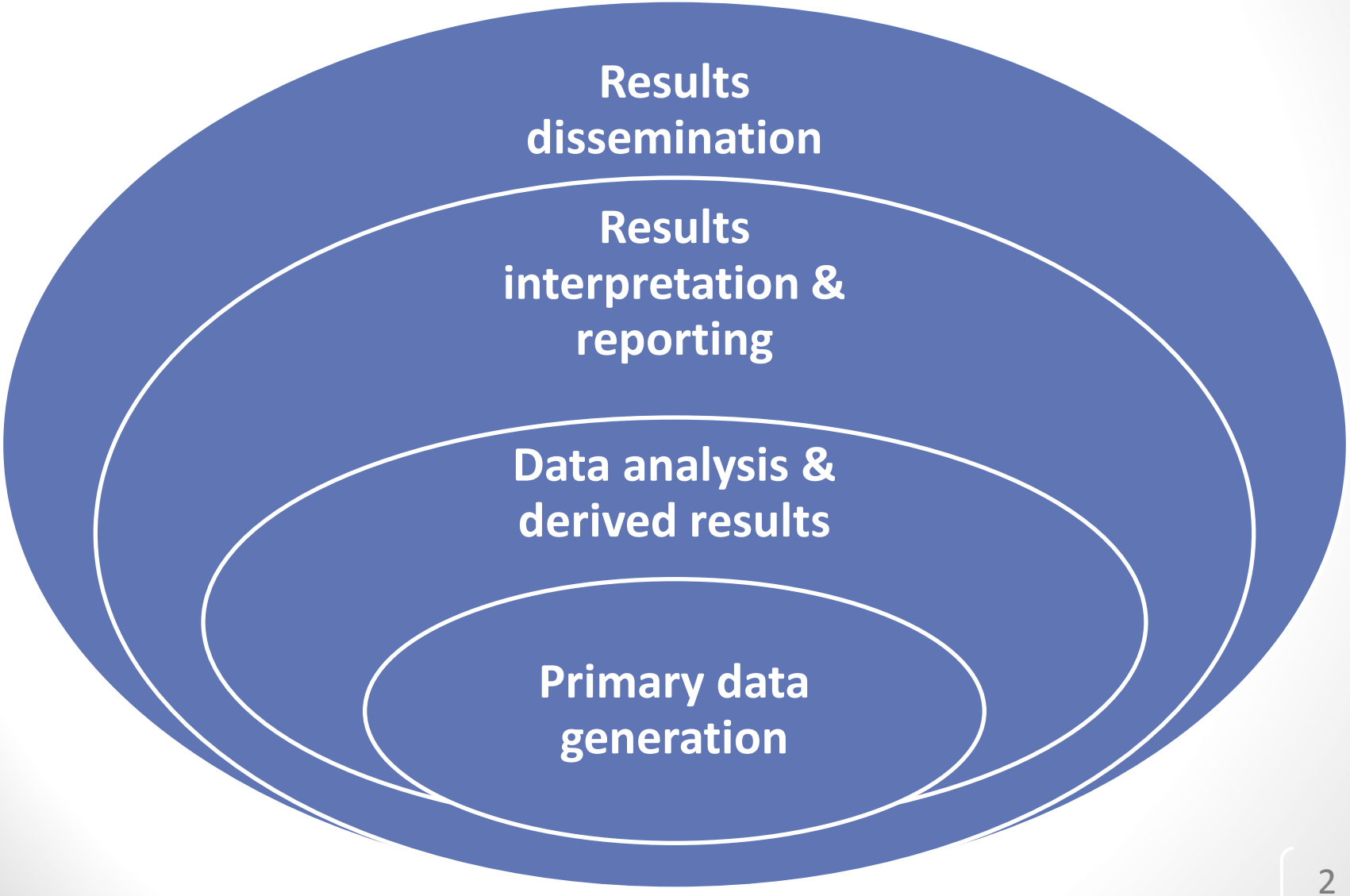
The Role of Reporting Guidelines in Promoting Reproducible Research

Presentation to the
NCI Board of Scientific Advisors

Lisa McShane, PhD
Biometric Research Branch, DCTD
National Cancer Institute

March 4, 2013

Propagation of Irreproducible Research



REMARK: **RE**porting guidelines for tumor **MARK**er prognostic studies

Lisa M. McShane, Douglas G. Altman, Willi Sauerbrei, Sheila E. Taube, Massimo Gion, and Gary M. Clark for the Statistics Subcommittee of the NCI-EORTC Working Group on Cancer Diagnostics (*J Natl Cancer Inst* 2005; 97:1180-1184, and simultaneously in *BJC, EJC, JCO, NCPO)*

Recommended reporting elements to facilitate

- Evaluation of **appropriateness & quality** of study design, methods, and analysis
- Understanding of **context** in which conclusions apply
- **Reproducibility**
- **Comparisons** across studies, including formal meta-analyses

REMARK: Target Studies

- Studies relating marker values to clinical events (e.g., recurrence, death, response)
- NOT primarily aimed at biological discovery studies, but use encouraged to extent possible
  - Patients
  - Specimens
  - Assays
- NOT sufficient for studies developing multiplex classifiers/risk scores (e.g., derived from omics data), but applicable to studies assessing them

State of the Tumor Marker Literature

VOLUME 25 · NUMBER 33 · NOVEMBER 20 2007

JOURNAL OF CLINICAL ONCOLOGY

ASCO SPECIAL ARTICLE

American Society of Clinical Oncology 2007 Update of Recommendations for the Use of Tumor Markers in Breast Cancer

Lyndsay Harris, Herbert Fritsche, Robert Mennel, Larry Norton, Peter Ravdin, Sheila Taube, Mark R. Somerfield, Daniel F. Hayes, and Robert C. Bast Jr

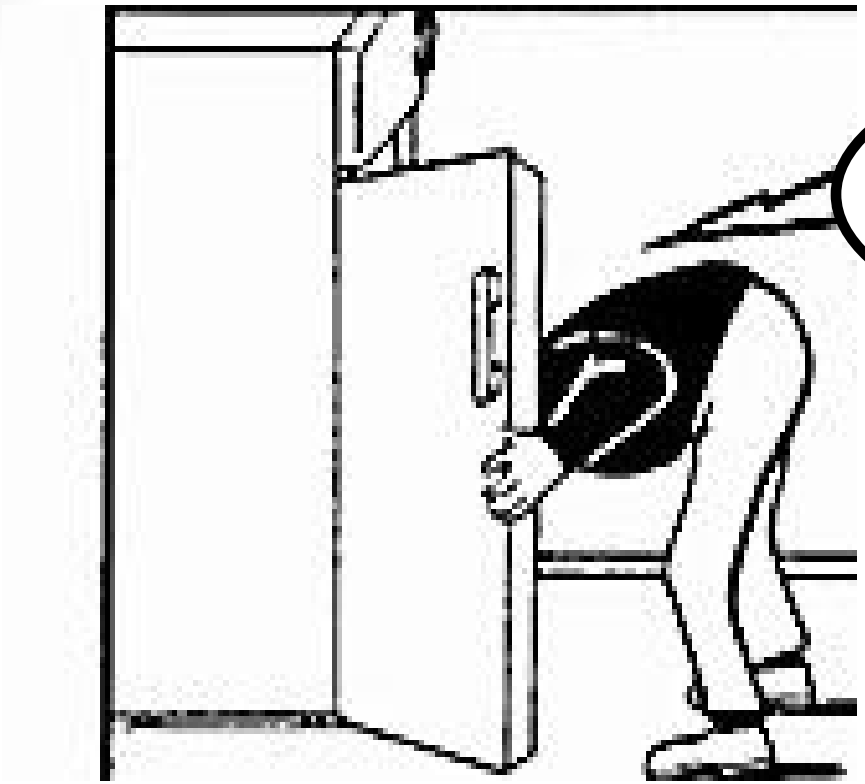
Purpose: To update the recommendations for the use of tumor marker tests in the prevention, screening, treatment, and surveillance of breast cancer.

". . . primary literature is characterized by studies that included small patient numbers, that are retrospective, and that commonly perform multiple analyses until one reveals a statistically significant result. . .many tumor marker studies fail to include descriptions of how patients were treated or analyses of the marker in different treatment subgroups. The Update Committee hopes that adherence to . . . REMARK criteria will provide more informative data sets in the future.

REMARK Elements: Introduction

- State all marker(s) examined
- Study objectives
- Pre-specified hypotheses

Common Tumor Marker Study Design



What can we do with our marker on these 89 specimens?

- "Convenience" specimens
- Heterogeneous patient characteristics
- Treatments: Unknown, non-randomized, not standardized
- Insufficient sample size (underpowered)
- Uncertain specimen and data quality

REMARK Elements: Materials & Methods

- Patients
  - Inclusion/exclusion (e.g., stage, subtype), source, treatments
- Specimen characteristics
  - Format, collection, preservation, storage
  - See BRISQ criteria (Moore et al, *Cancer Cytopathology* 2011; 119:92-101)

REMARK Elements: Materials & Methods (cont.)

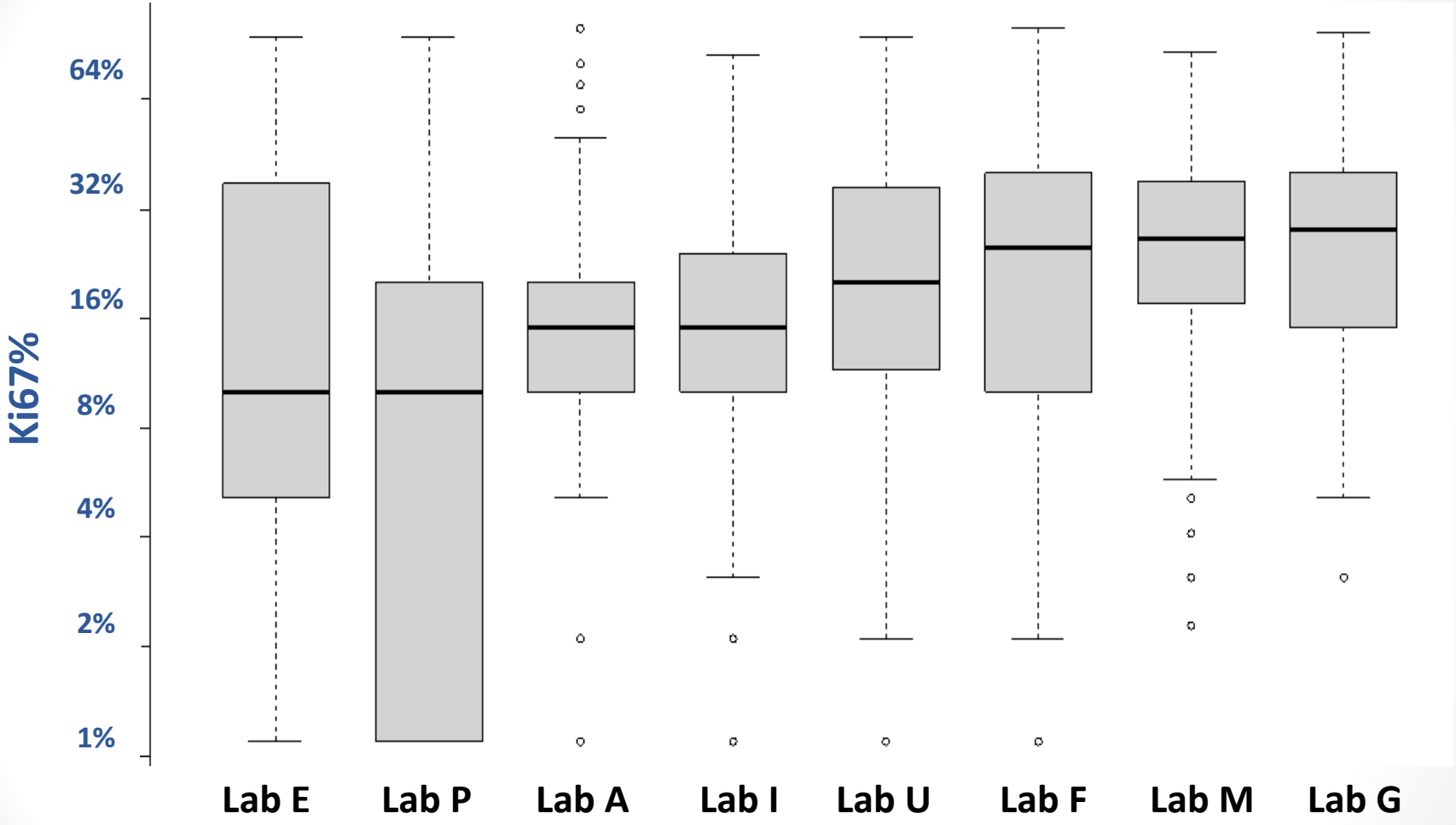
- Assay methods
  - Detailed protocol (reagents/kits), quantitation, scoring & reporting, reproducibility, blinding

Example: Systematic review (43 studies) of Ki67 in early breast cancer (Stuart-Harris et al, *The Breast* 2008; 17:323-334)

- English publication, Jan. 1995 – Sept. 2004
- ≥ 100 patients, OS or DFS endpoint

- Results
  - 7 different antibodies for IHC, single or combination
  - 19 different cutpoints, ranging from 0-30%
  - Significant between-study heterogeneity and publication bias

International Ki-67 Reproducibility Study
(Nielsen et al, *SABCS* 2012 abstract)



median: 10%

median: 28%

Consecutive TMA sections, single assay batch

REMARK Elements: Materials & Methods (cont.)

- Study design
  - Case selection (e.g., random, case-control), clinical endpoints, variables considered, sample size
- Statistical analysis methods
  - Models, variable selection, handling of missing data, multiple testing adjustments, validations

Statistical Analysis Methods

EUROPEAN JOURNAL OF CANCER 43 (2007) 2559–2579

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejconline.com

Almost all articles on cancer prognostic markers report statistically significant results

Panayiotis A. Kyzas[a], Despina Denaxa-Kyza[a], John P.A. Ioannidis[a,b,c,]*

[a]*Clinical and Molecular Epidemiology Unit, Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece*

[b]*Biomedical Research Institute, Foundation for Research and Technology-Hellas, Ioannina, Greece*

[c]*Institute for Clinical Research and Health Policy Studies, Department of Medicine, Tufts-New England Medical Center, Boston, USA*

"If you torture the data long enough they will confess to anything."

Source unknown

Statistical Analysis: Multiple Testing

- Multiple markers
- Multiple endpoints
- Multiple subgroups
- Multiple marker cutpoints
- Multiple models with multiple variables

Example: 8 subgroups defined by 3 binary factors

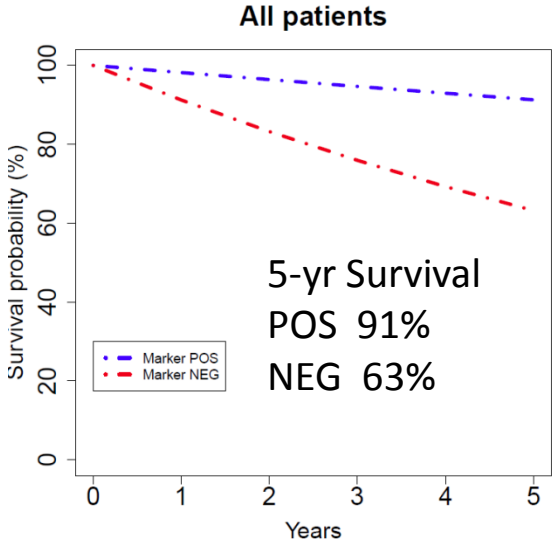
Number of independent tests ($\alpha = 0.05$ per test)	Probability observe ≥ 1 statistically significant ($p<0.05$) result
1	0.05
2	0.10
3	0.14
4	0.19
5	0.23
6	0.26
7	0.30
8	0.34
9	0.37
10	0.40

REMARK Elements: Results

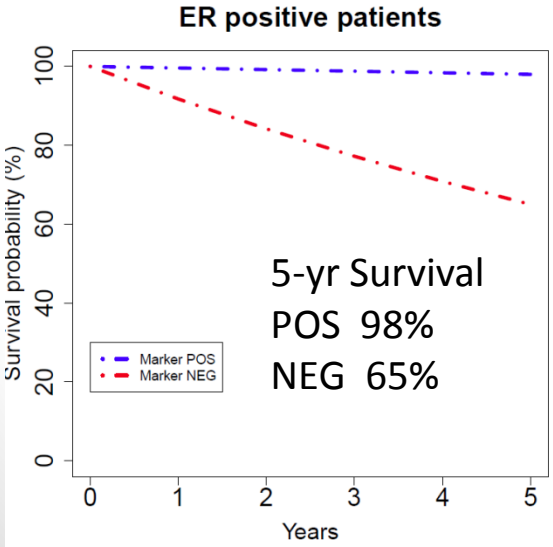
- Data
  - Numbers of patients and events
  - Demographic characteristics
  - Standard prognostic variable distribution
  - Tumor marker distribution
- Analysis & presentation
  - Univariate analyses (marker vs. standard prognostic variables, marker vs. outcome)
  - Multivariable analyses (association of marker with outcome after adjustment for standard prognostic variables)
  - Measures of uncertainty for reported effect estimates

REMARK Elements: Results (cont.)

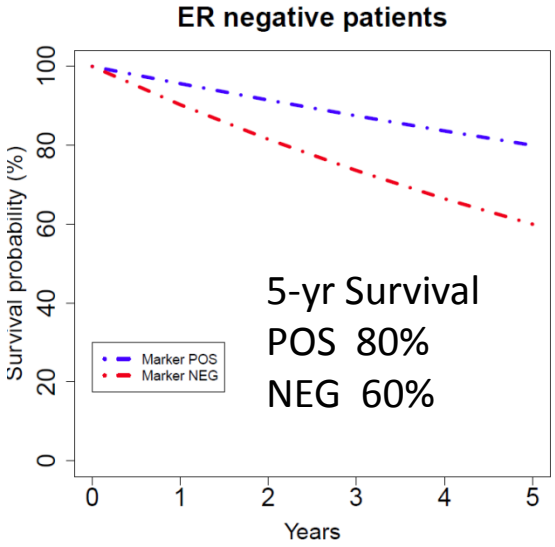
- Viewing in context of standard factors and treatments received



65%



35%



REMARK Elements: Discussion

- Interpretation in context of pre-specified hypotheses
- Relevance to other studies
- Limitations
- Future research
- Clinical value

REMARK Status & Future

- Explanation & Elaboration: Altman et al, *PLoS Medicine* 2012; 9(5):e1001216 (also *BMC Medicine* 2012; 10:51)
- Plans for "before vs. after" comparisons of reporting
- Journals stating REMARK adherence requirements: *Ann Oncol, Breast Cancer Res Treat, Clin Cancer Res, J Clin Oncol, J Natl Cancer Inst, J Pathol*

Statement of editorial intent

Annals of Oncology 2012; 23:1931-1932

"Studies of 'prognostic' markers of no real future clinical utility and single biomarker studies will not be considered. Reports of studies into prognostic markers should be prospective and have a clear view of the practical clinical applications of the results. Retrospective analysis of biomarkers can be considered, if done within the framework of data collected from a prospective trial, with appropriate statistics and with multivariate analysis that includes established predictive/prognostic markers. Reports of prognostic tumor marker studies should follow the REMARK guidelines (available from www.equator-network.org)."

J. B. Vermorken
Editor-in-Chief

Concluding Remarks

- Poor study reporting is a significant impediment to achieving reproducible research
- Reporting will improve only with effort from all stakeholders
- Complete & transparent reporting is more fair
- Effort spent on good reporting is a smaller burden than time, effort and resources wasted on false leads