Supplemental Material (Appendix)

Calculating the Net Effect Estimate

The statistical method is fully described here and a free online calculator is available at ebscohealth.com/innovations. In brief, while we add the point estimates for each effect estimate to determine the point estimate for the net effect estimate, the calculation of the 95% confidence interval requires a couple of formulas.

Suppose we have an effect estimate X for one outcome and Y for another outcome and we want to determine a combined or net effect estimate Z. The model to determine the net effect estimate Z as a summative or linear combination of effect estimate X and effect estimate Y is based on the same statistical principles for determination of confidence intervals for differences between means, using addition instead of subtraction. That is,

$$Z = X + Y.$$

Assumptions regarding effect estimates include they:

1) represent data conforming to the normal distribution,
2) are independent and not correlated with each other, and
3) are expressed using the same units of measure.
4) The mean (or point estimate) for a net effect estimate Z is simply the addition of the means (or point estimates) for effect estimates X and Y. That is,

$$\text{Mean Z} = \text{Mean X} + \text{Mean Y.}$$

A 95% confidence interval for the net effect estimate is determined by calculating Mean Z +/- 1.96 SD_{MeanZ} where SD_{MeanZ} = standard deviation [SD] of Mean Z.

For the net effect estimate Z, the SD of Mean Z is related to the SDs of the component estimates Mean X and Mean Y through the formula:

$$SD_{MeanZ}^2 = SD_{MeanX}^2 + SD_{MeanY}^2$$

Therefore,

$$SD_{MeanZ} = \sqrt{(SD_{MeanX}^2 + SD_{MeanY}^2)}$$

The 95% confidence interval for the net effect will be:

$$\text{Mean Z - 1.96 } SD_{MeanZ} \text{ to Mean Z + 1.96 } SD_{MeanZ}$$

The third assumption (that effect estimates X and Y are expressed using the same units of measure) is rarely true so we need to introduce a "standardization" or "normalization" of outcomes, and this can be

done based on their relative importance.   One approach is to assign a multiplier (M) to each outcome representing its importance or relative value compared to a reference outcome. The reference outcome can be external to the body of evidence, or can be one of the outcomes of interest (in which case the value of M for the reference outcome will be 1).

The mean (or point estimate) for a net effect estimate Z, expressed in units of multiples of the reference outcome, becomes:

$$\text{Mean Z} = (M_X \times \text{Mean X}) + (M_Y \times \text{Mean Y})$$

With the use of multipliers, the SD of the net effect estimate Z becomes related to the formula:

$$SD_{MeanZ}^2 = (M_X \times SD_{MeanX})^2 + (M_Y \times SD_{MeanY})^2$$

Therefore,

$$SD_{MeanZ} = \sqrt{(M_X^2 SD_{MeanX}^2 + M_Y^2 SD_{MeanY}^2)}$$

Note that if the SD_{Mean} is not directly reported for an individual effect estimate, it can be derived from the width of the 95% confidence interval (CIW) for the effect estimate:

$$SD_{MeanX} = CIW_X / 3.92$$

Using the data for the sacubitril-valsartan example (with units of hospitalization-equivalent events per 1000 patients) we get:

$SD_{\text{All-cause mortality outcome}}$ = CIW of 160 / 3.92 = 40.816

$SD_{\text{Hospitalization rate outcome}}$ = CIW of 27 / 3.92 = 6.888

$SD_{\text{Symptomatic hypotension rate outcome}}$ = CIW of 12 / 3.92 = 3.061

Applying $SD_{\text{MeanZ}} = \sqrt{(SD_{\text{MeanX}}^2 + SD_{\text{MeanY}}^2)}$ we get:

$SD_{\text{Net effect estimate}} = \sqrt{(SD_{\text{All-cause mortality}}^2 + SD_{\text{Hospitalization rate}}^2 + SD_{\text{Symptomatic hypotension rate}}^2)}$

$SD_{\text{Net effect estimate}} = \sqrt{(40.816^2 + 6.888^2 + 3.061^2)}$

$SD_{\text{Net effect estimate}} = \sqrt{(1665.9459 + 47.4445 + 9.3697)}$

$SD_{\text{Net effect estimate}} = \sqrt{(1722.7601)} = 41.5$

The 95% confidence interval for the net effect estimate is the mean +- 1.96 SD. For the lower boundary, this translates to 154 - (1.96)(41.5) = 154-81.34 = 72.66 (rounded to 73) and for the upper boundary, this would be 154 + 81.34 = 235.34 (rounded to 235).

We report a net effect estimate of a decrease in 154 hospitalization-equivalent events per 1000 patients (95% confidence interval for the net effect estimate being 73 fewer to 235 fewer hospitalization-equivalent events per 1000 patients).

Sensitivity analysis of the net effect estimate

The 95% confidence interval implies a range within which the true net effect is likely to occur. There are many factors that can affect the certainty that the true net effect is within this range.

If assumptions used in the model are not met, the results will not have accurate precision. If individual outcomes are correlated (such as increase in one benefit being correlated with an increase in another benefit), the "true" 95% confidence interval would be wider or less precise than the one estimated by our method. Alternatively, if individual outcomes are inversely correlated (such as an increase in a benefit being correlated with an increase in a harm, or correlated with a decrease in another benefit), then the "true" 95% confidence interval would be narrower or more precise than the one estimated by our method. In the latter case our proposed approach is conservative but less powerful. If outcomes have other dependencies or do not follow a normal distribution (such as a highly skewed distribution), then the 95% confidence interval may be inaccurate.

The statistical formulas can be adjusted with correlation coefficients if they can be estimated. The formula to determine the standard deviation of the mean of the net effect

$$SD_{MeanZ}^2 = (M_X \times SD_{MeanX})^2 + (M_Y \times SD_{MeanY})^2$$

is modified to

$$SD_{MeanZ}^2 = (M_X \times SD_{MeanX})^2 + (M_Y \times SD_{MeanY})^2 + (2 \times r \times M_X \times SD_{MeanX} \times M_Y \times SD_{MeanY})$$

where r = the correlation coefficient between X and Y. Correlation coefficients are rarely available but the maximum value of r that appears plausible can be used for a sensitivity analysis to address plausible correlations between outcomes.

For the sacubitril-valsartan example, there is data suggesting a small inverse correlation (r = -0.17) between all-cause mortality and hospitalization for heart failure among patients with heart failure (25). There is no data addressing correlations between drug-related symptomatic hypotension and the outcomes of mortality or hospitalization. Let's assume r = 0.5 for each of these as an upper bound of plausible correlations for a sensitivity analysis.

$SD_{Net\ effect\ estimate}$ = √ ($SD_{All-cause\ mortality}^2$ + $SD_{Hospitalization\ rate}^2$ + $SD_{Symptomatic\ hypotension\ rate}^2$) + $2r(SD_{Mortalty})$ $(SD_{Hospiltalization})$ + $2r(SD_{Mortalty})$ $(SD_{Symptomatic\ hypotension\ rate})$ + $2r(SD_{Symptomatic\ hypotension\ rate})$ $(SD_{Hospitalization})$

$SD_{Net\ effect\ estimate}$ = √ (40.816^2 + 6.888^2 + 3.061^2) + 2(-0.17)(40.816)(6.888) + 2(0.5)(40.816)(3.061) + 2(0.5)(3.061)(6.888)

$SD_{Net\ effect\ estimate}$ = √ (1722.7601) + (-95.5878) + (124.9378) + (21.0842)

$SD_{Net\ effect\ estimate}$ = √ (1773.1943) = 42.1

This net effect estimate (in a sensitivity analysis adjusting for known and plausible correlations among outcomes) is a decrease in 154 hospitalization-equivalent events per 1000 patients (95% confidence interval 71 fewer to 237 fewer hospitalization.

Example 1. Longer dual-antiplatelet therapy (DAPT) after drug-eluting stents

A systematic review comparing longer versus shorter durations of DAPT after drug eluting stent placement provides the summary of effect estimates for longer duration DAPT in Appendix Table 1 (26). Longer duration of DAPT ranged from 12 months to 42 months and shorter duration of DAPT ranged from 3 months to 18 months (26).

Appendix Table 1. Summary of findings for longer versus shorter durations of dual antiplatelet therapy after drug eluting stent placement

Outcome	Absolute effect estimate per 1000 patients (95% confidence interval)	Certainty of effect estimates
All-cause mortality	2 more (0 change to 4 more)	High*
Myocardial infarction	8 fewer (12 fewer to 2 fewer)	Moderate
Major bleeding	6 more (3 more to 10 more)	High
Any stroke	0 change (2 fewer to 2 more)	High*

* originally reported as moderate quality evidence with downgrade limited to precision. Precision downgrade for single outcome effect estimates are not relevant in this approach as the confidence intervals are being used in the determination of the net effect estimate.

Step 1. Determine the outcomes to be combined.

All four outcomes (mortality, myocardial infarction, major bleeding, and stroke) are considered impactful to include in net effect estimates. None are overlapping outcomes with the assumptions that hemorrhagic stroke contributes minimally to estimates of major bleeding, and fatal outcomes contribute minimally to estimates of myocardial infarction, major bleeding and stroke.

Step 2. Determine the quantified relative importance for each outcome.

Myocardial infarction-equivalent will be considered the reference unit. For example purposes, we will start with the assumption that patients would consider the importance of a myocardial infarction and major bleeding similarly, consider a stroke 3 times more important, and consider mortality 5 times more important. These assignments of relative importance of outcomes are derived from systematic review of evidence of relative importance of outcomes for myocardial infarction, major bleeding and stroke (17) and without empiric investigation for the mortality outcome (3).

Step 3. Combine the importance-adjusted effect estimates.

Importance-adjusted effect estimates are determined by multiplying each effect estimate by its relative importance multiplier. Our importance-adjusted effect estimates (in units of myocardial infarction-equivalent events per 1000 patients) are summarized in Appendix Table 2:

Appendix Table 2. Importance-adjusted effect estimates for longer versus shorter durations of dual antiplatelet therapy after drug eluting stent placement

Outcome	Absolute effect estimate (in units of myocardial infarction-equivalent events per 1000 patients) (95% confidence interval)
All-cause mortality	10 more (0 change to 20 more)
Myocardial infarction	8 fewer (12 fewer to 2 fewer)
Major bleeding	6 more (3 more to 10 more)
Any stroke	0 change (6 fewer to 6 more)

The effect estimates are combined using the online calculator at ebscohealth.com/innovations (see Appendix Part 2). The net effect estimate is an increase in 8 myocardial infarction-equivalent events per 1000 patients (95% confidence interval [CI] decrease in 5 to increase in 21 myocardial infarction-equivalent events per 1000 patients).

Appendix Figure 1. Effect estimates for longer versus shorter durations of dual antiplatelet therapy after drug eluting stent placement
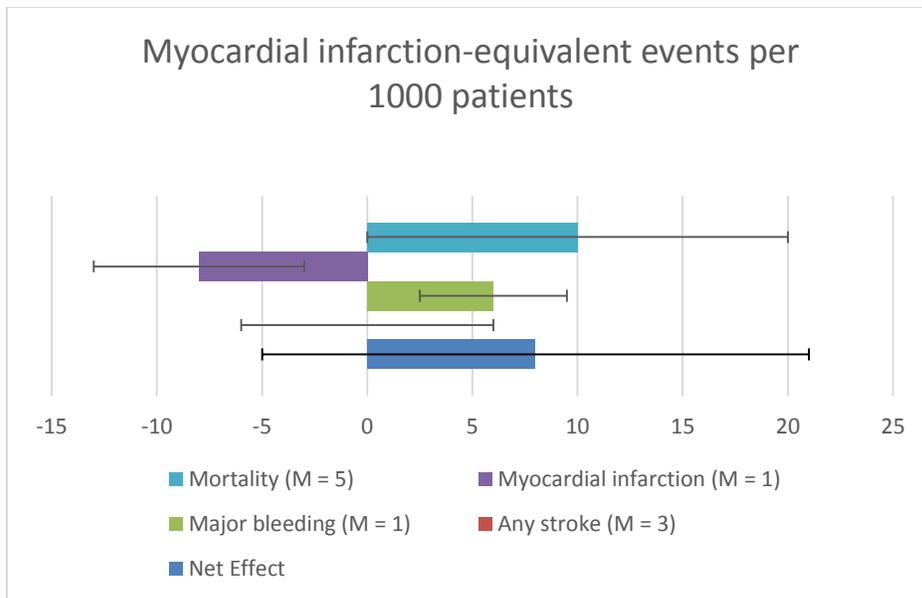
Step 4. Classify the precision of the net effect estimate.

The net effect point estimate is harmful, the lower bound of the confidence interval for the net effect estimate is beneficial, and the absolute value of the lower bound of the confidence interval is smaller than the absolute value of the net effect point estimate. This pattern is likely net harm, and consistent with a moderate certainty of net harm.

Step 5. Consider the certainty of effect estimates for outcomes that are critical to the likelihood of net benefit.

Mortality and major bleeding are critical outcomes (potential differentiators of the likelihood of net benefit) because removal of either outcome could change the pattern to one suggesting net benefit. Stroke and myocardial infarction have limited impact on the net effect classification. Both critical outcomes have high certainty of evidence so this does not change our moderate certainty of net harm.

Step 6. Consider the range of relative importance for outcomes. Determine if the net effect estimate across the range of relative importance changes the certainty of net benefit rating.

If patients considered reduction of myocardial infarction to have higher relative importance than mortality and major bleeding it is possible to derive a net benefit. Such relative importance ratings are plausible because myocardial infarction can have a greater contribution to long-term quality of life. Consideration of the range of relative importance for outcomes leads to a low certainty of net harm.

Completing the evidence-to-decision framework

With a low certainty of net harm, the expected result is a weak recommendation against longer duration DAPT after drug-eluting stent placement. The costs are relatively low and there are little adverse consequences related to acceptability, feasibility and equity, so guideline panels may consider to make a weak recommendation against longer duration DAPT.

At the current time, major guidelines have inconsistent recommendations for this concept. The American College of Chest Physicians makes a strong recommendation against DAPT (and for single antiplatelet therapy) after 12 months following drug-eluting stent placement (27). The American College of Cardiology makes a weak recommendation suggesting continuing DAPT beyond 12 months may be considered in patients receiving drug-eluting stents (28).

Example 2. Sacubitril-valsartan for symptomatic heart failure

This example is a decision or recommendation to use sacubitril-valsartan instead of an angiotensin-converting enzyme (ACE) inhibitor or angiotensin receptor blocker (ARB) in patients with symptomatic heart failure with reduced ejection fraction despite treatment with an ACE inhibitor or ARB. A systematic evidence review and GRADE evidence profile for such use of sacubitril-valsartan finds the effect estimates in Appendix Table 3 (29, 30), based on a single trial (31).

Appendix Table 3. Summary of findings for sacubitril-valsartan versus continued ACE inhibitor or ARB in symptomatic heart failure with reduced ejection fraction

Outcome	Absolute effect estimate per 1000 patients (95% confidence interval)	Certainty of evidence*
All-cause mortality	29 fewer (12 fewer to 44 fewer)	Moderate
Cardiovascular mortality	31 fewer (17 fewer to 45 fewer)	Moderate
Hospitalization for worsening heart failure	31 fewer (16 fewer to 43 fewer)	Moderate
Symptomatic hypotension	44 more (33 more to 57 more)	Moderate
Change in heart failure symptom score (scale 0-100)	1.64 points decrease (0.63-point decrease to 2.65-point decrease)	Moderate
Decline in renal function	4 fewer (3 fewer to 9 fewer)	Moderate

* Certainty of evidence ratings here do not rate down for imprecision.

Step 1. Determine the outcomes to be combined

Two outcomes were dropped from consideration because they were considered to have little to no impact on the net effect. In the sacubitril-valsartan example decline in renal function was considered not impactful for determination of the net effect estimate because the effect size is small and the outcome has low importance to patients. Change in heart failure symptom score was considered not impactful for determination of the net effect estimate because the effect size is small and the relative importance is uncertain and may be accounted for in other outcomes. Using means for a continuous score can be misleading when one considers the impact on individual patients who vary in their responses (i.e. assuming every patient experiences the mean effect is likely an erroneous assumption). The only data regarding the proportion of patients who have an important change in symptoms is the outcome of hospitalization for worsening heart failure, and authors of the study reported this outcome.

All-cause mortality is selected instead of cardiovascular mortality to avoid duplicate counting of mortality. The outcomes included in net effect estimation are all-cause mortality, hospitalization for worsening heart failure and symptomatic hypotension.

Step 2. Determine the quantified relative importance for each outcome

Hospitalization-equivalent events per 1000 patients will be considered the reference unit. We do not readily find empiric evidence for the relative importance of outcomes in patients with heart failure. We

will start with the assumption that patients would consider the outcome of all-cause mortality 5 times more important than an episode of hospitalization, and an outcome of symptomatic hypotension half as important as being hospitalized.

Step 3. Combine the importance-adjusted effect estimates

Importance-adjusted effect estimates are determined by multiplying each effect estimate by its relative importance multiplier. Our importance-adjusted effect estimates (in units of hospitalization-equivalent events per 1000 patients) are summarized in Appendix Table 4:

Appendix Table 4. Importance-adjusted effect estimates for sacubitril-valsartan versus continued ACE inhibitor or ARB in symptomatic heart failure with reduced ejection

Outcome	Absolute effect estimate (in units of hospitalization-equivalent events per 1000 patients) (95% confidence interval)
All-cause mortality	145 fewer (60 fewer to 220 fewer)
Hospitalization for heart failure	31 fewer (16 fewer to 43 fewer)
Symptomatic hypotension	22 more (16.5 more to 28.5 more)

The effect estimates are combined using the online calculator at ebscohealth.com/innovations and the calculations are shown in part in Appendix Part 2.

The net effect point estimate is a decrease in 154 hospitalization-equivalent events per 1000 patients. (-145 plus -31 plus +22 = -154)

The net effect estimate is a decrease in 154 hospitalization-equivalent events per 1000 patients. (95% CI 73 fewer to 235 fewer hospitalization-equivalent events per 1000 patients)

Appendix Figure 2. Effect estimates for sacubitril-valsartan versus continued ACE inhibitor or ARB in symptomatic heart failure with reduced ejection fraction

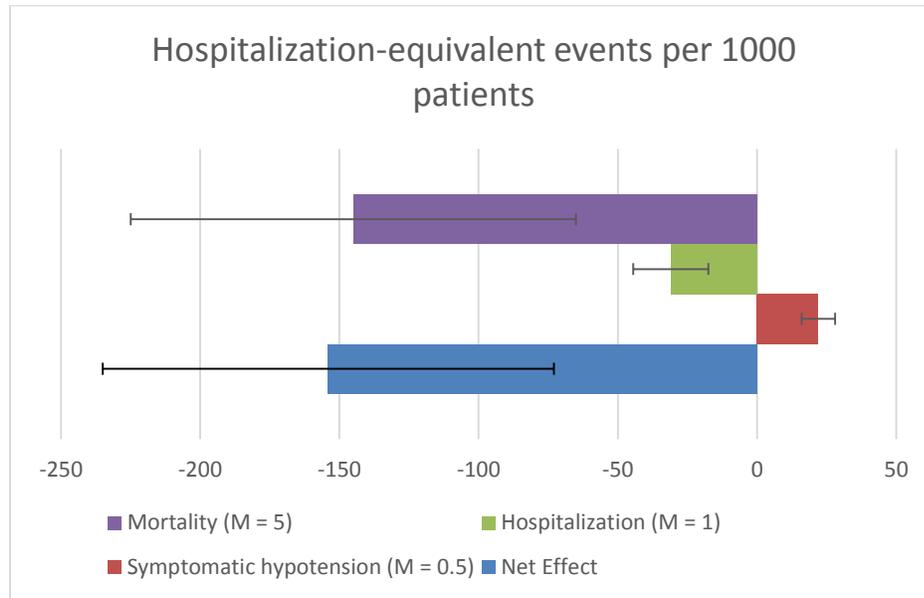**Step 4. Classify the precision of the net effect estimate.**

The entire confidence interval of the net effect estimate is beneficial so the pattern is net benefit, consistent with a high certainty of net benefit.

Step 5. Consider the certainty of effect estimates for outcomes that are critical to the likelihood of net benefit.

Mortality is potentially differentiating because removal of a mortality effect would change the net effect estimate from 154 fewer (95% CI 73 fewer to 235 fewer) hospitalization-equivalent events per 1000 patients to 9 fewer (95% CI 24 fewer to 6 more) events per 1000 patients, and the overall pattern would change from net benefit to likely net benefit.

Hospitalization for heart failure is not potentially differentiating because removal from the net effect estimate would not change the pattern from net benefit. The net effect estimate would be 123 fewer (95% CI 43 fewer to 203 fewer) events per 1000 patients. A result of increasing hospitalizations for heart failure is not a plausible likelihood. One could question whether total hospitalizations should be used as an outcome rather than cause-specific hospitalization. The outcome of total hospitalizations was not reported in the underlying evidence (28), and guideline panels would need to determine if such an outcome is impactful enough to reassess the overall balance of benefits and harms for this decision.

Symptomatic hypotension is initially not potentially differentiating because removal from the net effect estimate would not change the pattern from net benefit. Symptomatic hypotension can still be considered critical because a higher rate of symptomatic hypotension than observed in the underlying

evidence is plausible, especially related to the use of run-in periods excluding patients who did not tolerate study medications.

The critical outcomes have effect estimates with moderate certainty. This leads to a moderate certainty of net benefit.

Step 6. Consider the range of relative importance for outcomes. Determine if the net effect estimate across the range of relative importance changes the certainty of net benefit rating.

We started with an assumption that the average patient would consider the importance of all-cause mortality five times more important than an episode of hospitalization, and an outcome of symptomatic hypotension half as important as being hospitalized. To consider a range of relative importance for outcomes we should consider the lowest relative importance for all-cause mortality and highest relative importance for symptomatic hypotension that would occur among patients facing this decision and is considered reasonable to reflect the range of importance among common, rational people.   Some patients (such as those with terminal illness) may place higher importance on how they feel than mortality so for these patients they might consider mortality, symptomatic hypotension, and hospitalization to be equivalent.

Using assumptions of equivalence across these three outcomes the net effect estimate would be 16 fewer (95% CI 40 fewer to 9 <u>more</u>) events per 1000 patients.

With a reasonable limit for the range of relative importance (including most patients) weighted to support net harm, the net effect estimate changes from net benefit to likely net benefit.   If there were otherwise high certainty of net benefit this finding could reduce our certainty to moderate certainty of net benefit. As we already have a moderate certainty of net benefit, extreme assumptions reaching likely net benefit do not further change our certainty.

Completing the evidence-to-decision framework

In an assessment in 2015 the moderate certainty of net benefit justified a weak recommendation for sacubitril-valsartan (29, 30). The high cost further supported a weak recommendation. Four national guidelines have since made strong recommendations for the use of sacubitril-valsartan (32-35), though the findings have not been replicated in a second trial. A recommendations panel reconsidered the rationale across all four guidelines and reconfirmed a weak recommendation for sacubitril-valsartan based on a moderate certainty of evidence (limited to a single trial with potential selection bias related to the run-in period), a moderate certainty of net benefit (considering the range of quantitative estimates of importance of outcomes), and high cost with some uncertainty in the cost-benefit ratio (29). A different recommendations panel could generate different ratings, but the process allows explicit and transparent expression of what is being rated and how it is rated.

Example 3. Ivabradine for symptomatic heart failure

Ivabradine is a heart rate lowering drug which has been tested for clinical use in patients with heart failure in two large randomized trials (36, 37). In the first trial ivabradine was not associated with overall clinical benefit and was not associated with any decrease in death or hospitalization attributed to heart failure (36). In the second trial with more stringent selection criteria (left ventricular ejection fraction <= 35%, heart rate => 70 beats/minute) ivabradine reduced the rate of hospital admissions for worsening heart failure (37).

Outcome differences with ivabradine instead of placebo (from randomization until first event, up to 42 months) are summarized in Appendix Table 5 (38):

Appendix Table 5. Summary of findings for ivabradine versus placebo in symptomatic heart failure with reduced ejection fraction

Outcome	Absolute effect estimate per 1000 patients (95% confidence interval)	Certainty of effect estimates
All-cause mortality	13.9 fewer (31.8 fewer to 4 more)	Moderate
Cardiovascular mortality	11.9 fewer (29 fewer to 5.2 more)	Moderate
Death from heart failure	11.4 fewer (21 fewer to 1.8 fewer)	Moderate
Hospitalization for any cause	35.6 fewer (59.4 fewer to 11.8 more)	Moderate
Hospitalization for cardiovascular reason	42.3 fewer (65 fewer to 19.6 more)	Moderate
Hospitalization for worsening heart failure	47.3 fewer (66 fewer to 28.6 fewer)	Moderate
Bradycardia	33.5 more (26 more to 41 more)	High
Phosphenes (a visual adverse effect)	22.6 more (16.5 more to 28.8 more)	High
Atrial fibrillation	12.1 more (2.1 more to 22 more)	High

The second trial had a low risk of bias though the quality of evidence could be considered moderate for benefits based on inconsistency with the first trial. The adverse effects data could be considered as high quality evidence as the findings are consistent with the first trial (39).

Step 1. Determine the outcomes to be combined.

All-cause mortality, hospitalization for any cause, bradycardia, phosphenes, and atrial fibrillation are selected as non-overlapping outcomes.

Step 2. Determine the quantified relative importance for each outcome.

Hospitalization-equivalent relative importance will be estimated at 0.3 for each adverse effect and 5 for mortality.

Step 3. Combine the importance-adjusted effect estimates

The importance-adjusted effect estimates (in units of hospitalization-equivalent events per 1000 patients) are in Appendix Table 6.

Appendix Table 6. Importance-adjusted effect estimates for ivabradine versus placebo in symptomatic heart failure with reduced ejection fraction

Outcome	Absolute effect estimate (in units of hospitalization-equivalent events per 1000 patients) **(95% confidence interval)**	Confidence interval width (CIW)
All-cause mortality	69.5 fewer (159 fewer to 20 more)	179 per 1000
Hospitalization for any cause	35.6 fewer (59.4 fewer to 11.8 more)	71.2 per 1000
Bradycardia	10.05 more (7.8 more to 12.3 more)	4.5 per 1000
Phosphenes (a visual adverse effect)	6.78 more (4.95 more to 8.64 more)	3.69 per 1000
Atrial fibrillation	3.63 more (0.63 more to 6.6 more)	5.97 per 1000

The net effect point estimate is a decrease in 85 hospitalization-equivalent events per 1000 patients. (-69.5 plus -35.6 plus +10.05 plus +6.78 plus +3.63 = -84.64)

The net effect estimate is a decrease in 85 (95% confidence interval decrease in 181 to increase in 12) hospitalization-equivalent events per 1000 patients (see Appendix Figure 3).

Appendix Figure 3. Effect estimates for ivabradine versus placebo in symptomatic heart failure with reduced ejection fraction
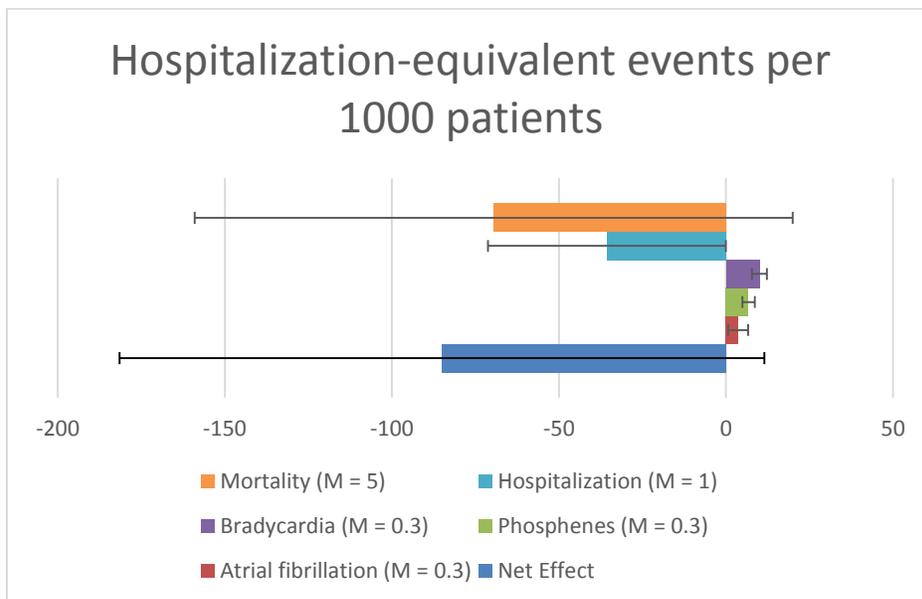
Step 4. Classify the precision of the net effect estimate.

The pattern is likely net benefit, consistent with a moderate certainty of net benefit.

Step 5. Consider the certainty of effect estimates for outcomes that are critical to the likelihood of net benefit.

As the effect estimates have at least moderate certainty of evidence there is moderate certainty of net benefit.

Step 6. Consider the range of relative importance for outcomes. Determine if the net effect estimate across the range of relative importance changes the certainty of net benefit rating.

A plausible range of relative importance could include consideration of death equivalent to hospitalization and other adverse effects 0.6 times as disruptive as hospitalization.   Such assumptions would lead to importance-adjusted effect estimates of:

- All-cause mortality: 13.9 fewer (95% CI 31.8 fewer to 4 more)
- Hospitalization for any cause: 35.6 fewer (95% CI 59.4 fewer to 11.8 more)
- Bradycardia: 20.1 more (95% CI 15.6 more to 24.6 more)
- Phosphenes (a visual adverse effect):   13.56 more (95% CI 9.90 more to 17.28 more)
- Atrial fibrillation: 7.26 more (95% CI 1.26 more to 13.2 more)

These estimates would result in a net effect estimate of a decrease in 9 (95% confidence interval decrease in 49 to increase in 32) hospitalization-equivalent events per 1000 patients.

This would be possible net benefit, and results in a low certainty of net benefit upon consideration across the range of relative importance that patients may have for the various effects.

Completing the Evidence-to-Decision Framework

A low certainty of net benefit supports a weak recommendation for ivabradine in patients meeting the selected criteria used in the trial suggesting benefit. Three of four current guidelines provide a weak recommendation for ivabradine in this setting (40-42) while one makes a strong recommendation (43).

Example 4. Second autologous stem cell transplant (ASCT) for patients with relapsed myeloma and response duration more than 2 years after first ASCT

A National Institute for Health and Clinical Excellence (NICE) guideline includes GRADE profiles for a second ASCT in relapsed myeloma including (44):

- Median overall survival from relapse – low quality evidence – absolute effect 2.1 years longer (95% CI not reported)
- Median time to progression – moderate quality evidence – absolute effect 13 months longer (95% CI not reported)
- No evidence identified for treatment-related morbidity and mortality, health-related quality of life, and adverse effects.

Step 1. Determine the outcomes to be combined.

The guideline panel considered overall survival and progression-free survival to be the most impactful outcomes for consideration. Because overall survival includes progression-free survival, progression events (time to progression) and death events (time to death) can be counted as non-overlapping outcomes.

Step 2. Determine the quantified relative importance for each outcome.

Time to death (overall survival) will be considered the reference unit. We will start with the assumption that patients would consider time to progression 0.2 times the importance of time to death.

Step 3. Combine the importance-adjusted effect estimates.

The importance-adjusted estimate for time to progression is median 2.6 months (0.2 x 13 months) and the estimate for time to death is median 2.1 years (or 25.2 months).

The point estimate for the net effect is the equivalent of 27.8 months of increased survival. Confidence intervals were not reported.

Step 4. Classify the precision of the net effect estimate.

This appears to start with a pattern of net benefit.   The statistical significance was not expressed but assuming the results were statistically significant the confidence intervals would be completely within estimates of benefit because no evidence was provided to suggest harm.

Step 5. Consider the certainty of effect estimates for outcomes that are critical to the likelihood of net benefit.

Overall survival is the critical outcome here and was reported as low certainty of evidence based on a single retrospective comparative study (and related consistent data in noncomparative studies). Even so the guideline panel could potentially consider this to represent a moderate certainty of a survival

benefit and a low certainty for a specific magnitude of effect.   This could lead to a moderate certainty of net benefit.

Step 6. Consider the range of relative importance for outcomes. Determine if the net effect estimate across the range of relative importance changes the certainty of net benefit rating.

In a model with no harms being considered as important outcomes, the range of relative importance is really an opportunity to consider how potential harms may affect the balance of benefits and harms. The guideline panel rationalized that harms would be similar to what patients experienced with their first ASCT and patients would thus have individual experience representing their individual harms estimates when considering the balance of harms and benefits.

Completing the Evidence-to-Decision Framework

In the context of harms mainly being considered burdens the patient would individually consider, the potential for increases in overall survival is considered a moderate certainty of net benefit. To reflect the importance of the patient weighing a personalized relative importance the guideline panel made a strong recommendation to offer the therapy (for the potential for net benefit) rather than recommend that the therapy should be administered.

Example 5. Avoiding 100% oxygen saturation in intensive care unit

A randomized trial with 480 adults admitted to the intensive care unit (ICU) found mortality in the ICU of 11.6% with target arterial oxyhemoglobin saturation (SpO2) 94%-98% versus 20.2% with target SpO2 97%-100% (absolute risk reduction 8.6%, 95% CI 1.7% to 15%) (45). This evidence may be considered to have moderate certainty due to early trial termination without use of a formal stopping rule.

Step 1. Determine the outcomes to be combined.

Although other outcomes were reported, there was no evidence of benefits for high-SpO2, so the mortality outcome can be considered the primary outcome for a net effect estimate.

Step 2. Determine the quantified relative importance for each outcome.

Step 3. Combine the importance-adjusted effect estimates.

These steps are irrelevant in this case and the net effect estimate is the estimate for ICU mortality, which can be considered inversely for the action of targeting an SpO2 97%-100% (absolute risk increase 8.6%, 95% confidence interval 1.7% to 15%).

Step 4. Classify the precision of the net effect estimate.

There is a net harm based on the confidence intervals of the effect estimate, consistent with a high certainty of net harm.

Step 5. Consider the certainty of effect estimates for outcomes that are critical to the likelihood of net benefit.

As the underlying evidence is a single trial with early unplanned termination, the moderate certainty of evidence may reduce the certainty of net harm to moderate.

Step 6. Consider the range of relative importance for outcomes. Determine if the net effect estimate across the range of relative importance changes the certainty of net benefit rating.

This is irrelevant following the decision to focus on a single outcome

Completing the Evidence-to-Decision Framework

A moderate certainty of net harm is sufficient to support a strong recommendation against an intervention with no apparent benefit.

Summary of Examples.

Five examples are presented to show how the model for defining and reporting the certainty of net benefit or certainty of net harm can provide more clear and explicit representations of evidence-based assessments and judgments supporting a recommendation spanning a broad continuum of complex situations. See Appendix Table 7.

Example 1 (Longer DAPT after drug-eluting stents) shows a net effect estimate suggesting a low certainty of net harm. Adjustment for certainty of evidence and the range of relative importance across outcomes is unnecessary as the certainty is already low. No other factors change the approach to the recommendation and we support a weak recommendation against. This may be clearer than the variations across current guidelines ranging from a weak recommendation for to a strong recommendation against.

Example 2 (Sacubitril-valsartan for symptomatic heart failure on standard therapy) is an example in which a net effect estimate shows net benefit based on a single trial using reasonable assumptions of relative importance of outcomes. The moderate certainty of evidence and the influence of reasonable extremes of relative importance assignments each led to ratings of a moderate certainty of net benefit. A moderate certainty of net benefit and high cost may support a weak recommendation for sacubitril-valsartan although many current guidelines provide a strong recommendation.

Example 3 (Ivabradine for symptomatic heart failure) shows a treatment with relatively smaller effects on benefits and harms with a closer balance between benefits and harms and moderate certainty of effect estimates. The resulting low certainty of net benefit supports a weak recommendation for ivabradine, and most current guidelines provide a weak recommendation for it.

Example 4 (Second ASCT for patients with relapsed myeloma and response duration more than 2 years after first ASCT) starts with low to moderate certainty in effect estimates for benefits and no direct comparative evidence to quantify harms. This leads to a higher certainty of net benefit, though still a moderate certainty of net benefit given the limited certainty in effect estimates. Despite not reaching a high certainty of net benefit the guideline panel reasoned that the harms for a second ASCT would be patient-specific (and patient-recognized based on the first ASCT) so provided a strong recommendation to offer the therapy and allow the patient to individually weigh the estimated benefits against their individualized harms. This is consistent with the GRADE approach which would provide a weak recommendation and encourage shared decision making. The guideline panel did not provide a strong recommendation for the intervention without shared decision making.

Example 5 (Avoiding 100% oxygen saturation in intensive care unit) is an example of an intervention with no apparent benefit and moderate certainty of net harm.   The quantitative effect estimates support a high certainty of net harm but the risk of bias (qualitative certainty) reduced the overall assessment to a moderate certainty of net harm. Even so, without any apparent benefit, a strong recommendation against would be justified.

Appendix Table 7. Certainty of net benefit and strength of recommendations in examples

Example	Certainty of Evidence for Critical Outcomes*	Certainty of Net Benefit	Strength of Recommendation
Longer dual-antiplatelet therapy (DAPT) after drug-eluting stents	Moderate to high certainty of evidence	Low certainty of net harm	Weak recommendation against
Sacubitril-valsartan for symptomatic heart failure on standard therapy	Moderate certainty of evidence	Moderate certainty of net benefit	Weak recommendation for
Ivabradine for symptomatic heart failure	Moderate certainty of evidence	Low certainty of net benefit	Weak recommendation for
Second autologous stem cell transplant (ASCT) for patients with relapsed myeloma and response duration more than 2 years after first ASCT	Low to moderate certainty of evidence	Moderate certainty of net benefit	Weak recommendation for (or strong recommendation for offering with shared decision making)
Avoiding 100% oxygen saturation in intensive care unit	Moderate certainty of evidence	Moderate certainty of net harm	Strong recommendation against

* Certainty of evidence ratings here do not rate down for imprecision.

References

References 1-24 are the same as those included in the index article.

1. URL: www.gradeworkinggroup.org. Accessed November 27, 2016.
2. Atkins D, Best D, Briss PA, et al; GRADE Working Group. Grading quality of evidence and strength of recommendations. BMJ 2004;328:1490. doi: 10.1136/bmj.328.7454.1490.
3. Balshem H, Helfand M, Schunemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol 2011;64:401-6. doi: 10.1016/j.jclinepi.2010.07.015.
4. Hultcrantz M, Rind D, Akl E, et al; GRADE Working Group. The GRADE Working Group clarifies the construction of certainty of evidence. J Clin Epidemiol 2017;87:4-13. doi: 10.1016/j.jclinepi.2017.05.006.
5. Andrews JC, Schünemann HJ, Oxman AD, et al. GRADE guidelines: 15. Going from evidence to recommendation-determinants of a recommendation's direction and strength. J Clin Epidemiol 2013;66:726-35. doi: 10.1016/j.jclinepi.2013.02.003.
6. Editors: Schunemann H, Brozek J, Guyatt G, Oxman A. GRADE Handbook. Handbook for grading the quality of evidence and the strength of recommendations using the GRADE approach. Updated October 2013. Section 6.3.2 Confidence in best estimates of magnitude of effects (quality of evidence). URL: gdt.guidelinedevelopment.org/app/handbook/handbook.html#h.s1fsemn2yjpm. Accessed November 27, 2016.
7. Guyatt GH, Oxman AD, Kunz R, et al; GRADE Working Group. Going from evidence to recommendations. BMJ 2008;336:1049-51. doi: 10.1136/bmj.39493.646875.AE.
8. Editors: Schunemann H, Brozek J, Guyatt G, Oxman A. GRADE Handbook. Handbook for grading the quality of evidence and the strength of recommendations using the GRADE approach. Updated October 2013. Section 6.3.3 Confidence in values and preferences. URL: http://gdt.guidelinedevelopment.org/app/handbook/handbook.html#h.i5hfweocv3qs. Accessed November 27, 2016.
9. Editors: Schunemann H, Brozek J, Guyatt G, Oxman A. GRADE Handbook. Handbook for grading the quality of evidence and the strength of recommendations using the GRADE approach. Updated October 2013. Section 6.2.1 Balance of desirable and undesirable consequences. URL: http://gdt.guidelinedevelopment.org/app/handbook/handbook.html#h.feb5g22tf6nv. Accessed November 27, 2016.
10. Editors: Schunemann H, Brozek J, Guyatt G, Oxman A. GRADE Handbook. Handbook for grading the quality of evidence and the strength of recommendations using the GRADE approach. Updated October 2013. Section 6.3.4 Resource use (cost). URL: http://gdt.guidelinedevelopment.org/app/handbook/handbook.html#h.y0ey8465h47v. Accessed November 27, 2016.
11. Editors: Schunemann H, Brozek J, Guyatt G, Oxman A. GRADE Handbook. Handbook for grading the quality of evidence and the strength of recommendations using the GRADE approach. Updated October 2013. Section 6.5 The Evidence-to-Decision framework. URL: http://gdt.guidelinedevelopment.org/app/handbook/handbook.html#h.xr5ac2p2khuq. Accessed November 27, 2016.
12. Sawaya GF, Guirguis-Blake J, LeFevre M, Harris R, Petitti D, U.S. Preventive Services Task Force. Update on the methods of the U.S. Preventive Services Task Force: estimating certainty and

magnitude of net benefit. Ann Intern Med 2007;147:871-5. doi: 10.7326/0003-4819-147-12-200712180-00007.

13. Krist AH, Wolff TA, Jonas DE, et al. Update on the methods of the U.S. Preventive Services Task Force: methods for understanding certainty and net benefit when making recommendations. Am J Prev Med 2018;54:S11-S18. doi: 10.1016/j.amepre.2017.09.011.

14. Editors: Schunemann H, Brozek J, Guyatt G, Oxman A. GRADE Handbook. Handbook for grading the quality of evidence and the strength of recommendations using the GRADE approach. Updated October 2013. Section 3 Selecting and rating the importance of outcomes. URL: https://gdt.gradepro.org/app/handbook/handbook.html#h.1i2bwkm8zpjo. Accessed April 6, 2019.

15. Hazlewood GS, Bombardier C, Tomlinson G, Marshall D. A Bayesian model that jointly considers comparative effectiveness research and patients' preferences may help inform GRADE recommendations: an application to rheumatoid arthritis treatment recommendations. J Clin Epidemiol 2018;93:56-65. doi: 10.1016/j.jclinepi.2017.10.003.

16. Fain KM, Yu T, Li T, et al. Evidence selection for a prescription drug's benefit-harm assessment: challenges and recommendations. J Clin Epidemiol 2016;74:151-7.

17. MacLean S, Mulla S, Akl EA, et al. Patient Values and Preferences in Decision Making for Antithrombotic Therapy: A Systematic Review: Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. Chest 2012;141:e1S-e23S.

18. Richards T, Schroter S, Price A, Godlee F. Better together: patient partnership in medical journals. BMJ 2018;362:k3798.

19. Schünemann HJ. Interpreting GRADES's levels of certainty or quality of the evidence: GRADE for statisticians, considering review information size or less emphasis on imprecision? J Clin Epidemiol 2016;75:6-15.

20. Efron B, Tibshirani RJ. An Introduction to the Bootstrap, 1st Edition Chapman and Hall/CRC. New York, NY 1993

21. Spiegelhalter DJ. Myles JP, Jones DR, Abrams KR. An introduction to Bayesian methods in health technology assessment. BMJ 1999;319: 508–512.

22. Zhang Y, Alonso Coello P, et al. GRADE Guidelines: 19. Assessing the certainty of evidence in the importance of outcomes or values and preferences -- Risk of bias and indirectness. J Clin Epidemiol 2018.

23. Smith CM. Origin and uses of primum non nocere—above all, do no harm! J Clin Pharmacol 2005;45:371-7. doi: 10.1177/0091270004273680.

24. Schunemann HJ. Guidelines 2.0: do no net harm-the future of practice guideline development in asthma and other diseases. Current allergy and asthma reports 2011;11:261-8.

25. Krumholz HM, Lin Z, Keenan PS, et al. Relationship between hospital readmission and mortality rates for patients hospitalized with acute myocardial infarction, heart failure, or pneumonia. JAMA 2013;309:587-93.

26. Spencer FA, Prasad M, Vandvik PO, et al. Longer versus shorter duration dual-antiplatelet therapy after drug-eluting stent placement: a systematic review and meta-analysis. Ann Intern Med 2015;163:118-26. doi: 10.7326/M15-0083.

27. Vandvik PO, Lincoff AM, Gore JM, et al. Primary and secondary prevention of cardiovascular disease: Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. Chest 2012;141:e637S-e666S. doi: 10.1378/chest.11-2306.

28. Levine GN, Bates ER, Blankenship JC, et al; American College of Cardiology Foundation; American Heart Association Task Force on Practice Guidelines; Society for Cardiovascular

Angiography and Interventions. 2011 ACCF/AHA/SCAI Guideline for Percutaneous Coronary Intervention. A report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines and the Society for Cardiovascular Angiography and Interventions. J Am Coll Cardiol 2011;58:e44-122. doi: 10.1016/j.jacc.2011.08.007.

29. DynaMed Plus [Internet]. Ipswich (MA): EBSCO Information Services. 1995-2016. Record No. 114099, Heart failure with reduced ejection fraction; [updated 2016 Oct 07, cited 2016 Nov 27]; [about 101 screens]. Available from http://www.dynamed.com/topics/dmp~AN~T114099/Heart-failure-with-reduced-ejection-fraction#Sacubitril-valsartan--ARNI-.

30. Database of Evidence Profiles [Internet]. Alper B, Colpitts C, Lang E, et al. Should sacubitril-valsartan vs. ACE inhibitor or ARB be used for symptomatic patients with heart failure and reduced ejection fraction? [updated 2015 Nov 17; cited 2016 Nov 27]. Available from http://dbep.gradepro.org/profile/1FEB6C9D-1F7A-EAE1-AEC5-564B423DF770.

31. McMurray JJ, Packer M, Desai AS, et al; PARADIGM-HF Investigators and Committees. Angiotensin-neprilysin inhibition versus enalapril in heart failure. N Engl J Med 2014;371:993-1004. doi: 10.1056/NEJMoa1409077.

32. WRITING COMMITTEE MEMBERS; Yancy CW, Jessup M, Bozkurt B, et al. 2016 ACC/AHA/HFSA Focused Update on New Pharmacologic Therapy for Heart Failure: An Update of the 2013 ACCF/AHA Guideline for the Management of Heart Failure: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Heart Failure Society of America. Circulation 2016;134:e282-93. doi: 10.1161/CIR.0000000000000435.

33. Ponikowski P, Voors AA, Anker SD, Bueno H, Cleland JG, Coats AJ, et al; Authors/Task Force Members; Document Reviewers. 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: The Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC). Developed with the special contribution of the Heart Failure Association (HFA) of the ESC. Eur J Heart Fail 2016;18:891-975. doi: 10.1002/ejhf.592.

34. Scottish Intercollegiate Guidelines Network (SIGN). National clinical guideline for management of chronic heart failure. SIGN 2016 Mar. SIGN publication no. 147.

35. National Institute for Health and Care Excellence (NICE). Sacubitril valsartan for treating symptomatic chronic heart failure with reduced ejection fraction. Technology appraisal guidance. Published: 27 April 2016. TA388.

36. Fox K, Ford I, Steg PG, Tendera M, Ferrari R; BEAUTIFUL Investigators. Ivabradine for patients with stable coronary artery disease and left-ventricular systolic dysfunction (BEAUTIFUL): a randomised, double-blind, placebo-controlled trial. Lancet 2008;372:807-16. doi: 10.1016/S0140-6736(08)61170-8.

37. Swedberg K, Komajda M, Bohm M, et al; SHIFT Investigators. Ivabradine and outcomes in chronic heart failure (SHIFT): a randomised placebo-controlled study. Lancet 2010;376:875-85. doi: 10.1016/S0140-6736(10)61198-1.

38. ClinicalTrials.gov. Effects of Ivabradine on Cardiovascular Events in Patients With Moderate to Severe Chronic Heart Failure and Left Ventricular Systolic Dysfunction. A Three-year International Multicentre Study (SHIFT). NCT02441218. Last Updated May 9, 2016.

39. ClinicalTrials.gov. The BEAUTIFUL Study: Effects of Ivabradine in Patients With Stable Coronary Artery Disease and Left Ventricular Systolic Dysfunction. NCT00143507. Last Updated June 22, 2015.

40. Writing Committee Members, Yancy CW, Jessup M, Bozkurt B, Butler J, Casey DE Jr, et al. 2016 ACC/AHA/HFSA Focused Update on New Pharmacological Therapy for Heart Failure: An Update

of the 2013 ACCF/AHA Guideline for the Management of Heart Failure: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Heart Failure Society of America. Circulation 2016;134:e282-93. doi: 10.1161/CIR. 0000000000000435.

41. Ponikowski P, Voors AA, Anker SD, Bueno H, Cleland JG, Coats AJ, et al; Authors/Task Force Members. 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: The Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC). Developed with the special contribution of the Heart Failure Association (HFA) of the ESC. Eur Heart J 2016;37:2129-200. doi: 10.1093/eurheartj/ehw128.

42. National Institute for Health and Care Excellence (NICE). Ivabradine for treating chronic heart failure. Technology appraisal guidance. Published: 28 November 2012. TA267.

43. Scottish Intercollegiate Guidelines Network (SIGN). Management of chronic heart failure: A national clinical guideline. Published: March 2016. SIGN publication no. 147. http://www.sign.ac.uk/assets/sign147.pdf

44. National Collaborating Centre for Cancer. Myeloma: diagnosis and management. NICE Guideline 35, Full guideline, February 2016. https://www.nice.org.uk/guidance/ng35/evidence/full-guideline-2306487277

45. Giardis M, Busani S, Damiani E, et al. Effect of conservative vs conventional oxygen therapy on mortality among patients in an intensive care unit: The Oxygen-ICU Randomized Clinical Trial. JAMA 2016;316:1583-1589. doi: 10.1001/jama.2016.11993.