Title
On Lumping, Splitting, and the Nosology of Clinical Trial Populations and End Points

Journal
Journal of Clinical Oncology, 32(10)

ISSN
0732-183X

Authors
Mell, Loren K
Zakeri, Kaveh
Rose, Brent S

Publication Date
2014-04-01

DOI
10.1200/jco.2013.54.4429

Peer reviewed

Published Ahead of Print on February 18, 2014 as 10.1200/JCO.2013.54.4429
The latest version is at http://jco.ascopubs.org/cgi/doi/10.1200/JCO.2013.54.4429

On Lumping, Splitting, and the Nosology of Clinical Trial Populations and End Points

TO THE EDITOR: Wildiers et al[1] espouse conventional wisdom in reaching the conclusion that "all clinical trials in oncology should be without an upper age limit." We agree greater representation of elderly patients in clinical trials is needed but disagree with this conclusion, and feel two premises should be challenged:

1. "Clinical Trials Need to be Representative of the Whole Population in Whom the Treatment Will be Used Later"[1(p3715)]

Ideally, clinical trial populations should represent those in which its findings are applied; however, this is not an unqualified need. Clinical trials naturally create selective, idealized conditions that necessitate extrapolating their findings to a wider population than represented by their participants.[2] Furthermore, evidence that outcomes differ as a function of cancer trial participation is weak.[3] The authors emphasize that excluding elderly patients limits generalizability of treatment, and extrapolating findings from younger populations risks overutilization of expensive, ineffective treatments. However, underutilization of effective treatments is as much a concern as overutilization of ineffective treatments. They consider relaxed inclusion/exclusion criteria "the price society has to pay if it wants to ensure that older patients are not subjected to toxic therapies that provide no tangible clinical benefit."[1(p3712)] Although this seems correct, it sidesteps an inconvenient truth: this can substantially increase trial costs, without greatly augmenting knowledge. Including more patients unlikely to benefit from treatment increases sample size, without necessarily increasing power.[4-8] In resource-constrained settings, optimizing allocation strategies becomes paramount.[9] Using age as a selection factor does not imply exclusion from clinical trials generally. Different trials can be tailored to address specific patients' needs, which might be more consistent with personalized medicine. Rather than liberalizing entry criteria, investigators should define populations most likely to benefit from treatment, to maximize efficiency and value of information gained. Reproducing trials in elderly populations assuming that results from other trials are nongeneralizable is misguided. Forgoing the assurance from such studies might be considered the price we pay to increase the dimensionality of scientific knowledge.

2. "Overall Survival is Considered the Gold Standard in Clinical Trials, Especially When Evaluating the Superiority of New Treatments"[1(p3712)]

Wildiers et al[1] do not acknowledge that overall survival (OS) is itself a composite end point, comprised of competing causes of death. They admit composite end points are justified when "the expected effects on each component are similar based on clinical/biologic plausibility."[1(p3713)] This condition is generally unrealistic with respect to OS, because typically, if a cancer therapy is to be beneficial, it should reduce mortality from cancer, not other causes, and we hope this benefit is not offset by increased treatment-related or competing mortality. Previously, we have shown that effects on a composite end point can be interpreted as the weighted average of effects on cause-specific events.[4] As such, it is problematic when composite effects, but not cause-specific effects, are not designated explicitly, since an effect on cancer mortality or noncancer mortality tells us nothing about the effect on either outcome. Contrary to Wildiers et al's statement that using composite end points increases statistical efficiency, their use can predispose studies to both type I[10-11] and type II error[4-8]: type I error, because a positive result may be attributable in whole or part to an effect that is inconsistent with the treatment mechanism; type II error, because a negative result may be attributable to overestimating power in the presence of competing risks. Predisposition to both type I and type II error is the calling card of a bad model, leading us to think "OS as the gold standard"[1(p3712)] should go the way of the actual gold standard, formally abandoned in 1976.[12]

An approach the authors briefly address, which we advocate, is to consider coprimary end points in competing risks settings. Whether type I and/or II error must be adjusted for multiple testing depends on the framework applied. For example, coprimary end points can be aggregated, assigning type I and II error for the composite end point, with cause-specific effects designated explicitly.[4,13] Wildiers et al[1] imply that for this approach to be valid, a precondition "is that cause of death can be reliably ascertained." If ascertainment were biased, we agree that would be a problem, but if it were just imprecise, we are not convinced this precondition is necessary,[14] and fear it will discourage investigators from determining cause of death when it might be challenging, but not impossible. A concern with composite end points is that effects on the end point may be driven in part or whole by one of its components.[15] We agree that disease-specific event probabilities should be reported, but would add that competing event probabilities, and effects on competing events should be reported, to safeguard against publication bias. Since it is not typically reasonable to hypothesize that a cancer therapy will favorably affect noncancer mortality, we should view claims that treatments improve survival in the elderly with utmost scrutiny, particularly when the mechanism for this benefit cannot be evinced using clinical data. The authors list increased sample size among the limitations of coprimary end points; in contrast, we consider this a justifiable cost to verify the mechanism by which beneficial effects are achieved.

Specificity is the hallmark of every science, and one we should aspire to in clinical oncology. In general, we believe readers should be wary of rules for defining trial populations and end points without reference to a specific question. Age limits should not be off-limits for clinical trials. If one's hypothesis is that effects are homogeneous with respect to age, then it makes no difference if elderly patients are included, save concerns regarding efficiency. If one's hypothesis is that effects are heterogeneous, then elderly subpopulations could be studied separately. Generalizability in geriatric subpopulations is most likely to be a significant concern when there is reason to expect the cause-specific treatment effects (not composite effects) would differ

with age. Whether to lump or split populations or end points should depend on one's fundamental assumptions regarding effect heterogeneity, similar to the field of genetics.[16] If an effect is assumed to be homogeneous, lumping may be advantageous. However, in geriatric oncology, this assumption is often untenable, necessitating a conceptual framework with greater specificity.

Loren K. Mell and Kaveh Zakeri
University of California San Diego, La Jolla, CA

Brent S. Rose
Harvard Radiation Oncology Program, Harvard Medical School, Boston, MA

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The author(s) indicated no potential conflicts of interest.

REFERENCES

1. Wildiers H, Mauer M, Pallis A, et al: End points and trial design in geriatric oncology research: A joint European Organisation for Research and Treatment of Cancer–Alliance for Clinical Trials in Oncology–International Society of Geriatric Oncology position article. J Clin Oncol 31:3711-3718, 2013

2. Post PN, de Beer H, Guyatt GH: How to generalize efficacy results of randomized trials: Recommendations based on a systematic review of possible approaches. J Eval Clin Pract 19:638-643, 2013

3. Peppercorn JM, Weeks JC, Cook EF, et al: Comparison of outcomes in cancer patients treated within and outside clinical trials: Conceptual framework and structured review. Lancet 363:263-270, 2004

4. Mell LK, Jeong JH: Pitfalls of using composite primary end points in the presence of competing risks. J Clin Oncol 28:4297-4299, 2010

5. Gómez G, Lagakos SW: Statistical considerations when using a composite endpoint for comparing treatment groups. Stat Med 32:719-738, 2013

6. Rose BS, Jeong JH, Nath SK, et al: Population-based study of competing mortality in head and neck cancer. J Clin Oncol 29:3503-3509, 2011

7. Manola JB, Gray RJ: When bad things happen to good studies. J Clin Oncol 29:3497-3499, 2011

8. Latouche A, Porcher R: Sample size calculations in the presence of competing risks. Stat Med 26:5370-5380, 2007

9. Dilts DM, Cheng SK: The importance of doing trials right while doing the right trials. Clin Cancer Res 18:3-5, 2012

10. Zakeri K, Rose BS, Gulaya S, et al: Competing event risk stratification may improve the design and efficiency of clinical trials: Secondary analysis of SWOG 8794. Contemp Clin Trials 34:74-79, 2013

11. Mell LK, Carmona R, Gulaya S, et al: Cause-specific effects of radiotherapy and lymphadenectomy in stage I-II endometrial cancer: A population-based study. J Natl Cancer Inst 105:1656-1666, 2013

12. Elwell CK: Brief history of the gold standard in the United States. Washington, DC, Congressional Research Service Report for Congress 7-5700 R41887, 2011. http://www.fas.org/sgp/crs/misc/R41887.pdf

13. Mell LK, Lau SK, Rose BS, et al: Reporting of cause-specific treatment effects in cancer clinical trials with competing risks: A systematic review. Contemp Clin Trials 33:920-924, 2012

14. Korn EL, Dodd LE, Freidlin B: Measurement error in the timing of events: Effect on survival analyses in randomized clinical trials. Clin Trials 7:626-633, 2010

15. Freemantle N, Calvert M, Wood J, et al: Composite outcomes in randomized trials: Greater precision but with greater uncertainty? JAMA 289:2554-2559, 2003

16. McKusick VA: On lumpers and splitters, or the nosology of genetic disease. Perspect Biol Med 12:298-312, 1969

DOI: 10.1200/JCO.2013.54.4429; published online ahead of print at www.jco.org on February 19, 2014

