the following three categories:

1. Low risk: if the criterion is adequately fulfilled in the study, that is, the study is at a low risk of bias for the given criterion.
2. High risk: if the criterion is not fulfilled in the study, that is, the study is at high risk of bias for the given criterion.
3. Unclear: if the study report does not provide sufficient information to allow for a clear judgement or if the risk of bias is unknown for one of the criteria listed above.

Presenting and reporting results

The selection process, including search results and the reason for exclusion, will be reported in a flow diagram, as recommended by the PRISMA statement.[33]

Data obtained from each eligible study will be summarised by providing descriptive tables reporting authors' names,

---

Box 1 Domains and items of the EPIPHANY data extraction form

1. Analysis type
   – Type of study analysis.
2. Study characteristics (model development)
   – First author.
   – Journal title.
   – Year of publication.
   – Cancer type.
   – Cancer stage.
   – Broad treatment type.
   – Was this study a secondary data analysis?
   – Source of data/study design.
   – Was the prognostic model incorporating PRO the primary objective of the study?
   – If no, what was the primary objective?
   – If yes, was sample size determination based on the primary objective (ie, prognostic model development/validation)?
   – Primary outcome predicted by the prognostic model.
   – Secondary outcomes predicted by the prognostic model.
   – Was the prognostic model with PRO identified in the abstract?
   – Any rationale for including PRO in the prognostic model?
   – Please specify when were baseline PROs collected.
   – Type of PRO instruments used.
3. Multivariable model and predictors (model development)
   – Multivariable model used.
   – Were model assumptions assessed and reported.
   – Was the number of events (deaths) checked in relation to the number of candidate predictors (Events per Variable)?
   – The rationale for the choice of PRO candidate variables used to build the final model.
   – Strategy to build the final model.
   – If stepwise, forward or backward, which stopping criterion was used?
   – Was multicollinearity among baseline PRO variables investigated?
   – If yes, which technique was used to account for baseline PRO multicollinearity?
   – Were any interactions among predictors examined?
   – How were continuous baseline PRO predictors handled?
   – If any continuous baseline PRO predictor, did the authors report its range?
   – For continuous baseline PRO that was categorised/dichotomised, how was this done?
   – Have the authors been clear on the number of categories of baseline PRO predictors?
   – Did the authors evaluate the performances of the prognostic model?
   – Which kind of measures were used to evaluate model performance?
   – Any additional analyses carried out?
4. Missing data (model development)
   – Were missing covariates mentioned in the paper?
   – The extent of missing baseline PRO questionnaires was reported.
   – If yes:
   – Report the number of returned PRO questionnaires.
   – Report the number of due PRO questionnaires.
   – The extent of missing baseline PRO data was given for each scale separately.
   – Is there any evidence of comparison of baseline sociodemographic and clinical characteristics of patients with and without completed baseline PRO questionnaires?

Continued

Box 1 Continued

- – Is there any evidence of comparison of survival between patients with and without completed baseline PRO questionnaires?
5. Results (model development)
   - – Overall number of patients analysed.
   - – Age of patients.
   - – Follow-up time.
   - – Number of events (ie, deaths).
   - – The overall number of candidate covariates considered prior to any model building and/or univariate screening.
   - – Number of baseline PRO candidate covariates considered prior to any model building and/or univariate screening.
   - – If (multi)collinearity was investigated:
   - – Was (multi)collinearity among baseline PRO variables detected?
   - – If yes, please specify (multi)collinear PRO covariates.
   - – If yes, was any PRO excluded from model building process for (multi)collinearity reasons?
   - – Total number of candidate variables entering in the multivariable model.
   - – Total number of baseline PRO candidate variables entering the multivariable model.
   - – Were any known risk factors forced into the model prior to building the final model?
   - – Was the prognostic model adjusted for sociodemographic or clinical characteristics?
   - – The overall number of predictors included in the final model.
   - – Did the final model include any baseline PRO predictor?
   - – If Yes, please report the number of baseline PRO predictors included in the final model.
   - – Instruments of baseline PRO included in the FINAL model.
   - – For each PRO included in the final model, please report the corresponding regression coefficient/HR.
   - – For each PRO included in the final model, please report the corresponding SD/CI.
   - – Please specify the confidence level of the reported CI.
   - – Did the authors reported HRs for the number of points increase different from one unit (eg, for every ten-points increase)?
   - – Were p values reported for all baseline PRO candidates entering in the multivariable model?
   - – Standard errors or CIs reported for all coefficients or HRs of baseline PRO candidates entering in the multivariable model?
   - – Was a prognostic index developed?
   - – Please report the value obtained for model performance.
   - – Comparison of model performance with and without baseline PROs was done.
   - – If yes, did baseline PROs increase predictive accuracy?
6. Model validation
   - – Did the authors also validate the model?
   - – If yes, how was this done?
   - – Did the authors evaluate the performances of the validation prognostic model?
   - – If yes, Which kind of measures were used to evaluate validation model performance?
   - – Did the authors discuss or mention concerns of overfitting/optimism?
   - – Have any modifications been suggested to the model based on the validation process?
   - – Based on the validation findings, did the authors refit or recalibrated the model?

Continued

Box 1 Continued

- – If the authors updated the model, briefly describe what did they do.
7. Model validation (In case of external validation only)
   - – Were any of the authors of the validation study an author of the development study?
   - – Is the outcome (ie, OS) of the validation study different from that of the development study?
   - – Please report OS of the population of the validation study.
   - – Did the authors discuss the distribution of baseline patients' characteristics of the validation study in comparison to those of the development study (either discussed or in tabulation)?
   - – If yes, did they report that they are:
   - – Were missing baseline PRO data mentioned in the validation study?
   - – Were baseline PRO predictors coded the same as in the development study?
   - – If no, please report what were the discrepancies and the reasons.
   - – If the prognostic model contained continuous baseline PRO risk predictors, were the ranges valid?
   - – Did the authors validate the developed prognostic model or a simplified version of it?
8. Discussion/Conclusions
   - – Did the authors discuss the PRO findings and their prognostic value on OS?
   - – Did the authors compare the model to another existing model?
9. Comments

OS, overall survival; PRO, patient-reported outcome.

publication year, study sample, type of cancer, instrument used to measure PRO and the main methodological findings (eg, multivariable model used, selection procedure, type of model validation, missing data information and model performance).

All relevant issues related to the methodological quality will be carefully reported and described separately. Results from the risk of bias assessment will be presented as counts and percentages, highlighting the most critical domains affecting the overall risk of bias.

Patient and public involvement
No patients will be involved in this study.

Ethics and dissemination
Given that this systematic review will use only published data, ethical permissions will not be required. Findings from this review will be published in peer-reviewed scientific journals and presented at international conferences relevant to various research fields of cancer outcomes research, biostatistics and QoL.

CONCLUSION
The risk of bias and potential usefulness of prediction models can be adequately assessed only by clear and full reporting of information, about all steps regarding their development and validation. Well-conducted and reported prognostic model studies with a PRO predictor

have great potential to improve healthcare delivery, for example, by improving the management of patients in clinical practice and by informing a more accurate stratification of patients in clinical trials. Our results will contribute to identifying key areas of improvement for conducting these analyses and will lay the groundwork for the development of evidence-based guidelines to improve the quality of research in this field.

REFERENCES

1. Secord AA, Coleman RL, Havrilesky LJ, *et al*. Patient-reported outcomes as end points and outcome indicators in solid tumours. *Nat Rev Clin Oncol* 2015;12:358–70.
2. Basch E. The missing voice of patients in drug-safety reporting. *N Engl J Med* 2010;362:865–9.
3. US Food and Drug Administration. Clinical outcome assessment qualification program. http://www.fda.gov/Drugs/DevelopmentApproval Process/DrugDevelopmentToolsQualificationProgram/ucm284077.htm (accessed 13 Mar 2018).
4. Efficace F, Fayers P, Pusic A, *et al*. Quality of patient-reported outcome reporting across cancer randomized controlled trials according to the CONSORT patient-reported outcome extension: A pooled analysis of 557 trials. *Cancer* 2015;121:3335–42.
5. US Food and Drug Administration. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims. 2009 http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf (accessed 19 Oct 2018).
6. European Medicine Agency, 2016. The use of patient-reported outcome (PRO) measures in oncology studies. Appendix 2 to the guideline on the evaluation of anticancer medicinal products in man. http://www.ema.europa.eu/docs/en_GB/document_library/Other/2016/04/WC500205159.pdf (accessed 9 Jul 2017).
7. Doward LC, Gnanasakthy A, Baker MG. Patient reported outcomes: looking beyond the label claim. *Health Qual Life Outcomes* 2010;8:89.
8. Efficace F, Cottone F, Abel G, *et al*. Patient-reported outcomes enhance the survival prediction of traditional disease risk classifications: an international study in patients with myelodysplastic syndromes. *Cancer* 2018;124:1251–9.
9. Gotay CC, Kawamoto CT, Bottomley A, *et al*. The prognostic significance of patient-reported outcomes in cancer clinical trials. *J Clin Oncol* 2008;26:1355–63.
10. Zikos E, Coens C, Quinten C, *et al*. The added value of analyzing pooled health-related quality of life data: a review of the EORTC PROBE initiative. *J Natl Cancer Inst* 2016;108:djv391.
11. Quinten C, Coens C, Mauer M, *et al*. Baseline quality of life as a prognostic indicator of survival: a meta-analysis of individual patient data from EORTC clinical trials. *Lancet Oncol* 2009;10:865–71.
12. Quinten C, Martinelli F, Coens C, *et al*. A global analysis of multitrial data investigating quality of life and symptoms as prognostic factors for survival in different tumor sites. *Cancer* 2014;120:302–11.
13. Ediebah DE, Coens C, Zikos E, *et al*. Does change in health-related quality of life score predict survival? Analysis of EORTC 08975 lung cancer trial. *Br J Cancer* 2014;110:2427–33.
14. Efficace F, Bottomley A, Smit EF, *et al*. Is a patient's self-reported health-related quality of life a prognostic factor for survival in non-small-cell lung cancer patients? A multivariate analysis of prognostic factors of EORTC study 08975. *Ann Oncol* 2006;17:1698–704.
15. Efficace F, Biganzoli L, Piccart M, *et al*. Baseline health-related quality-of-life data as prognostic factors in a phase III multicentre study of women with metastatic breast cancer. *Eur J Cancer* 2004;40:1021–30.
16. Efficace F, Cartoni C, Niscola P, *et al*. Predicting survival in advanced hematologic malignancies: do patient-reported symptoms matter? *Eur J Haematol* 2012;89:410–6.
17. Bottomley A, Coens C, Efficace F, *et al*. Symptoms and patient-reported well-being: do they predict survival in malignant pleural mesothelioma? A prognostic factor analysis of EORTC-NCIC 08983: randomized phase III study of cisplatin with or without raltitrexed in patients with malignant pleural mesothelioma. *J Clin Oncol* 2007;25:5770–6.
18. Cella D, Traina S, Li T, *et al*. Relationship between patient-reported outcomes and clinical outcomes in metastatic castration-resistant prostate cancer: post hoc analysis of COU-AA-301 and COU-AA-302. *Ann Oncol* 2018;29:392–7.
19. Efficace F, Gaidano G, Breccia M, *et al*. Prognostic value of self-reported fatigue on overall survival in patients with myelodysplastic syndromes: a multicentre, prospective, observational, cohort study. *Lancet Oncol* 2015;16:1506–14.
20. Roncolato FT, Gibbs E, Lee CK, *et al*. Quality of life predicts overall survival in women with platinum-resistant ovarian cancer: an AURELIA substudy. *Ann Oncol* 2017;28:1849–55.
21. Fiteni F, Vernerey D, Bonnetain F, *et al*. Prognostic value of health-related quality of life for overall survival in elderly non-small-cell lung cancer patients. *Eur J Cancer* 2016;52:120–8.
22. Wu X, Ye Y, Barcenas CH, *et al*. Personalized prognostic prediction models for breast cancer recurrence and survival incorporating multidimensional data. *J Natl Cancer Inst* 2017;109.
23. Movsas B, Hu C, Sloan J, *et al*. Quality of life analysis of a radiation dose-escalation study of patients with non-small-cell lung cancer: a secondary analysis of the radiation therapy oncology group 0617 randomized clinical trial. *JAMA Oncol* 2016;2:359–67.
24. Collins GS, Reitsma JB, Altman DG, *et al*. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ* 2015;350:g7594.
25. Mauer M, Bottomley A, Coens C, *et al*. Prognostic factor analysis of health-related quality of life data in cancer: a statistical methodological evaluation. *Expert Rev Pharmacoecon Outcomes Res* 2008;8:179–96.
26. Van Steen K, Curran D, Kramer J, *et al*. Multicollinearity in prognostic factor analyses using the EORTC QLQ-C30: identification and impact on model selection. *Stat Med* 2002;21:3865–84.
27. Naughton MJ, Herndon JE, Shumaker SA, *et al*. The health-related quality of life and survival of small-cell lung cancer patients: results of a companion study to CALGB 9033. *Qual Life Res* 2002;11:235–48.
28. Harrell FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 1996;15:361–87.
29. Deliu N, Efficace F, Collins GS, *et al*. Modelling strategies to improve estimates of prognostic factors analyses with patient reported outcomes: a simulation study. 24th Annual Conference of the International Society for Quality of Life Research (ISOQOL), Philadelphia, PA. *Quality Of Life Research* 2017;26:34–5. Abs. 204.2.
30. 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.
31. Moons KG, de Groot JA, Bouwmeester W, *et al*. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. *PLoS Med* 2014;11:e1001744.
32. Geersing GJ, Bouwmeester W, Zuithoff P, *et al*. Search filters for finding prognostic and diagnostic prediction studies in Medline to enhance systematic reviews. *PLoS One* 2012;7:e32844.
33. Moher D, Liberati A, Tetzlaff J, *et al*. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *J Clin Epidemiol* 2009;62:1006–12.
34. Harris PA, Taylor R, Thielke R, *et al*. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009;42:377–81.