

# Analyzing composite outcomes in cardiovascular studies: traditional Cox proportional hazards versus quality-of-life-adjusted survival approaches

DEAN T EURICH, SUMIT R MAJUMDAR, FINLAY A McALISTER, ROSS T TSUYUKI, YUTAKA YASUI, JEFFREY A JOHNSON

## ABSTRACT

**Background:** Composite outcomes that weight each component equally are commonly used to study treatment effects. We hypothesized that each component of a composite outcome would differentially affect patients' overall health-related quality of life (HRQL).

**Methods:** We tested our hypothesis using data from 2 published clinical studies of treatment for heart failure, one comparing metformin and sulfonylurea and the other comparing digoxin and placebo. We applied the quality-adjusted survival (QAS) approach, which incorporates HRQL data to accommodate differential weights for 2 components (in this analysis, death or admission to hospital) of a commonly used composite end point. For each of the 2 studies, the composite outcome was partitioned into its components, to which utility weights derived from the literature were assigned. Total QAS time determined for each treatment by the QAS analysis was compared with the results from traditional survival analyses based on Cox proportional hazards regression.

**Results:** In the observational study of metformin in heart failure, the risk of the composite outcome of death or admission to hospital was lower for those receiving metformin therapy than for those who received sulfonylurea (event rate 160 [77%] v. 658 [85%]; hazard ratio [HR] 0.83, 95% confidence interval [CI] 0.70–0.99). With traditional survival analysis, the net gain was 0.82 years (95% CI 0.26–1.37), whereas the difference in QAS time was less, at 0.54 years (95% CI 0.20–0.89). In the randomized trial of digoxin therapy, the risk of the composite outcome was lower for those receiving the intervention than for those receiving placebo (event rate 1291 [38%] v. 1041 [31%]; HR 0.75, 95% CI 0.69–0.82). With traditional survival analysis, the net gain was 0.06 years (95% CI 0.02–0.16), whereas the difference in QAS time was greater, at 0.11 years (95% CI 0.06–0.16).

**Interpretation:** Studies that assume equal weights for the components of composite outcomes may overestimate or underestimate treatment effects. By incorporating HRQL into survival analyses, the impact of the various components of the outcome can be assessed more directly.

**Dean T Eurich**, BSP, PhD, is assistant professor, Department of Public Health Sciences, School of Public Health, University of Alberta, Edmonton, Alberta, Canada. **Sumit R Majumdar**, MD, MPH, is associate professor, Department of Medicine, Faculty of Medicine and Dentistry, University of Alberta Hospital. **Finlay A McAlister**, MD, MSc, is associate professor, Department of Medicine, Faculty of Medicine and Dentistry, University of Alberta Hospital. **Ross T Tsuyuki**, BSc(Pharm), PharmD, MSc, is professor, Faculty of Medicine and Dentistry, University of Alberta Hospital. **Yutaka Yasui**, PhD, is professor, Department of Public Health Sciences, School of Public Health, University of Alberta. **Jeffrey A Johnson**, PhD, is professor, Department of Public Health Sciences, School of Public Health, University of Alberta.

**Competing interests:** None declared.

**Funding:** Dean Eurich held a Doctoral Research award with the Canadian Institutes for Health Research (CIHR) and is currently a Population Health Investigator with the Alberta Heritage Foundation for Medical Research (AHFMR) and a New Investigator with the CIHR. Sumit Majumdar and Finlay McAlister are Health Scholars with the AHFMR. Finlay McAlister and Ross Tsuyuki hold the Merck Frosst/Aventis Chair in Patient Health Management. Yutaka Yasui holds a Canada Research Chair in Biostatistics and Epidemiologic Methods and is a Health Scholar supported by the AHFMR. Jeffrey Johnson holds a Canada Research Chair in Diabetes Health Outcomes, is a Health Scholar supported by the AHFMR and is chair of a New Emerging Team (NET) grant to the Alliance for Canadian Health Outcomes Research in Diabetes (ACHORD). The ACHORD NET grant is sponsored by the Canadian Diabetes Association, the Heart and Stroke Foundation of Canada, the Kidney Foundation of Canada, the CIHR Institute of Nutrition, Metabolism and Diabetes, and the CIHR Institute of Circulatory and Respiratory Health. These study sponsors did not play any role in the design or conduct of the study; collection, analysis or interpretation of the data; writing of the report; or the decision to submit the article for publication.

**Correspondence:** Dr. Dean Eurich, Department of Public Health Sciences, School of Public Health 2-040 HRIF East, University of Alberta, Edmonton AB T6G 2E1; tel: 780 492-6333; fax: 780 492-7455; deurich@ualberta.ca

**C**OMPOSITE OUTCOMES ARE COMMON IN CLINICAL research, especially in trials of cardiovascular disease. The potential advantages of composite outcomes include higher event rates, better power and statistical efficiency, and integration of clinically important events into a single quantifiable outcome. However, several concerns have been raised, particularly regarding the interpretation of results when the components of the composite outcome affect patients' health differentially or are associated with competing risks.<sup>1</sup> For example, the composite outcome of death or admission to hospital is often used in studies of heart failure, but it could be difficult to compare the effects of 2 treatments if one treatment reduces mortality but increases admissions and the other treatment has no effect on mortality but reduces admissions.

One concern with composite outcomes is the assumption that each component of the outcome (i.e., each health state) is equally important. Although this assumption is valid in certain cases, more often the individual components of a composite outcome will affect patients' overall quality of life differently. One way to address this situation is to account for potential differences in health states by assigning differential weights to the components. In many instances where this has been attempted, the weights assigned have been based on expert opinion.<sup>2-5</sup> However, the appropriateness of these expert-derived weights is controversial, which has probably contributed to the limited adoption of weighted methods. Because patients ultimately experience the events in question, it seems reasonable to incorporate their perspectives when weighting composite outcomes in clinical research.

One method for assigning weights that reflects the patient's perspective is the use of health-related quality of life (HRQL). If HRQL is incorporated into survival analyses, index measures (i.e., utilities) can be used to adjust for the unequal impact of health states used in composite outcomes, thereby providing a "weighted" survival outcome that accounts for different degrees of quality and quantity of life.<sup>6</sup> Furthermore, incorporation of HRQL into survival analysis would align clinical research with the methods of economic evaluations, for which cost utility analyses (i.e., cost per quality-adjusted life year)<sup>7</sup> are recommended, thereby improving the assessment of health care interventions. This concept of using HRQL to adjust survival analyses for different health outcomes has been used in oncology,<sup>6</sup> where it is referred to as quality-adjusted survival (QAS) analysis.<sup>8</sup> To our knowledge, it has not been employed in cardiovascular disease

or specifically for patients with heart failure. Here, we apply this method to the analysis of cardiovascular disease, using heart failure as our base example, since many studies of this condition evaluate the composite outcome of death or admission to hospital as the primary end point. Most patients and clinicians, however, would consider these disparate end points. Therefore, analyses of such composite outcomes may be improved with the QAS approach.

## Methods

**Overview.** We applied the QAS technique as originally proposed by Cole and associates<sup>8</sup> to the commonly reported composite outcome of death or admission to hospital using 2 previously published studies: an observational study of antidiabetic therapies in patients with diabetes and heart failure (the metformin study)<sup>9</sup> and a randomized controlled trial comparing digoxin with placebo in patients with heart failure (the Digitalis Investigation Group [DIG] study).<sup>10</sup> A sample data set and SAS code have been provided with this article to illustrate the application of the QAS technique (see online Appendices A and B). Data for the DIG study were obtained from the National Heart, Lung, and Blood Institute, whereas the metformin study data were previously acquired by the investigators from Saskatchewan Health.

**Data sources.** The data sources and population for the study of metformin use in heart failure have been described in detail elsewhere.<sup>9</sup> Briefly, 1833 patients with diabetes newly treated with oral antidiabetic agents and incident heart failure were identified using the administrative databases of Saskatchewan Health. The patients were categorized into 3 mutually exclusive groups: 773 (42%) were treated with sulfonylurea therapy alone, 208 (11%) with metformin alone and 852 (47%) with combinations of sulfonylurea and metformin. For the purposes of the current article, only patients who received either metformin or sulfonylurea monotherapy were included. All patients were followed prospectively until death, termination of Saskatchewan Health coverage or December 31, 1999. Maximum follow-up was 9 years.

With standard Cox proportional hazards regression techniques, after adjustment for potentially confounding variables (i.e., age, sex, a modified chronic disease score,<sup>11,12</sup> therapies known to affect heart failure outcomes [angiotensin-converting enzyme inhibitors, angiotensin II blockers,  $\beta$ -blockers, antiplatelet agents, nitrates, lipid-lowering therapies, antiarrhythmic agents and spironolactone] and total physician visits before

diagnosis of heart failure), there was a reduction in the hazard of events in favour of the metformin group compared to sulfonylurea therapy for both all-cause mortality (69 [33%] v. 404 [52%]; hazard ratio [HR] 0.70, 95% confidence interval [CI] 0.54–0.91) and the composite outcome of “all-cause mortality or all-cause hospitalization” (160 [77%] v. 658 [85%]; HR 0.83, 95% CI 0.70–0.99).<sup>9</sup>

The rationale, design and results of the DIG study have been reported in detail elsewhere.<sup>10</sup> A total of 6800 patients with heart failure and left ventricular ejection fraction of 0.45% or less were randomly assigned to receive either digoxin or placebo. After an average follow-up of 37 months, there was no difference between the study groups with respect to the primary outcome of all-cause mortality (1194 [35%] in the placebo group v. 1181 [35%] in the digoxin group; HR 0.99, 95% CI 0.91–1.07). There was a trend toward lower risk of death related to heart failure in the digoxin group than in the placebo group (394 [12%] v. 449 [13%]; HR 0.88, 95% CI 0.77–1.01). In addition, the risk for the composite outcome of death due to worsening heart failure or admission to hospital related to that diagnosis was lower in the digoxin group (1041 [31%] v. 1291 [38%]; HR 0.75, 95% CI 0.69–0.82).<sup>10</sup>

**Weighted composite outcome.** We applied the QAS analysis to the composite outcome of “all-cause mortality or all-cause hospitalization” in the metformin study and to the composite outcome of “heart failure-specific mortality or hospitalization” in the DIG study. In both of these case studies, we considered 3 successive stages through which a patient might transition: state H<sub>1</sub>, representing the patient’s initial health state before admission to hospital, death or censoring at the end of the study (i.e., the mean event-free survival time corresponding to the area under the composite outcome curve); state H<sub>2</sub>, representing the patient’s health state after a hospital stay until either death or censoring at the end of the study (i.e., area between the composite outcome survival curve and the mortality survival curve); and state H<sub>3</sub>, death (no time is associated with this health state).

For the QAS analyses, each state was associated with a specific HRQL, represented by a utility coefficient.<sup>8</sup> We applied validated utility coefficients for patients with heart failure taken from the literature, since HRQL measurements had not been collected in either of the 2 case studies. For state H<sub>1</sub>, the utility coefficient was set at 0.81, based on Health Utilities Index Mark 3 (HUI3) scores observed for patients with heart disease and diabetes in the Canadian population.<sup>13</sup> This value is similar

to utility weights observed in patients with heart failure alone.<sup>14</sup> Among patients with heart failure, admission to hospital reduces HRQL by 30%,<sup>14</sup> resulting in a utility coefficient of 0.57 for state H<sub>2</sub>. By convention, the health state “dead” (H<sub>3</sub>) was assigned a utility score of 0.<sup>13</sup> For the traditional survival analysis, which assumes equal importance for each health state (i.e., H<sub>1</sub> and H<sub>2</sub>), utility equal to 1 (i.e., perfect health) was used. We used the HUI3 utility scores in our analyses because this is the only utility measure of which we are aware that has specifically evaluated the effect of hospital admission on HRQL for patients with heart failure. We acknowledge, however, that it has not necessarily been applied to patients with heart failure and diabetes; therefore, we were limited to the estimate for patients with heart disease and diabetes in the Canadian population.

**Analysis.** For the metformin cohort study, the transitional survival functions for each health state (i.e., all-cause death or all-cause admission to hospital) for the 2 treatment groups were estimated using Cox proportional hazards models (i.e., SAS code PHREG) with adjustments for the confounding covariates and the product-limit estimate of the baseline survival function.<sup>8</sup> The survival functions for the treatment groups were estimated using overall mean values for each covariate. In the DIG study, there were no significant differences in baseline characteristics between the patients in the digoxin and placebo groups because of the randomized study design.<sup>10</sup> Therefore, we estimated the survival functions using Cox proportional hazards models without adjustments for covariates. We checked proportional hazards assumptions using log–log plots and time interactions, with no violations noted for either study.

For each study, we first estimated the Cox proportional hazards regression survival function (i.e., survival curves) for the health state “dead” (H<sub>3</sub>) as the event of interest in each treatment group. We then estimated the survival function for the composite outcome (death or admission to hospital) as the event of interest for each group. The mean time spent in each health state was calculated by integrating the estimated survival function from 0 to the maximum limit of 7.8 years for the metformin study and 4.9 years for the DIG study. Time in state H<sub>1</sub> (initial health state of the patient) was estimated as the integrated survival time for the composite outcome (death or admission to hospital). Time in state H<sub>2</sub> (health state of the patient after a hospital stay) was calculated by integrating the overall survival time for death and subtracting the mean time spent in state H<sub>1</sub>. QAS time in

each health state was then calculated by multiplying the mean time spent in each health state by the respective utility coefficient.<sup>8</sup> Summation of these quality-adjusted times provided an estimate of the overall QAS time during the study period for each treatment group.

To generate estimates of the variability of the mean QAS time for the treatment groups, we used 500 bootstrap samples for each study, with the 95% CI and corresponding *p* values calculated according to the percentile method,<sup>15</sup> whereby the QAS estimates spanning the 2.5 and 97.5 percentiles of the bootstrapped distribution represented the 95% CI.

Sensitivity analyses were performed to evaluate the effect of different utility coefficients for the health states on the results for both studies. The utility coefficients for health states H<sub>1</sub> and H<sub>2</sub> were increased or decreased by 0.03, which is considered a clinically important difference on the HUI3.<sup>16</sup> All analyses were conducted using SAS for Windows, version 9.1 (SAS Institute, Cary, N.C.).

## Results

**Metformin use in heart failure.** In the metformin study, sulfonylurea users spent a mean of 1.21 years in health state H<sub>1</sub> (survival without hospital admission) and 2.89 years in health state H<sub>2</sub> (after admission to hospital until either death or censoring at the end of follow-up) (Fig. 1A, Table 1). Conversely, metformin users spent a mean of 1.52 years in health state H<sub>1</sub> and 3.40 years in health state H<sub>2</sub> (Fig. 1B, Table 1). Using traditional survival analysis, which assumes that those who are alive are in perfect health (i.e., utility of 1 for health states H<sub>1</sub> and H<sub>2</sub>) until the time of death, we determined that

sulfonylurea users had a mean expected adjusted survival of 4.10 years, calculated as (1.21 years × utility of 1) + (2.89 years × utility of 1) and metformin users a mean expected adjusted survival of 4.92 years, calculated as (1.52 years × utility of 1) + (3.40 years × utility of 1). This translates into an average gain of 0.82 life-years (95% bootstrapped CI 0.26–1.37) associated with metformin use relative to sulfonylurea use. However, using the QAS approach, with an expected utility of 0.81 for health state H<sub>1</sub> and 0.57 for health state H<sub>2</sub>, we determined that sulfonylurea users had a mean expected QAS time of 2.63 years, calculated as (1.21 years × utility of 0.81) + (2.89 years × utility of 0.57) (Fig. 1A, Table 1), and metformin users a mean expected QAS time of 3.17 years, calculated as (1.52 years × utility of 0.81) + (3.40 years × utility of 0.57) (Fig. 1B, Table 1). Thus, metformin users had a significant net increase of 0.54 quality-adjusted life years (QALYs) relative to those using sulfonylurea (95% bootstrapped CI 0.20–0.89, *p* < 0.002). In this case, traditional survival analysis led to a 34% overestimation of potential benefits related to use of metformin in the treatment of patients with both diabetes and heart failure.

**DIG study.** In the DIG study, placebo users spent a mean of 3.32 years in health state H<sub>1</sub> (survival with no admission to hospital for heart failure) and 1.03 years in health state H<sub>2</sub> (after admission to hospital for heart failure until either death from heart failure or censoring at the end of follow-up) (Fig. 2A, Table 2). Conversely, digoxin users spent a mean 3.64 years in health state H<sub>1</sub> and 0.77 years in health state H<sub>2</sub> (Fig. 2B, Table 2). Assuming perfect health before death,

as in traditional survival analysis (i.e., utility of 1 for all health states), placebo users had a mean expected adjusted survival of 4.35 years, calculated as (3.32 years × utility of 1) + (1.03 years × utility of 1), whereas digoxin users had a mean expected adjusted survival of 4.41 years, calculated as (3.64 years × utility of 1) + (0.77 years × utility of 1). This translates into a net survival benefit of 0.06 years (95% bootstrapped CI 0.00–0.12) for

**Table 1: Integrated survival time for sulfonylurea and metformin groups according to health state**

Group	Survival, years (95% CI)*			
	In health state H <sub>1</sub>	In health state H <sub>2</sub>	Total	Quality-adjusted
Sulfonylurea	1.21 (1.06–1.32)	2.89 (2.65–3.17)	4.10 (3.82–4.41)	2.63 (2.45–2.82)
Metformin	1.52 (1.27–1.82)	3.40 (2.81–3.92)	4.92 (4.35–5.41)	3.17 (2.81–3.48)

CI = confidence interval, H<sub>1</sub> = patient's initial health state before admission to hospital, death or censoring at the end of the study, H<sub>2</sub> = patient's health state after a hospital admission until either death or censoring at the end of the study.

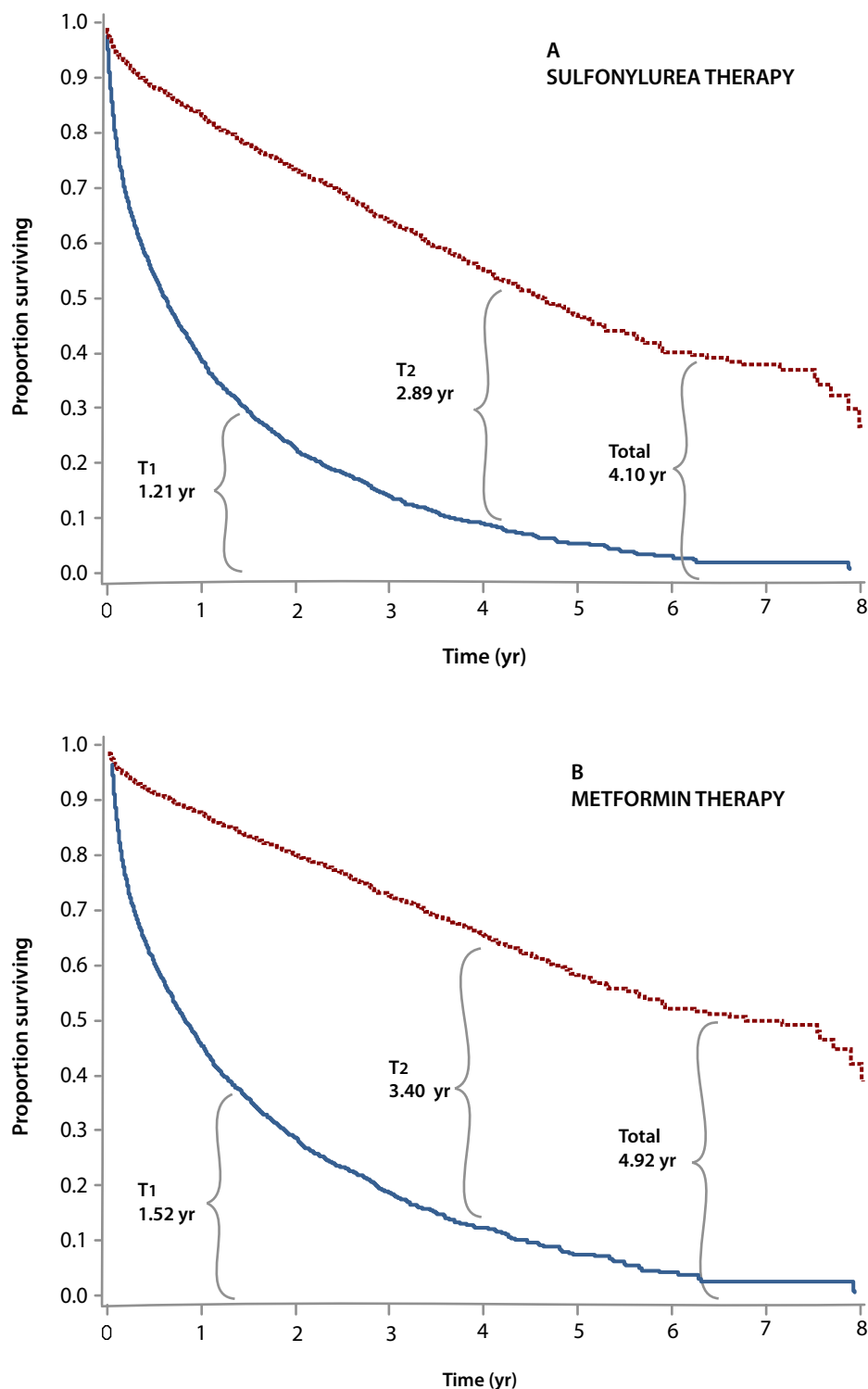
\*95% CI values generated from the bootstrap procedure (500 samples).

**Table 2: Integrated survival time for placebo and digoxin groups according to health state**

Group	Survival, years (95% CI)*			
	In health state H <sub>1</sub>	In health state H <sub>2</sub>	Total	Quality-adjusted
Placebo	3.32 (3.25–3.39)	1.03 (0.98–1.09)	4.35 (4.31–4.39)	3.28 (3.24–3.31)
Digoxin	3.64 (3.56–3.70)	0.77 (0.72–0.83)	4.41 (4.37–4.45)	3.39 (3.35–3.42)

CI = confidence interval, H<sub>1</sub> = patient's initial health state before admission to hospital, death or censoring at the end of the study, H<sub>2</sub> = patient's health state after a hospital admission until either death or censoring at the end of the study.

\*95% CI values generated from the bootstrap procedure (500 samples).



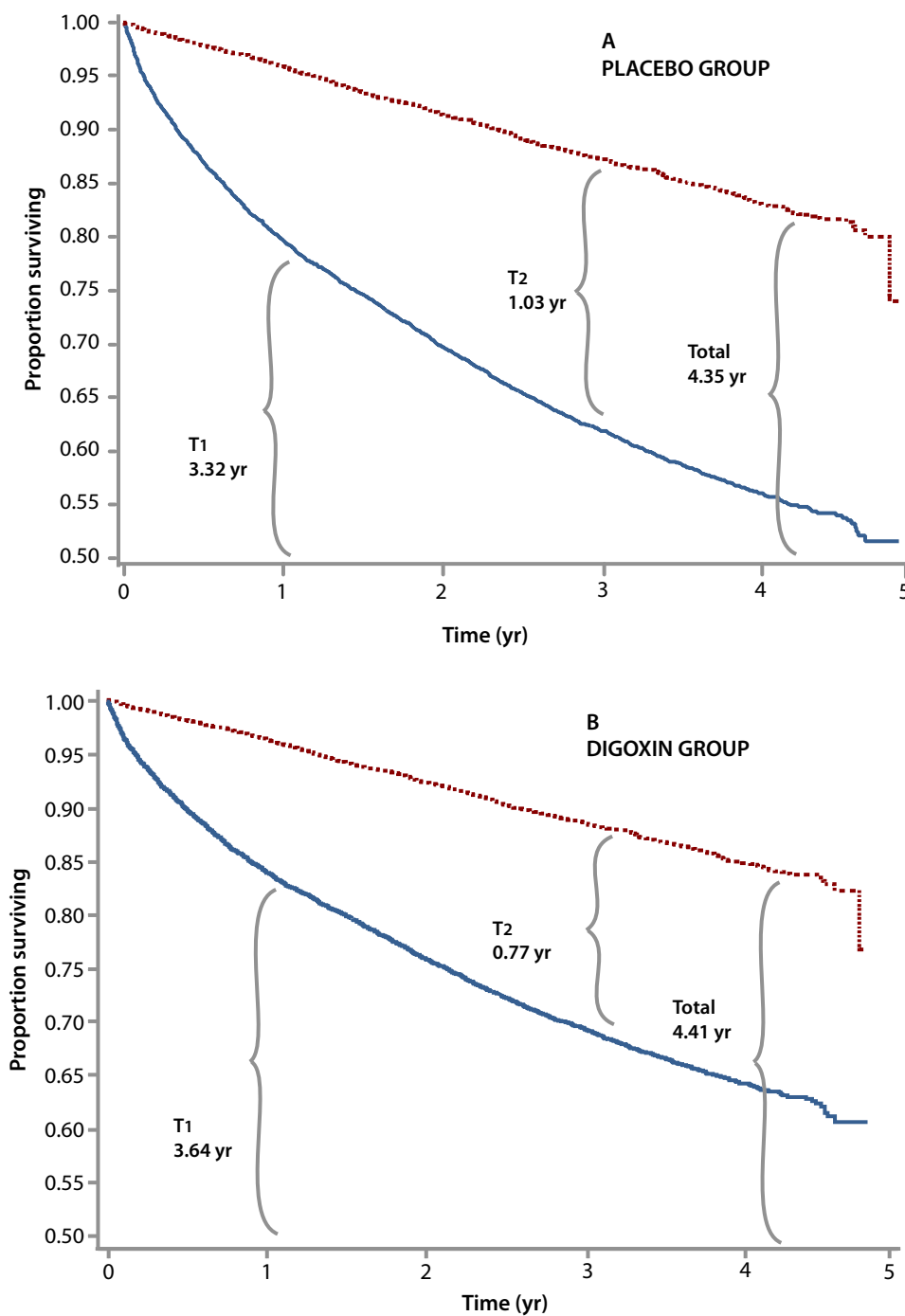
**Figure 1: Partitioned survival functions for all-cause mortality and all-cause admission to hospital in an observational study of antidiabetic agents in patients with heart failure.** T1 = time spent in health state 1. T2 = Time spent in health state 2. Dashed red line = all-cause mortality. Solid blue line = all-cause hospital admission or mortality.

digoxin users relative to placebo users. However, according to the QAS approach, placebo users had a mean expected QAS time of 3.28 years, calculated as (3.32 years  $\times$  utility of 0.81) + (1.03 years  $\times$  utility of 0.57), whereas the

digoxin users had a mean expected QAS time of 3.39 years, calculated as (3.64 years  $\times$  utility of 0.81) + (0.77 years  $\times$  utility of 0.57). Therefore, after taking into account the greater time spent before first admission for heart failure, the digoxin group had a net gain of 0.11 QALYs over the placebo group (95% bootstrapped CI 0.06–0.16,  $p < 0.001$ ). Furthermore, this estimate is 83% higher than the standard survival estimate based on equally weighted outcomes. In this case, traditional analytic methods led to an underestimation of the potential benefits related to using digoxin for treatment of patients with heart failure.

**Sensitivity analyses.** In the base-case analyses described above, we applied deterministic utility coefficients to health states taken from the literature. We also conducted a sensitivity analysis to assess the effect of varying the utility coefficients by a clinically important amount (0.03) on the QALY estimates for health state H<sub>1</sub> and H<sub>2</sub> for the metformin and DIG studies. Variations in the utility coefficients resulted in changes to both the individual QALYs calculated for each health state and the differences between study groups. All sensitivity analyses confirmed that the standard survival estimates overestimated the benefits of metformin by 30% to 35% in the observational metformin study and underestimated benefits of digoxin by 50% to 117% in the DIG study (see online Appendix

C for a full report of the sensitivity results). Additional analyses in which the HUI3 utilities were varied by an exceptionally large clinical change of 0.1 (i.e., change expected if a person with diabetes were to be cured of heart disease)<sup>13</sup>



**Figure 2: Partitioned survival functions for death due to worsening heart failure or admission to hospital related to that diagnosis in a randomized trial of digoxin.**

T1 = time spent in health state 1. T2 = Time spent in health state 2. Dashed red line = mortality related to heart failure. Solid blue line = heart-failure-related hospital admission or mortality.

consistently showed that traditional survival analyses overestimated treatment effects in the metformin study. Similarly, traditional survival analyses underestimated treatment effects in the digoxin study, with the exception of the analysis in which the HUI3 utility was decreased by 0.1 for health state H<sub>1</sub> and increased by 0.1 for health state H<sub>2</sub>.

## Interpretation

Using heart failure as an example, we have applied and demonstrated a method that adjusts survival to deal with the potentially unequal impact of the individual components of composite outcomes. If an estimate of the impact on patients' HRQL is incorporated into survival analyses, a weighted composite outcome may provide a representation of the benefits (or harms) associated with a therapy that is more pertinent to the individual patient. In the 2 case studies examined here, incorporation of HRQL into the survival analyses resulted in quality-adjusted survival estimates that ranged from 34% lower to 83% higher than those derived from traditional survival analyses using equally weighted components for the composite outcomes.

Composite outcomes are commonly reported, but the tacit assumption that each component is of equal importance to patients, providers and payers seems untenable and is rarely met. Increasing evidence suggests that components of a composite end point that have the greatest effect on patients, such as death, contribute little to the estimates of composite end points in cardiovascular-related research, whereas those components with the least impact on patients contribute the most.<sup>17</sup> Not surprisingly, the results of previous studies have suggested that the use of equally weighted components in composite end points may lead to biased conclusions and to overestimates of

treatment effects in cardiovascular-related trials.<sup>5</sup> Our results also illustrate that bias may occur, but in either direction.

As a result, researchers evaluating the merits of studies employing composite end points should be cautious of the influence of the individual components of such composite end points. It is clear that additional techniques,

like QAS, could improve the clinical interpretability of composite end points. Notably, the QAS approach does not influence either the estimated hazard ratios or the survival curves often reported in cardiovascular research; rather, it represents complementary information that allows for adjustment of the survival time, thereby “weighting” the survival outcome by incorporating different degrees of quality and quantity of life. Clearly, alternative analytical methods such as competing risk models or multistage models can also be employed to address the issue of multiple end points and competing events observed in composite outcomes. However, in contrast to the QAS approach, both of these methods weight each component equally, and neither places a “value” on the relative survival time gained or lost in association with the competing events.

Our primary intention in undertaking this study was to examine differences in the magnitude of the QAS estimates between the treatment groups (e.g., digoxin and placebo in the DIG study) and to determine how differential weighting of composite outcomes might change the interpretation of the data. The QAS approach generates survival time associated with “perfect health” rather than the overall survival generated by the traditional unweighted approach. For example, in the DIG trial, the risk of hospital admission or death due to heart failure was reported to be significantly lower in the digoxin group (HR 0.75, 95% CI 0.69–0.82).<sup>10</sup> However, the analysis that we have outlined above quantifies this benefit as a gain of only 0.06 life-years with the traditional unweighted approach and only 0.11 QALYs with the weighted QAS approach. This difference is of perhaps questionable clinical significance for the DIG study, but such may not always be the case. Conversely, the gain of 0.54 QALYs observed with metformin therapy is similar to that observed with the use of other proven efficacious therapies (relative to placebo) in patients with heart failure.<sup>18</sup> Furthermore, the difference in the QAS estimates for the unweighted and weighted approaches for metformin therapy (0.82 v. 0.54) would be similar to the HRQL effects of choosing not to prescribe angiotensin-converting enzyme inhibitors for patients with heart failure.<sup>18</sup> In these 2 examples, both therapies remained beneficial after adjustment for HRQL, but one could easily foresee clinical trials in which marginal benefits determined with unweighted composite outcomes might be nonsignificant if the weighted approach were used or vice versa. In addition, unlike most traditional survival analyses evaluating composite outcomes, the QAS approach is directly amenable to economic evaluations

(e.g., cost utility analysis incorporating QALYs). Thus, we believe that better-informed treatment decisions may result from incorporating measures of clinical impact into summary effect estimates.

Importantly, appropriate selection of the utility weights is required, as shown by the modest effect on the QAS estimates in our sensitivity analyses. Prospective collection of utilities in the setting of randomized controlled trials or cohort studies would permit incorporation of study-specific utilities into the QAS method. For example, collection of utility estimates at baseline, at the time of clinically important events and at regularly scheduled time points (e.g., every 6 months) might allow more complete evaluation of the effect of therapies on HRQL in clinical trials. Furthermore, since the periods of evaluation are consecutive, these time-specific health states could be easily incorporated into the QAS approach. In studies where prospective collection of utility estimates is not possible, carefully selected literature-based estimates (i.e., utility estimates from patient populations that are similar in terms of socio-demographic characteristics [age, sex, ethnicity], severity of disease, location of care [hospital, community] and health care systems) or threshold utility analyses for unknown utility estimates<sup>6,8</sup> may be used, but they should be justified and subject to sensitivity analyses such as we have illustrated.

Although there are several advantages to using QAS analyses, there are also some limitations. First, the quality-adjusted survival estimate that we have described must be restricted to a set time limit. Thus, in our analyses, the integration range was limited to the maximum follow-up of 7.8 years for the metformin study and 4.9 years for the DIG study, identical with that used in the traditional survival analysis. As such, this method cannot provide “lifetime” estimates associated with the therapy itself, but could be incorporated into projected life expectancy models to provide such estimates. Second, we used the simplest presentation of this method, which assumes a progressive health state model where admission to hospital precedes death. Although this is appropriate for many disease conditions and most composite outcomes, it may not be suitable for all; consider, for example, a composite outcome of admission to hospital or coronary artery bypass grafting. Parametric methods have been developed to overcome these limitations.<sup>19</sup> In addition, these models may be extended to account for repeated failure times and the use of time-varying covariates.<sup>8</sup> Third, all limitations and assumptions associated with Cox proportional hazards regression also apply

to the QAS method.<sup>20</sup> Fourth, we used literature-derived utility estimates, which may not be representative for the populations of the 2 studies analyzed here. Importantly, other preference-based utility measures are responsive to HRQL changes in patients with heart failure.<sup>21</sup> Therefore, although the actual utility weights assigned to the respective health states may vary and it is conceivable for our study that people with diabetes and heart failure had lower initial utility scores than patients with heart failure alone, the relative change in HRQL and associated differences in QAS between groups would be expected to be similar, irrespective of the utility measure. This assumption is supported by the sensitivity analysis with different utility coefficients. In addition, we assumed that utilities remained constant within each health state over time. The evidence to date is limited, but it is likely that patients with heart failure experience improvements and decrements in HRQL over time; however, given the high mortality among patients with heart failure, decrements in HRQL over the longer term would be expected. The assumption of constant utility over time may have led to an overestimation or underestimation of the true impact of patients' health state on their outcomes. To avoid this limitation, prospective collection of health utilities should be considered in future trials of heart failure therapy, as previously suggested. In addition, the QAS method uses average group utility estimates, not individual patient utilities. Even though average utilities will provide appropriate estimates for overall treatment effects, they may not be appropriate for assessment of treatment impact at the individual level. Finally, our examples were restricted to one type of composite outcome evaluated for only one common condition. However, this method can be extended to incorporate multiple end points comprising a composite outcome. Although such an analysis would be computationally more complex, at a conceptual level it may help to remember that in the area of cardiovascular research, most composite outcomes comprise nonfatal and fatal end points. Thus, complex composite outcomes could be distilled into a simple 2 health state model, with the utility assigned to the nonfatal end point weighted according to the distribution of nonfatal events.

In conclusion, thoughtful and well-constructed composite outcomes are important in observational studies and clinical trials, but traditional methods of analysis may not capture the whole picture. By incorporating patient-reported HRQL into survival analyses, the potential impact of the individual components of the composite outcome on the patient's health can be assessed more directly. As such, the potential benefits, harms and

costs associated with therapy may be more precisely quantified and may be made transparent for patients, providers and policy-makers.

---

**Contributors:** All of the authors participated in the study conception, design, analysis, interpretation of results and revision, and all of the authors reviewed and approved the final version. Dean Eurich drafted the initial manuscript and acts as guarantor. He had full access to all the data in the study and had final responsibility for the decision to submit for publication.

---

**Acknowledgements:** We extend our gratitude to Qi Lui for her assistance with the data analysis.

The Digitalis Investigation Group (DIG) is conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with the DIG Study Investigators. This manuscript was prepared using a limited-access data set obtained from the NHLBI and does not necessarily reflect the opinions or views of the DIG study or the NHLBI.

This study was based, in part, on nonidentifiable data provided by the Saskatchewan Department of Health. The interpretation and conclusions contained herein do not necessarily represent those of the Government of Saskatchewan or the Saskatchewan Department of Health.

---

## REFERENCES

- O'Connor CM, Gattis WA, Ryan TJ. The role of clinical nonfatal end points in cardiovascular phase II/III clinical trials. *Am Heart J* 2000;139(4):S143–S154.
- Taylor AL, Ziesche S, Yancy C, Carson P, D'Agostino R, Ferdinand K, et al. Combination of isosorbide dinitrate and hydralazine in blacks with heart failure. *N Engl J Med* 2004;351(20):2049–2057.
- DeMets DL, Califf RM. Lessons learned from recent cardiovascular clinical trials: Part I. *Circulation* 2002;106(6):746–751.
- Neaton JD, Gray G, Zuckerman BD, Konstam MA. Key issues in end point selection for heart failure trials: composite end points. *J Card Fail* 2005;11(8):567–575.
- Ferreira-González I, Busse JW, Heels-Ansdell D, Montori VM, Akl EA, Bryant DM, et al. Problems with use of composite end points in cardiovascular trials: systematic review of randomised controlled trials. *BMJ* 2007;334(7597):786.
- Glasziou PP, Simes RJ, Gelber RD. Quality adjusted survival analysis. *Stat Med* 1990(11);9:1259–1276.
- Guidelines for economic evaluation of pharmaceuticals: Canada*. 2nd ed. Ottawa: The Canadian Coordinating Office for Health Technology Assessment; 1997.
- Cole BF, Gelber RD, Goldhirsch A. Cox regression models for quality adjusted survival analysis. *Stat Med* 1993;12(10):975–987.
- Eurich DT, Majumdar SR, McAlister FA, Tsuyuki RT, Johnson JA. Improved clinical outcomes associated with metformin in patients with diabetes and heart failure. *Diabetes Care* 2005;28(10):2345–2351.
- The Digitalis Investigation Group. The effect of digoxin on mortality and morbidity in patients with heart failure. *N Engl J Med* 1997;336(8):525–533.
- Eurich DT, Majumdar SR, Tsuyuki RT, Johnson JA. Reduced mortality associated with the use of ACE inhibitors in patients with type 2 diabetes. *Diabetes Care* 2004;27(6):1330–1334.

12. Johnson JA, Majumdar SR, Simpson SH, Toth EL. Decreased mortality associated with metformin use compared to sulfonylurea monotherapy in type 2 diabetes mellitus. *Diabetes Care* 2002;25(12):2244–2248.
13. Maddigan SL, Feeny DH, Johnson JA. Health-related quality of life deficits associated with diabetes and comorbidities in a Canadian National Population Health Survey. *Qual Life Res* 2005;14(5):1311–1320.
14. Nichol G, Kaul P, Huszti E, Bridges JFP. Cost-effectiveness of cardiac resynchronization therapy in patients with symptomatic heart failure. *Ann Intern Med* 2004;141(5):343–351.
15. Efron B. Censored data and the bootstrap. *J Am Stat Assoc* 1981;76:312–319.
16. Horsman J, Furlong W, Feeny D, Torrance G. The Health Utilities Index (HUI): concepts, measurement properties and applications. *Health Qual Life Outcomes* 2003;1:54.
17. Lim E, Brown A, Helmy A, Mussa S, Altman DG. Composite outcomes in cardiovascular research: a survey of randomized trials. *Ann Intern Med* 2008;149(9):612–617.
18. Boyko WL Jr, Glick HA, Schulman KA. Economics and cost-effectiveness in evaluating the value of cardiovascular therapies. ACE inhibitors in the management of congestive heart failure: comparative economic data. *Am Heart J* 1999;137(5):S115–S119.
19. Cole BF, Gelber RD, Anderson KM. Parametric approaches to quality-adjusted survival analysis. International Breast Cancer Study Group. *Biometrics* 1994;50(3):621–631.
20. Southern DA, Faris PD, Brant R, Galbraith PD, Norris CM, Knudtson ML, et al. Kaplan–Meier methods yielded misleading results in competing risk scenarios. *J Clin Epidemiol* 2006;59(10):1110–1114.
21. Eurich DT, Johnson JA, Reid KJ, Spertus JA. Assessing responsiveness of generic and specific health related quality of life measures in heart failure. *Health Qual Life Outcomes* 2006;4:89.

---

**Citation:** Eurich DT, Majumdar SR, McAlister FA, Tsuyuki RT, Yasui Y, Johnson JA. Analyzing composite outcomes in cardiovascular studies: traditional Cox proportional hazards versus quality-of-life-adjusted survival approaches. *Open Med* 2010;1(1):40–48.

**Published:** 23 February 2010. Updated 8 March 2010.

**Copyright:** Open Medicine applies the Creative Commons Attribution Share Alike Licence, which means that anyone is able to freely copy, download, reprint, reuse, distribute, display or perform this work and that authors retain copyright of their work. Any derivative use of this work must be distributed only under a licence identical to this one and must be attributed to the authors. Any of these conditions can be waived with permission from the copyright holder. These conditions do not negate or supersede Fair Use laws in any country.

---