Competing risk and heterogeneity of treatment effect in clinical trials

David M Kent*1, Alawi Alsheikh-Ali1,2 and Rodney A Hayward3,4

Address: 1Institute for Clinical Research and Health Policy Studies, Department of Medicine, Tufts Medical Center, 750 Washington St, #63, Boston, MA, 02111, USA, 2Cardiac Arrhythmia Center, Division of Cardiology, Department of Medicine, Tufts Medical Center, Tufts University School of Medicine, 750 Washington Street, NEMC #108, Boston, Massachusetts 02111, USA, 3Veterans Affairs Ann Arbor Health Services Research and Development Service Center of Excellence and 4Department of Internal Medicine, University of Michigan, 3110 Taubman Health Care Center, Ann Arbor, MI 48109-0368, USA

Email: David M Kent* - dkent1@tufts-nemc.org; Alawi Alsheikh-Ali - AAlsheikh-Ali@tufts-nemc.org; Rodney A Hayward - rhayward@umich.edu

* Corresponding author

Abstract

It has been demonstrated that patients enrolled in clinical trials frequently have a large degree of variation in their baseline risk for the outcome of interest. Thus, some have suggested that clinical trial results should routinely be stratified by outcome risk using risk models, since the summary results may otherwise be misleading. However, variation in competing risk is another dimension of risk heterogeneity that may also underlie treatment effect heterogeneity. Understanding the effects of competing risk heterogeneity may be especially important for pragmatic comparative effectiveness trials, which seek to include traditionally excluded patients, such as the elderly or complex patients with multiple comorbidities. Indeed, the observed effect of an intervention is dependent on the ratio of outcome risk to competing risk, and these risks—which may or may not be correlated—may vary considerably in patients enrolled in a trial. Further, the effects of competing risk on treatment effect heterogeneity can be amplified by even a small degree of treatment related harm. Stratification of trial results along both the competing and the outcome risk dimensions may be necessary if pragmatic comparative effectiveness trials are to provide the clinically useful information their advocates intend.

Introduction

Recent commentaries have highlighted several fundamental limitations of clinical trials in providing an evidence-base for medical practice. It has been pointed out that many patients seen in routine clinical practice, particularly older and complex patients with multiple comorbid conditions, are excluded from clinical trials [1,2]. To address this, there has been a call for pragmatic comparative effectiveness trials with broader inclusion criteria, with the goal of enrolling a diverse patient population more representative of patients seen in routine clinical practice [3]. However, other commentaries have highlighted another limitation in clinical trials: substantial treatment-effect heterogeneity within trials often makes the overall summary result difficult to interpret and apply [4-7]. Enrolling a greater diversity of patients will increase this within trial heterogeneity. Thus, while some argue for broader inclusion criteria to make results more "generalizable", increasing patient heterogeneity yields overall
results that are more likely to be uninformative or even misleading.

While a consensus has yet to fully emerge on how best to deal with treatment-effect heterogeneity, the limitations of conventional subgroup analyses are well-appreciated. Since patients have multiple attributes that might affect the risks and benefits of an intervention – they are male or female, young or old, with or without diabetes, have a high or low blood pressure, blood count, cholesterol, urinary protein excretion, ejection fraction, etc. – it is statistically impractical to consider each potentially important attribute in a one at a time manner [6,7]. It has therefore been suggested that patient characteristic be combined by risk models that describe fundamental dimensions of risk likely to underpin treatment-effect heterogeneity [6,9-11].

Prior work has demonstrated that variation in outcome-risk (i.e. a patient's baseline risk of having the outcome of interest) is a fundamental determinant of the opportunity for treatment benefit, and of the risk-benefit trade-offs when there is any treatment-related harm [6,9-11]. Because variation in outcome-risk among patients enrolled in clinical trials is ubiquitous, frequently large and typically skewed, a relatively small subgroup of high-risk patients often account for most trial outcomes and have a disproportionate influence on overall trial results [12]. Indeed, the summary result of a clinical trial might not even accurately reflect the tested intervention's treatment-effect in a typical patient within the trial [6,12]. Because of this, and because outcome-risk variation can often be well-described with a simple multivariate risk model, routine stratification of trial results by outcome-risk has been recommended [6,9-11,13]. In addition to outcome-risk, it is also recognized that, for treatments with serious and non-rare adverse effects (e.g. surgery or fibrinolytic therapy), individual patient variation in vulnerability to treatment-related harm can give rise to important treatment-effect heterogeneity; thus, it may in some circumstances be appropriate to stratify patients based on their risk of treatment-related harm [6,9,14,15]. However, another dimension of risk heterogeneity from which clinically significant differences in treatment-effect may emerge is relatively neglected and may be of particular importance for comparative effectiveness trials: variation in competing risk. Competing risk is the risk of an event that interferes with the probability of experiencing the disease-specific outcome of interest [16]. It is not merely a statistical issue affecting Kaplan-Meier [16] or sample size [17] estimates, but a clinical issue especially important when considering treatments in older or complex patients with multiple comorbidities for whom competing events may limit the likelihood of treatment benefit. This paper considers how – even for treatments with uniform treatment efficacy-understanding the complex interplay between baseline risk, treatment-related harm and competing risk is important in making good individual-patient recommendations and decisions, and how analyzing the effects of competing and outcome risks in clinical trials – normally obscured by overall trial results – may better inform clinical decision-making.

**Table 1: Interactions between baseline risk, treatment-related harm (Rx-harm) and competing risk (CR) when chemotherapy reduces breast cancer mortality by 15%**

| Risk of Breast CA Death | No Rx-harm or CR | Rx-harm (1.5% absolute rate) but no CR | Rx-harm & CR is: |
|-------------------------|------------------|---------------------------------------|------------------|
| (10-yr CA Mortality)    |                  | Absolute Risk Reduction/Number Needed to Treat |
| Low (10%)               | .015/67          | 0/8                                   | -.002/667§       | -.004/-267§ |
| Moderate (25%)          | .038/27          | .023/44                               | .019/53          | .013/76   |
| High (50%)              | .075/13          | .060/17                               | .053/19          | .041/24   |

§ The negative sign denotes instances in which chemotherapy does net harm, meaning that the statistics represent absolute risk increase and number need to harm.

**Discussion**

**When Competing Risk is Uncorrelated with Outcome-Risk**

To illustrate the interplay between outcome-risk and competing risk, consider the use of adjuvant chemotherapy for breast cancer. Adjuvant chemotherapy reduces the risk of breast cancer death in both node positive and node negative cancers [18]. Since the treatment carries a non-trivial risk of serious complications, evidence-based guidelines often strongly recommend that the patient's risk of cancer recurrence and death (based upon cancer grade and stage) play an important role in determining who should receive chemotherapy [19]. However, breast cancer frequently behaves like a chronic disease, with events occurring over a decade or more, and there is tremendous variation in the risk of both breast cancer and non-breast cancer death across the breast cancer population.

Table 1 shows how the benefit of chemotherapy will vary according to both the baseline risk from the cancer itself and the degree of competing risk for mortality, even
assuming a constant relative effect of treatment for all patients (relative risk reduction = 15%) and a constant absolute rate of serious treatment-related harm (15 events over 10 years per 1000 people treated). Based on these assumptions, and estimated survival rates for stages 1, 2 and 3 breast cancer [20], consistent with our previous work [6,9,10], those who are at very high risk of dying from breast cancer usually benefit substantially despite the risks of treatment-related harm, and not surprisingly, even when there is substantial competing risk. This is because when the baseline breast cancer risk is high the disease-specific risk overwhelms competing risks from comorbid illnesses, and an efficacious treatment produces a large amount of absolute benefit, far outweighing the small risk of treatment-related harm (see Table 1).

For patients with a more favorable prognosis, however, the absolute amount of benefit of the same treatment is much more modest, such that treatment-related harm and competing risk can greatly attenuate or reverse the net benefits of treatment. For example, a patient with a non-trivial 10% percent breast cancer mortality risk would appear to be a good candidate in the absence of other risks (number needed to treat [NNT] = 67); however, a small treatment-related risk would nullify their potential benefit and the presence of relatively modest competing risks would cause the treatment to result in net harm. Even for patients with a substantial 25% risk of breast cancer death in the absence of competing risks, a high rate of competing risk results in a greatly attenuated treatment-effect; the NNT increases from 44 (with no competing risks) to 267 (with a 50% 10-year competing risk of mortality). If patients with high competing risk also had twice the normal risk of treatment-related harm (and the risk of treatment-related harm is often influence by comorbid illness) then chemotherapy would result in net harm (number needed to harm [NNH] = 91).

Note that a 10-year competing risk of mortality of 50% is not extreme for breast cancer. Approximately a third of breast cancer patients are over the age of 70. The 10-year risk of competing mortality would be approximately 50% for a 70 year old at only slightly higher than average risk [21]. Indeed, a median age breast cancer patient (age 61) with asymptomatic class I CHF would also have approximately a 50% 10-year risk of competing mortality [22].

Table 1 demonstrates that the overall measured effectiveness of adjuvant chemotherapy in a clinical trial depends on the distribution of the competing and outcome risks of trial enrollees. By excluding older patient or those with comorbidities, [23,24] oncology trials enhance their likelihood of detecting treatment benefit, but their results are directly applicable only to patients with low competing risks. While enrolling older patients, and patients with multiple comorbidities would attenuate the treatment-effect in the summary results, it would still not provide the clinically useful knowledge about who to treat unless analyses included risk-based stratification, incorporating both competing and outcome risk.

When Competing Risk is Correlated with Outcome-Risk

While the presence or absence of comorbidities should not substantially alter the likelihood of a more aggressive versus a more indolent form of cancer, in many circumstances, competing risk can be highly correlated with the disease-specific outcome-risk. For example, an implantable cardiac defibrillator (ICD) would be of most benefit in patients with a high risk for sudden cardiac death (SCD) but little risk for death from other causes [25], since implanting these devices (costly and not risk-free) in patients destined to die from non-arrhythmic causes is highly undesirable. However, the criteria used to identify eligible patients at high risk for sudden cardiac death (SCD), namely a left ventricular ejection fraction of 35% or less, also identifies patients at risk of cardiac death from pump failure [26,27]. Separating these risks has proven difficult, as factors that predict mortality from SCD also usually predict non-SCD mortality [28].

The Seattle Heart Failure Model (SHFM), developed on a database of pooled clinical trials consisting of 10,538 ambulatory patients with heart failure [29], predicts total mortality in patients with congestive heart failure based on easily obtainable clinical variables. Both SCD and pump failure death substantially increase across risk strata [30]. However, the ratio between SCD and pump failure death dramatically decrease in higher risk compared to lower risk patients; low risk patients with risk scores of zero have roughly a 7 to 1 SCD to pump failure death ratio, while this ratio was 1 to 2 in patients with risk scores of 4 [30].

As Figure 1 demonstrates, these different ratios across risk strata can have important effects on the measured effectiveness of ICDs. As mortality risk increases, the relative risk reduction of ICDs dramatically decreases. This is because the relative risk reduction is inversely related to the ratio of preventable disease-specific (SCD) to non-preventable competing risk of mortality. More surprisingly, the absolute risk reduction across risk strata is described by a non-linear, inverted U-shaped function (Figure 1b). Intermediate risk patients are most likely to benefit, as low risk patients are unlikely to have an arrhythmic death even in the absence of treatment while benefit in the highest risk group is limited by the high rate of non-arrhythmic death. Folding in treatment-related harm could further amplify this treatment-effect heterogeneity, particularly if sicker patients were more prone to ICD-induced adverse events.
Indeed, the phenomena depicted in Figure 1 is consistent with a risk-stratified analysis of the Multicenter Automatic Defibrillator Implantation Trial (MADIT)-II [31,32]. While overall MADIT-II found a 31% reduction in all-cause two-year mortality associated with the ICD [31], the ICD did not appear to reduce all-cause mortality in either the very low risk group (defined as no more than 1 of multiple risk factors) or a very high risk group of patients (defined as those with blood urea nitrogen ≥ 50 mg/dL or serum creatinine ≥ 2.5 mg/dL) [32]. While the intermediate mortality risk group may represent the "sweet spot" for ICD therapy, more efficient targeting of these devices could presumably be achieved if risk factors for SCD that do not predict pump failure could be identified. Again, summary trial results would be quite dependent on the risk profile of the enrolled patients, and may be misleading in some circumstances.

Summary

As seen in the clinical conditions above, variation in competing risk can cause variation in treatment-effect. Even for treatments with a constant efficacy, the apparent relative risk reduction of an intervention is directly related to the ratio between disease-specific and competing risk. When these two risk dimensions are correlated, increasing outcome-risk may not uniformly lead to increasing benefit – particularly when the primary outcome is a combination of disease-specific and competing events. Even when outcome and competing risks are not correlated, understanding how to treat individual patients can depend on the interplay between competing and outcome risks, and these effects can be greatly amplified by even small amounts of treatment-related harm, especially when those with higher competing risk are also at greater risk of harm from treatment. Calls for large simple clinical trials [33] with broad inclusion criteria, including older or complex patients, designed to provide results generalizable to "real-world" patients [3], have generally ignored the fact that reporting a summary treatment-effect based on the arithmetic mean across all patients may at times be misleading. The application of the results of meta-analysis without a careful consideration of clinical heterogeneity is also problematic [34]. Since treatment benefit depends on the ratio between competing and outcome risks, it may be necessary to stratify these real world effectiveness trials along these two important risk dimensions, as in Table 1, or to account for these risks using appropriate multivariate models.

Some might point out that thoughtful and experienced clinicians attempt to do this in clinical practice, such that overall clinical trial results can still be customized at the bedside. Doctors specializing in prostate cancer, for example, are known to apply the so-called 10-year rule [35], an implicit assessment based on age and co-morbidity, whereby aggressive therapy might be offered to patients...
likely to live long enough to benefit. However, physicians in general (and oncologists specifically) are prognostically inaccurate and systematically over-optimistic, when estimating overall life expectancy [36,37]. In addition to being inaccurate, implicit clinical judgment is an inadequate basis for clinical practice in an era where decisions are expected to conform to guidelines and will be evaluated based on performance measures. While tools which help formally assess comorbidities and competing risks may be helpful [38], considering and comparing patient-specific competing risks to patient-specific disease-specific risks adds a dimension of complexity likely to render simple bedside heuristics inadequate since these may be determined by different or similar factors, and particularly where the time course for benefits and harms of therapies can vary. Further, when results of clinical trials themselves are aggregated across patients with greatly varying disease-specific and competing risks, the underlying treatment-effect that should be incorporated in to any decisional framework at the bedside may be totally obscure.

Conclusion
In order to build an evidence-base that can support guidelines for patients who have multiple diseases simultaneously, relaxing clinical trial eligibility criteria to include older and complex patient must be accompanied by analyses that examine how a treatment’s net benefit varies by an individual’s disease-specific risk, chance of treatment-related harm, and competing risks. Research is needed to develop and test reliable ways to capture competing risk for different conditions [35,39], to develop sound methodologies to examine treatment-effects across multiple dimensions of risk, to develop a consensus to standardize analytic approaches and to identify which circumstances and clinical conditions these more complex analytic approaches might be justified and necessary.

Competing interests
The authors declare that they have no competing interests.

Authors’ contributions
All authors contributed equally to the development of the concepts in this manuscript. DMK had the initial idea for the manuscript and prepared the first draft based on contributions from all authors. All authors participated in manuscript revision.

Acknowledgements
Dr Kent was supported in by a methods development grant from Pfizer, Inc. Dr Alawi Alsheikh-Ali was supported by a Tufts-NEMC-Pfizer Scholars Award. The sponsor had no role in the development of the concepts or in the preparation of the manuscript or in the decision to submit for publication.

References
1. Gurwitz JH, Col NF, Avorn J: The exclusion of the elderly and women from clinical trials in acute myocardial infarction. JAMA 1992, 268:1417-1422.
2. Cameron HJ, Williams BO: Clinical trials in the elderly should we do more? Drugs Aging 1996, 9:307-310.
3. Tunis SR, Stryer DB, Clancy CM: Practical clinical trials: increasing the value of clinical research for decision making in clinical and health policy. JAMA 1996, 276:1624-1632.
4. Rothwell PM: Can overviews of clinical trials be applied to all patients? Lancet 1995, 345:1616-1619.
5. Kravitz RL, Duan N, Braslow J: Evidence-based medicine, heterogeneity of treatment-effects, and the trouble with averages. Milbank Q 2004, 82:661-687.
6. Kent DM, Hayward RA: Subgroup Analyses in Randomized Controlled Trials: Quantifying the Risks of False Positives and False-Negatives. Health Technol Assess 2001, 5:1-68.
7. Kent D, Hayward R: When averages hide individual differences in clinical trials: analyzing the results of clinical trials to expose individual patient’s risks might help doctors make better treatment decisions. Am Sci 2007, 95(1):50(9).
8. Brookes ST, Whitley E, Peters Tj, Mulheran PA, Eggar M, Davey Smith G: Subgroup Analyses in Randomized Controlled Trials: Quantifying the Risks of False Positives and False-Negatives. Health Technol Assess 2001, 5:1-68.
9. Hayward RA, Kent DM, Vijan S, Hofer TP: Multivariable risk prediction can greatly enhance the statistical power of clinical trial subgroup analysis. BMC Med Res Methodol 2006, 6:18.
10. Hayward RA, Kent DM, Vijan S, Hofer TP: Reporting clinical trial results to inform providers, payers, and consumers. Health Aff (Millwood) 2005, 24:1571-1581.
11. Rothwell PM, Mehta Z, Howard SC, Guntikov SA, Warlow CP: Treating individuals 3: from subgroups to individuals: general principles and the example of carotid endarterectomy. Lancet 2005, 365:256-265.
12. Ioannidis JP, Lau J: The impact of high-risk patients on the results of clinical trials. J Clin Epidemiol 1997, 50:1089-1098.
13. Ioannidis JP, Lau J: Heterogeneity of the baseline risk within patient populations of clinical trials: a proposed evaluation algorithm. Am J Epidemiol 1998, 148:1117-1126.
14. Rothwell PM, Warlow CP: Prediction of benefit from carotid endarterectomy in individual patients: a risk-modelling study. European Carotid Surgery Trialists’ Collaborative Group. Lancet 1999, 353:2105-2110.
15. Kent DM, Ruthazer R, Selker HP: Are some patients likely to benefit from recombinant tissue-type plasminogen activator for acute ischemic stroke even beyond 3 hours from symptom onset? Stroke 2003, 34:464-467.
16. Gooey TA, Leisenring W, Crowley J, Storer BE: Estimation of failure probabilities in the presence of competing risks: new representations of old estimators. Stat Med 1999, 18:695-706.
17. Maki E: Power and sample size considerations in clinical trials with competing risk endpoints. Pharm Stat 2006, 5:159-171.
18. Early Breast Cancer Trialists’ Collaborative Group (EBCTCG): Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet 2005, 365:1687-1717.
19. National Comprehensive Cancer Network (NCCN) guideline. [http://www.nccn.org/professionals/physician_gls/default.asp]. Accessed 20 November 2007.
20. National Cancer Institute Surveillance Epidemiology and End Results (2008) [http://seer.cancer.gov/faststats/sites.php?site=BreastCancer&stat=Survival].
21. National Vital Statistics Report, United States Life Tables 2003, revised March 28, 2007 [http://www.cdc.gov/nchs/data/nvss/nvss54/nvss54_14.pdf]. Accessed 6 December 2007.
22. Seattle Failure Model [http://depts.washington.edu/shfm/index.php].
23. Hutchins L, Unger JM, Crowley JJ, Coltman CA Jr, Albain KS: Under-representation of patients 65 years of age or older in cancer-treatment trials. N Engl J Med 1999, 341:2061-2067.
24. Murtby V, Krumholz HM, Gross CP: Participation in cancer clinical trials: race-, sex-, and age-based disparities. JAMA 2004, 291:2720-2726.
25. Owens DK, Sanders GD, Heidenreich PA, McDonald KM, Hlatky MA: Effect of risk stratification on cost-effectiveness of the
implantable cardioverter defibrillator. Am Heart J 2002, 144:440-448.
26. Zipes DP, Camm AJ, Borggrefe M, Buxton AE, Chaitem B, Fromer M, Gregoratos G, Klein G, Moss AJ, Myerburg RJ, Priori SG, Quinones MA, Roeden DM, Silka MJ, Tracy C, Smith SC Jr, Jacobs AK, Adams CD, Antman EM, Anderson JL, Hunt SA, Halperin JL, Nishimura R, Ornato JP, Page RL, Riegel B, Priori SG, Blanc JJ, Budaj A, Camm AJ, Dean V, Deckers JW, Despres C, Dickstein K, Lekakis J, McGregor K, Metra M, Morais J, Osterspey A, Tamargo JL, Zamorano JL. American College of Cardiology; American Heart Association Task Force; European Society of Cardiology Committee for Practice Guidelines: ACC/AHA/ESC 2006 guidelines for management of patients with ventricular arrhythmias and the prevention of sudden cardiac death: a report of the American College of Cardiology/American Heart Association Task Force and the European Society of Cardiology Committee for Practice Guidelines (Writing Committee to Develop Guidelines for Management of Patients With Ventricular Arrhythmias and the Prevention of Sudden Cardiac Death). J Am Coll Cardiol 2006, 48:e247-346.
27. Buxton AE, Lee KL, Hafley GE, Pires LA, Fisher JD, Gold MR, Josephson ME, Lehmann MH, Prystowsky EN, for the MUSTT Investigators: Limitations of ejection fraction for prediction of sudden death risk in patients with coronary artery disease: lessons from the MUSTT study. J Am Coll Cardiol 2007, 50:1150-1157.
28. Every N, Hallstrom A, McDonald KM, Parsons L, Thom D, Weaver D, Hickey MA: Risk of sudden versus nonsudden cardiac death in patients with coronary artery disease. Am Heart J 2002, 144:390-396.
29. Levy WC, Mozaffarian D, Linker DT, Sutrathar SC, Anker SD, et al.: The Seattle Heart Failure Model: prediction of survival in heart failure. Circulation 2006, 113:1424-1433.
30. Mozaffarian D, Anker SD, Anand I, Linker DT, Sullivan MD, Cleland JG, Carson PE, Maggioni AP, Mann DL, Pocock B, Poole-Wilson PA, Levy WC: Prediction of mode of death in heart failure: the Seattle Heart Failure Model. Circulation 2007, 116:392-398.
31. Moss AJ, Zareba W, Hall WJ, Klein H, Wilber DJ, Cannom DS, Daubert JP, Higgins SL, Brown MW, Andrews ML. Multicenter Automatic Defibrillator Implantation Trial II Investigators: Prophylactic implantation of a defibrillator in patients with myocardial infarction and reduced ejection fraction. N Engl J Med 2002, 346:877-883.
32. Goldenberg I, Vyas AK, Hall WJ, Moss AJ, Wang H, He H, Zareba W, McNitt S, Andrews ML, for the MADIT-II Investigators: Risk stratification for primary implantation of a cardioverter defibrillator in patients with ischemic left ventricular dysfunction. J Am Coll Cardiol 2008, 51:288-296.
33. Yusuf S, Collins R, Peto R: Why do we need some large, simple randomized trials? Stat Med 1984, 3:409-420.
34. Lau J, Ioannidis JP, Schmid CH: Summing up evidence: one answer is not always enough. Lancet 1998, 351(9096):123-7.
35. Levey MS, Greenfield S, Elkin EP, Lubeck DP, Brower JM, Kaplan SH: Assessment of prognosis with the total illness burden index for prostate cancer: aiding clinicians in treatment choice. Cancer 2007, 109:1777-1783.
36. Christakis NA, Lamont EB: Extent and determinants of error in physicians’ prognoses in terminally ill patients: prospective cohort study. BMJ 2000, 320:469-472.
37. Glare P, Virik K, Jones M, Hudson M, Eychmuller S: A systematic review of physicians’ survival predictions in terminally ill cancer patients. BMJ 2003, 327:195.
38. Braithwaite RS, Concauto J, Chang CC, Roberts MS, Justice AC: A framework for tailoring clinical guidelines to comorbidity at the point of care. Arch Intern Med 2007, 167:2361-2365.
39. Charlson ME, Pompei P, Alyks KL, MacKenzie CR: A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987, 40:373-383.