Introduction to REMARK: Reporting tumour marker prognostic studies

Doug Altman

The EQUATOR Network
Centre for Statistics in Medicine, Oxford, UK

Reliable research

- "Reliability is determined by proper attention to study design and interpretation. No study can be perfect, but every study must be interpreted fairly. In particular, results must not be overinterpreted. The primary responsibility for interpretation and for specifying a study's limitations belongs to the investigator, although reviewers and editors have some role."

[Ransohoff, *J Clin Epidemiol* 2007]

Frequent methodological shortcomings of published prognostic studies

- **Design and data**
  - Poorly defined or unrepresentative cohort
  - Imprecise measurements
  - Unknown quality of tissue samples (when relevant)
  - Missing data
  - Unknown completeness of follow up
  - Some important predictors may be unavailable
  - Sample (much) too small
- **Analysis**
  - Unclear which variables have been adjusted for (& how)
  - Data-dependent (biased) analysis
    - e.g. data-derived cutpoint
- **Poor reporting of methods, data and results**

What should be reported?

Methods

- **All key aspects of how the study was done**
  - "Describe statistical methods with enough detail to enable a knowledgeable reader with access to the original data to verify the reported results."

    [International Committee of Medical Journal Editors]

Results

- **Main findings (corresponding to pre-specified plan)**

Biomarkers (prognostic markers) in cancer

- **Many markers have some prognostic potential**
- **As yet few markers have been demonstrated to be clinically useful**
  - oestrogen receptor (ER) status for breast cancer
  - c-erbB-2 (Her-2) for breast cancer
  - prostate specific antigen (PSA) for prostate cancer
- **Fewer still have been shown to predict who will benefit from a particular treatment**
  - ER and tamoxifen for breast cancer
  - Her-2 and herceptin for breast cancer

equator network

p53 as a prognostic marker in bladder cancer

- **168 published studies**
- **>10000 patients**

- **Interpretation:**
  "After 10 years of research, evidence is not sufficient to conclude whether changes in P53 act as markers of outcome in patients with bladder cancer."

  [Malats et al, *Lancet Oncology* 2005]

- **More studies like these won't clarify this question!**

Review of recent publications
[Mallett et al, *BJC* 2010]

- **Sample of 50 prognostic tumour marker studies from high impact cancer journals in 2006**
  - survival analysis; single biomarker; multivariate; not microarray
- **30% clearly defined the outcomes examined**
- **55% reported the number of eligible patients**
- **47% reported the number of outcome events**
- **55% and 23% reported patient and event numbers for all variables in univariate analyses;
  51% and 28% for multivariate analyses**
- **42% reported estimates for effect size (e.g. hazard ratios) for all variables included in univariate analyses;
  64% for multivariate analyses**

equator network

COMMENTARY

Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK)

Lisa M. McShane, Douglas G. Altman, Willi Sauerbrei, Sheila E. Taube, Massimo Gion, Gary M. Clark for the Statistics Subcommittee of the NCI-EORTC Working Group on Cancer Diagnostics

REMARK reporting guidelines
McShane et al: *JNCI, BJC, JCO, EJC* 2005

Guidelines for the REporting of tumor MARKer Studies (REMARK)

Introduction

1. State the marker examined, the study objectives, and any prespecified hypotheses.

Materials and Methods

Patients

2. Describe the characteristics (eg, disease stage or comorbidities) of the study patients, including their source and inclusion and exclusion criteria.

3. Describe treatments received and how chosen (eg, randomized or rule-based).

Specimen characteristics

4. Describe the type of biological material used (including control samples) and methods of preservation and storage.

Assay methods

5. Specify the assay method used and provide (or reference) a detailed protocol, including specific reagents or kits used, quality control procedures, reproducibility assessments, quantitation methods, and scoring and reporting protocols. Specify whether and how assays were performed blinded to the study end point.

Study design

6. State the method of case selection, including whether the study design was prospective or retrospective and whether stratification or matching (eg, by stage of disease or age) was used. Specify the time period from which cases were taken, the end of the follow-up period, and the median follow-up time.

7. Precisely define all clinical end points examined.

8. List all candidate variables initially examined or considered for inclusion in models.

9. Give rationale for sample size; if the study was designed to detect a specified effect size, give the target power and effect size.

Statistical analysis methods

10. Specify all statistical methods, including details of any variable selection procedures and other model-building issues, how model assumptions were verified, and how missing data were handled.

11. Clarify how marker values were handled in the analyses; if relevant, describe methods used for cutpoint determination.

Results

Data

12. Describe the flow of patients through the study, including the number of patients included in each stage of the analysis (a diagram may be helpful) and reasons for dropout. Specifically, both overall and for each subgroup extensively examined report the numbers of patients and the number of events.

13. Report distributions of basic demographic characteristics (at least age and sex), standard (disease-specific) prognostic variables, and tumor marker, including numbers of missing values.

Analysis and presentation

14. Show the relation of the marker to standard prognostic variables.

15. Present univariate analyses showing the relation between the marker and outcome, with the estimated effect (eg, hazard ratio and survival probability). Preferably provide similar analyses for all other variables being analyzed. For the effect of a tumor marker on a time-to-event outcome, a Kaplan-Meier plot is recommended.

16. For key multivariable analyses, report estimated effects (eg, hazard ratio) with confidence intervals for the marker and, at least for the final model, all other variables in the model.

17. Among reported results, provide estimated effects with confidence intervals from an analysis in which the marker and standard prognostic variables are included, regardless of their statistical significance.

18. If done, report results of further investigations, such as checking assumptions, sensitivity analyses, and internal validation.

Discussion

19. Interpret the results in the context of the prespecified hypotheses and other relevant studies; include a discussion of limitations of the study.

20. Discuss implications for future research and clinical value.

Reporting vs conduct: study methods

METHODS – each aspect of the methods

	Done well	Done poorly	Not done
Fully reported (=reproducible)			
Ambiguously or incompletely reported			
Not reported			

REMARK

Patients

Item 2 Describe the characteristics (e.g. disease stage or co-morbidities) of the study patients, including their source and inclusion and exclusion criteria.

"Inclusion criteria for the 2810 patients from whom tumour or cytosol samples were stored in our tumour bank (liquid nitrogen) were: primary diagnosis of breast cancer between 1978 and 1992 (at least 5 years of potential follow-up); no metastatic disease at diagnosis; no previous diagnosis of carcinoma, with the exception of basal cell skin carcinoma and cervical cancer stage I; no evidence of disease within 1 month of primary surgery … Patients with inoperable T4 tumours and patients who received neoadjuvant treatment before primary surgery were excluded." [Foekens et al, 1999]

"Patient Plasma Samples
After institutional review board approval, archived plasma samples were obtained from tumor banks at The University of Texas M. D. Anderson Cancer Center, Cedars-Sinai Medical Center, and Fox Chase Cancer Center. Our cohort included 164 patients with invasive epithelial ovarian carcinoma (EOC), 49 patients with benign ovarian neoplasms, and 75 unaffected age-matched controls. All patients were surgically staged based on the International Federation of Gynecology and Obstetrics (FIGO) staging system. Blood was collected before surgery, and plasma was separated by 2 rounds of centrifugation at 1000 g to remove cellular contamination; supernatant fluid was used for DNA extraction."   [Kamat et al, *Cancer* 2010]

Consider this example against REMARK items 2, 6, 9

equator network

REMARK

Specimen characteristics

Item 4 Describe type of biological material used (including controls), and methods of preservation and storage.

REMARK

Assay methods

Item 5 Specify the assay method used and provide (or reference) a detailed protocol, including specific reagents or kits used, quality control procedures, reproducibility assessments, quantitation methods, and scoring and reporting protocols. Specify whether and how assays were performed blinded to the study endpoint.

HSP-27 HAS NO DIAGNOSTIC OR PROGNOSTIC SIGNIFICANCE IN PROSTATE OR BLADDER CANCERS*

F. KRISTIAN STORM, M.D.
DAVID M. MAHVI, M.D.
KENNEDY W. GILCHRIST, M.D.

From the Departments of Surgery (Section of Oncology) and Pathology,
University of Wisconsin School of Medicine and
Comprehensive Cancer Center, Madison, Wisconsin

15

Example: hsp-27 in prostate and bladder cancer
[Storm et al, *Urology* 1993]

- **hsp-27 in prostate and bladder cancers**
  - 36 patients with prostate cancer
  - 24 patients with bladder cancer
- **No significant association found with survival**
- **"We conclude that hsp-27 expression has neither diagnostic nor prognostic significance and will not serve as a predictive biologic marker"**
- **Editorial comment: "…we can conclude with some confidence that hsp-27 will not provide the needed marker for either prostate or bladder cancer"**

Heat shock protein expression independently predicts clinical outcome in prostate cancer
[Cornford et al, *Cancer Res* 2000]

- **"We suggest that hsp27 expression provides novel diagnostic and prognostic information on individual patient survival which … would assist in determining tumor-specific management strategies."**
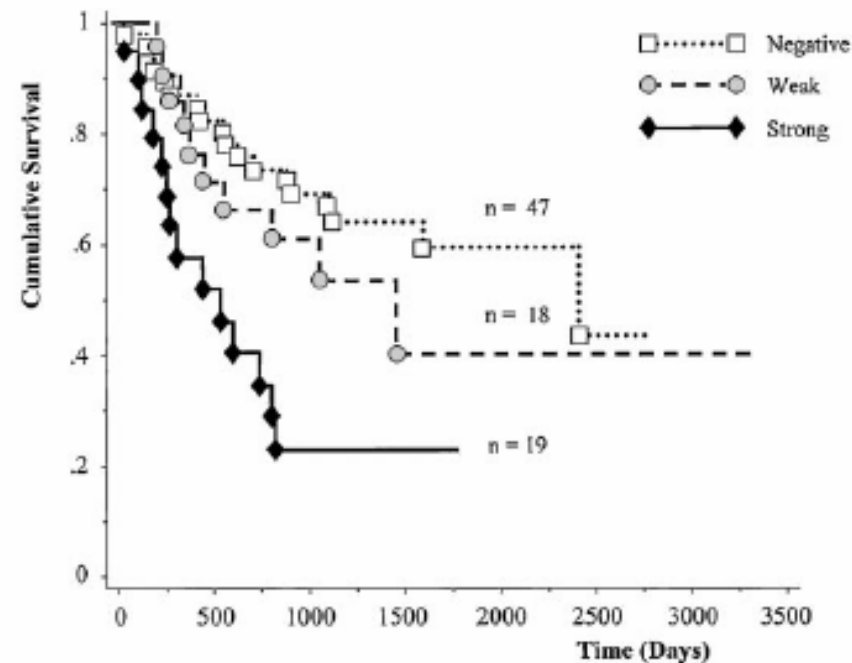

Fig. 3. hsp27 expression predicts outcome in advanced prostate cancers ($n = 85$; $\chi^2 = 20.78$; log-rank $P = 0.0001$).

REMARK

Item 12 Describe the flow of patients through the study, including the number of patients included in each stage of the analysis (a diagram may be helpful) and reasons for dropout. Specifically, both overall and for each subgroup extensively examined report the numbers of patients and the number of events.

Item 13 Report distributions of basic demographic characteristics (at least age and sex), standard (disease-specific) prognostic variables, and tumor marker, including numbers of missing values.

Table 2 REMARK profile of patients, variables and statistical analyses (Study profile for Pfisterer *et al*, 1994). The REMARK profile is shown for illustrative purposes in an adaptable format. Additional rows can be included for each multivariable analysis, subgroup analysis or further outcome investigated

(a) Patients, treatment and variables

Study and marker	Remarks
Marker (If non-binary: how was marker analysed? continuous or categorical. If categorical, how are cutpoints determined?)	M = ploidy (diploid, aneuploid)
Further variables (variables collected, variables available for analysis, baseline variables, patient and tumour variables)	v1 = age, v2 = histological type, v3 = grade, v4 = residual tumour, v5 = stage, v6 = ascites[a], v7 = oestrogen[a], v8 = progesterone[a], v9 = CA-125[a]

Patients	*n*	Remarks
Assessed for eligibility	257	*Disease*: advanced ovarian cancer, stage III and IV *Patient source*: Surgery 1982–1990, University Medical Center Freiburg *Sample source*: archived specimens available
Excluded	73	General exclusion criteria[b], non-standard therapy[b], CV > 7%[b]
Included	184	Previously untreated.
		Treatment: all had platinum-based chemotherapy after surgery
With outcome events	139	Overall survival: death from any cause

(b) Statistical analyses of survival outcomes

Analysis	Patients	Events	Variables considered	Results/remarks
A1: Univariable (Provide for all variables. Give numbers as range if variables have different numbers of missing values)	184	139	M,v1 to v5	Tab 2, Fig1
A2: Multivariable	174	133	M,v1,v3 to v5	Tab 3 (v2 omitted because of many missing data; backward selection, see text)
A3: Effect for ploidy adjusted for v4	184	139	M, 4	Fig 2 (based on the result of A2)
A4: Interaction ploidy and stage	175	133	M, v1, v2, v4, v5	See text
A5: Ploidy in stage subgroups				
v5 = III	128	88	M	Fig 3
v5 = IV	56	51	M	Fig 4

[a]Not considered for survival outcome as these factors are not considered as 'standard' factors and/or number of missing values are relatively large. [b]Values not given in the paper.