The 'Surprise question' in heart failure: a prospective cohort study

Valentina Gonzalez-Jaramillo [1,2] Luisa Fernanda Arenas Ochoa,[3] Clara Saldarriaga,[3,4] Alicia Krikorian [5] John Jairo Vargas,[5,6] Nathalia Gonzalez-Jaramillo,[1,2] Steffen Eychmüller,[7] Maud Maessen[1,7]

(Check for updates)

ABSTRACT

Objective The Surprise Question (SQ) is a prognostic screening tool used to identify patients with limited life expectancy. We assessed the SQ's performance predicting 1-year mortality among patients in ambulatory heart failure (HF) clinics. We determined that the SQ's performance changes according to sex and other demographic (age) and clinical characteristics, mainly left ventricular ejection fraction (LVEF) and the New York Heart Association (NYHA) functional classifications.

Methods We conducted a prospective cohort study in two HF clinics. To assess the performance of the SQ in predicting 1-year mortality, we calculated the sensitivity, specificity, positive and negative likelihood ratios, and the positive and negative predictive values. To illustrate if the results of the SQ changes the probability that a patient dies within 1 year, we created Fagan's nomograms. We report the results from the overall sample and for subgroups according to sex, age, LVEF and NYHA functional class.

Results We observed that the SQ showed a sensitivity of 85% identifying ambulatory patients with HF who are in the last year of life. We determined that the SQ's performance predicting 1-year mortality was similar among women and men. The SQ performed better for patients aged under 70 years, for patients with reduced or mildly reduced ejection fraction, and for patients NYHA class III/IV.

Conclusions We consider the tool an easy and fast first step to identify patients with HF who might benefit from an advance care planning discussion or a referral to palliative care due to limited life expectancy.

Key messages

What was already known?
⇒ The performance of the Surprise Question (SQ) screening tool to predict 1-year mortality had been assessed among inpatient populations with heart failure (HF) and populations with HF in an emergency setting.

What are the new findings?
⇒ The SQ's performance predicting 1-year mortality in an ambulatory setting, a HF clinic, where patients with HF are more stable.
⇒ The SQ's performance predicting 1-year mortality for a population with HF according to patients' sex, age, New York Heart Association functional class and left ventricular ejection fraction.

What is their significance?

Clinical:
⇒ The SQ's new psychometric profile allows for determining the appropriateness of its use in clinical practice at HF clinics and its predictive value for subgroup populations.

Research:
⇒ This study contributes to filling gaps in the knowledge of the SQ's performance in patients with HF in the ambulatory setting.

INTRODUCTION

People living with heart failure (HF), especially those in advanced stages of the disease, might present with uncontrolled symptoms such as shortness of breath, pain, sleep disorders and fatigue[1] ; they also often suffer from comorbidities, such as depression and anxiety.[2] Advance care planning (ACP) has proven to improve the quality of life and patient satisfaction with end-of-life care for patients with HF by promoting their autonomy concerning medical decisions.[3] Identifying those with HF who are in the last year of their life is paramount to guide discussions and determine other strategies that are part of ACP, including possible referrals to specialised palliative care (PC).

The Surprise Question (SQ) is a prognostic screening tool used to identify patients with limited life expectancy. The

SQ is a reflective question a physician or other healthcare providers ask themselves about a patient's prognosis: 'would I be surprised if this patient dies within the next 12 months?' An SQ is positive (+SQ) if the healthcare provider's answer is 'no, I would not be surprised'. There are also versions of the SQ within the context of 3 or 6 months. However, 1 year is the most common version.

The SQ's performance has been evaluated for patients with both oncological[4–7] and non-oncological diseases, including chronic kidney disease[8] and chronic obstructive pulmonary disease.[9] For patients with oncological diseases, the sensitivity of the 1-year SQ has varied from 58% to 85%.[4–7] For patients with non-oncological diseases, the performance also varies even within the same disease.[8–10] Despite the heterogeneity among performance measures, two systematic reviews suggest that, in general, the performance of the SQ predicting 1-year mortality is better for patients with oncological disease than for patients with non-oncological disease.[10 11] Recently, the SQ's performance was assessed for an inpatient population with HF[12] and for a population with HF in an emergency setting.[13] The results were promising with a sensitivity of 85% for hospitalised patients and 79% for the emergency setting. However, the SQ's performance in ambulatory settings, such as HF clinics where patients are more stable, is unknown. Furthermore, understanding the SQ's performance for men and women is important to ensure equity in the delivery of ACP and PC. To our knowledge, the SQ's performance for patients with HF, stratified according to the patient's sex, has not yet been reported. Finally, given that HF clinic admission criteria and samples vary, it is important to understand the SQ's performance according to other demographic and clinical characteristics to increase the generalisability of the results. Performance of the SQ is likely to differ across subgroups due to previous knowledge of the staff answering the question. For example, evidence has shown a trend of increasing mortality rates with increasing New York Heart Association (NYHA) stage,[14] increasing age[15] and decreasing left ventricular ejection fraction (LVEF).[16 17]

Therefore, our objectives included (1) assessing the performance of the SQ predicting 1-year mortality among patients in ambulatory HF clinics, and (2) assessing whether performance changes according to sex, age, NYHA classification and LVEF category.

METHODS

This study was conducted and reported in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology guidelines.[18]

Study design and setting

The prospective cohort included 174 ambulatory patients with HF who were recruited from two HF clinics in Medellín, Colombia between November 2017 and November 2018. One-year vital status was determined by consulting the national mortality register in Colombia. Both clinics are part of tertiary care institutions that are referral centres for patients with cardiovascular disease. They offer comprehensive, multidisciplinary care that includes clinical follow-up by HF cardiologists, nursing education and telephone follow-up, cardiac rehabilitation, and a psychoeducational programme for patients and their families.

Participants

Patients were potentially eligible for the study if they were 18 years or older and existing patients at the HF clinic who had at least two prior consultations. Since the first two consultations provide the cardiologist an opportunity to optimise treatment if necessary and to get to know patients under optimal treatment circumstances according to clinical guidelines, we did not enrol newer HF clinic patients. We enrolled consecutive eligible patients in the study. There were no exclusion criteria.

Ethical aspects

We conducted the study in accordance with the ethical guidelines from the Declaration of Helsinki[19]; our study was approved by the research ethics committees of the institutions involved in the study. Informed consent was collected before participants enrolled in the study.

Data sources and measurements

We obtained sociodemographic characteristics (age, sex and marital status) from electronic medical records, along with values of the clinical variables: LVEF, number of hospitalisations in the last year, presence of cardiac implantable devices, NYHA functional class, comorbidities and current medications. Comorbidities included clinical depression, atrial fibrillation, type 2 diabetes mellitus, kidney disease, lung disease, coronary artery disease, obstructive sleep apnoea and hypothyroidism. Current medications included ACE inhibitors, beta-blockers and angiotensin receptor blockers.

When a patient met the eligibility requirements for study inclusion, the treating cardiologist answered the SQ for that patient. For patients that the cardiologist would not be surprised if they died within the next year, we coded as +SQ. For patients that the cardiologist would be surprised if they died within the next year, we coded as a negative SQ (−SQ).

Statistical methods

To describe the study sample, we used mean and SD to summarise continuous variables in case of normal distribution. In cases of skewed distribution, we used median and IQR. We assessed normality using Q–Q plots. We summarised categorical variables as frequencies and percentages.

To assess the SQ's performance predicting 1-year mortality, we calculated the sensitivity and specificity, as well as the SQ's positive (+LR) and negative (–LR) likelihood ratios and the positive (PPV) and negative (NPV) predictive values (online supplemental tables 1 and 2). We interpreted the effect of the +LR on the likelihood of dying within 1 year based on the following classifications: no change if +LR=1; minimal increase if +LR between 1 and 2; small increase if +LR between 2 and 5; moderate increase if +LR between 5 and 10; and substantial increase if +LR >10.[20]

To illustrate how the result of the SQ changes the probability that a patient dies within 1 year, we created Fagan's nomograms.[21] We also conducted a univariable regression to assess the relation between a +SQ and 1-year mortality.

We performed subgroup analysis, comparing groups according to sex, age, LVEF and NYHA functional class. We created a categorical variable for median age (70 years) and another for LVEF (reduced LVEF ≤40%; mildly reduced LVEF 41%–49%; and preserved LVEF ≥50%).[22]

We performed all analyses with STATA V.15 (Stata Corp, College Station, Texas, USA).

RESULTS

Participants
Of the 184 patients who met the inclusion criteria and were potentially eligible, 178 consented to participate in the study. Among these 178 participants, 4 were excluded because their 1-year vital status was unknown (figure 1).

Table 1 shows baseline demographic and clinical characteristics of the study's 174 participants. The sample had a median age of 70 (58–77), was predominantly male, had reduced LVEF and NYHA class II.

The prevalence of a +SQ was 48%. After 1 year, 20 patients had died, giving an overall mortality rate of 12%.

Performance of the SQ predicting 1-year mortality
After 1 year, mortality among those with a +SQ was 21%; mortality among those with a –SQ was 3% (p<0.001). Participants with a +SQ had 7.5 times higher odds of death at 1 year compared with those with a –SQ (OR 7.6, 95% CI 2.1 to 26.9).

The +LR is the probability that patients with +SQ will die within 1 year divided by the probability that patients with +SQ will be alive in 1 year.[20] The +LR of the SQ was 1.98. The +SQ was nearly twice as likely for patients who died within 1 year than it was for patients who were alive in 1 year (figure 2). According to the classification of the +LR's effect on the likelihood of dying within 1 year, it is a minor increase. The 1-year mortality rate for our study was 12%. With the pretest probability of a patient dying within 1 year at 12% and a +LR of 1.98, the post-test probability of dying within 1 year is 20% (figure 3). A +SQ increased the probability that an ambulant patient with HF died within 1 year by 8 percentage points. The –LR was 0.26. The –SQ was nearly four times more likely assigned to participants who were alive in 1 year than it was for patients who died within 1 year.

For our sample, the probability that the SQ correctly identified an individual who would die in the course of a year, the sensitivity was 85% (95% CI 69% to 100%). The probability that the SQ would correctly identify an individual who would survive over a year, the specificity was 57% (95% CI 49% to 65%) (figure 2).
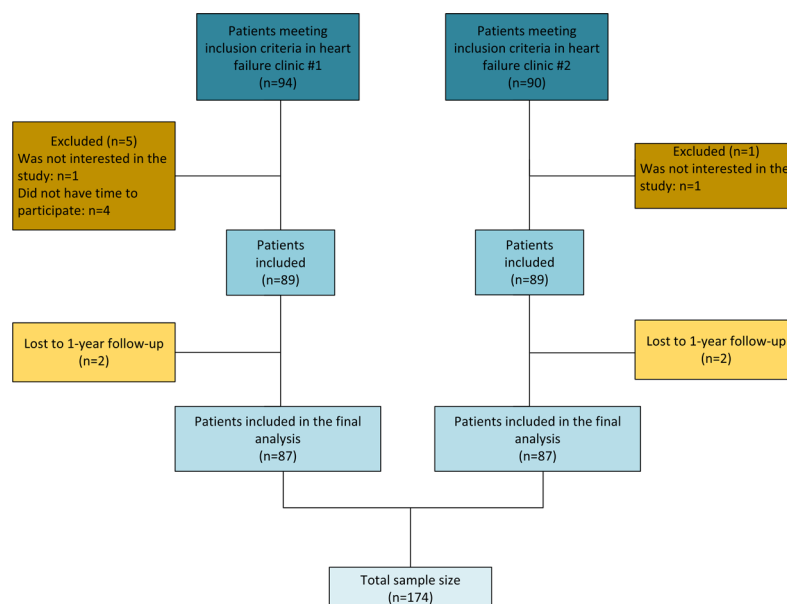

Figure 1 Flow chart of the patients included in the study.

Table 1 Clinical and demographic characteristics of study participants

	Patients included (n=174)
Age	70 (58–77)
Sex	
Women	78 (44.8%)
Men	96 (55.2%)
Marital status	
Single	25 (14.4%)
Married	91 (52.3%)
Divorced/separated	11 (6.3%)
Widow/widower	47 (27.0%)
Aetiology	
Ischaemic	62 (35.6%)
Hypertensive	15 (8.6%)
Idiopathic	69 (39.7%)
Valvular	22 (12.6%)
Toxic	6 (3.5%)
LVEF (%)	32 (25–44)
Classification according to LVEF	
HFrEF	116 (66.7%)
HFmrEF	28 (16.1%)
HFpEF	30 (17.2%)
NYHA functional class	
I	57 (32.8%)
II	82 (47.1%)
III	34 (19.5%)
IV	1 (0.6%)
Comorbidities	
Type 2 diabetes mellitus	55 (31.6%)
Chronic kidney disease	84 (48.3%)
Lung disease	25 (14.4%)
Coronary artery disease	64 (36.8%)
Atrial fibrillation	51 (29.3%
Depression	18 (10.3%)
Anxiety	7 (4.0%)
Obstructive sleep apnoea	13 (7.5%)
Obesity	20 (11.5%)
Hospitalisations in the last year	2 (1–2)
Implantable cardioverter defibrillator	42 (24.1%)
+SQ	83 (47.7%)
Mortality	20 (11.5%)
Medication	
Beta-blockers	164 (94.3%)
ACE inhibitors	136 (78.2%)
ARBs	130 (74.7%)

Data presented as number of patients (%) for categorical data or as median (IQR) for continuous data.

Those patients for whom the cardiologist would not be surprised if the patient died within the next year were coded as a +SQ.

ARBs, angiotensin receptor blockers; HFmrEF, heart failure with mildly reduced ejection fraction; HFpEF, heart failure with preserved ejection fraction; HFrEF, heart failure with reduced ejection fraction; LVEF, left ventricular ejection fraction; SQ, Surprise Question.

Subgroup analyses

We compared the SQ's performance between women and men. Whereas other parameters were similar, the SQ's sensitivity was 5 percentage points higher for women. With higher sensitivity, higher specificity and higher +LR, the SQ performed better for participants younger than 70 years (figure 2). Regarding LVEF's classification, sensitivity, NPV and –LR were better for patients with reduced LVEF. Specificity, PPV and +LR were better for patients with mildly reduced LVEF. We observed the SQ's worst performance for patients with preserved LVEF (figure 2). Among patients classified as NYHA III/IV, the SQ's sensitivity was perfect (100%). This subgroup also had the best values of PPV and NPV. However, the SQ's specificity was very low (31%) for patients at NYHA III/IV classification (figure 2).

Based on Fagan's nomograms, we accounted for the clinical application of the +LR. We observed the biggest probability changes for patients with +SQ dying within 1 year for those aged under 70 years and for patients with mildly reduced LVEF. Having a +SQ increased the probability of dying within 1 year by 12 and 14 percentage points, respectively (online supplemental figures 1–8). According to the +LR's effect on the likelihood of dying within 1 year, the +LRs showed a small increase in the likelihood among women; those aged under 70 years and patients with mildly reduced LVEF. In the remaining subgroups, +LR showed a minimal increase in the likelihood of dying within 1 year.

DISCUSSION
Key results
Our primary objective was to assess the SQ's performance predicting 1-year mortality among patients in ambulatory HF clinics. With a sensitivity of 85%, the SQ is a good tool to screen ambulatory HF clinic patients who might be in the last year of life. The SQ's performance predicting 1-year mortality was similar among women and men. We also assessed the SQ's performance according to sex and other demographic (age) and clinical characteristics. The SQ performed better for patients aged under 70 years; for those with reduced or mildly reduced ejection fraction; and for patients at NYHA III/IV classification. For the whole sample and the different subgroups, +LRs showed a minor or a small increase in the likelihood of a patient dying within 1 year, which is not good enough to consider that a patient has a life expectancy limited to 1 year.

Performance of the SQ by subgroup
The risk factors for developing HF differ between men and women, as do responses to treatment, symptom burden, and comorbidities due to both biological and cultural factors.[23] Because of this, sex-specific results should be presented in research.[24] Accounting for sex, determining survival rates for patients with HF has been inconclusive. Initially, the Framingham study showed better survival rates after HF diagnosis for women than for men.[25] Later, other studies suggested worse survival rates for women,[15] which was supported

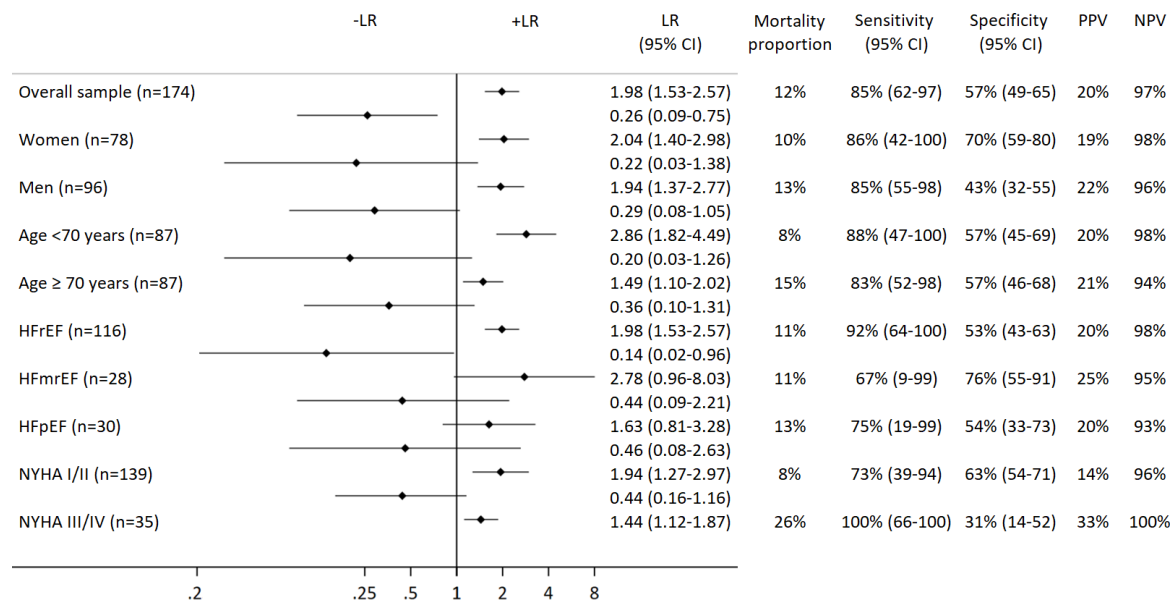|	-LR	+LR	LR (95% CI)	Mortality proportion	Sensitivity (95% CI)	Specificity (95% CI)	PPV	NPV
Overall sample (n=174)			1.98 (1.53-2.57)	12%	85% (62-97)	57% (49-65)	20%	97%
			0.26 (0.09-0.75)					
Women (n=78)			2.04 (1.40-2.98)	10%	86% (42-100)	70% (59-80)	19%	98%
			0.22 (0.03-1.38)					
Men (n=96)			1.94 (1.37-2.77)	13%	85% (55-98)	43% (32-55)	22%	96%
			0.29 (0.08-1.05)					
Age <70 years (n=87)			2.86 (1.82-4.49)	8%	88% (47-100)	57% (45-69)	20%	98%
			0.20 (0.03-1.26)					
Age ≥ 70 years (n=87)			1.49 (1.10-2.02)	15%	83% (52-98)	57% (46-68)	21%	94%
			0.36 (0.10-1.31)					
HFrEF (n=116)			1.98 (1.53-2.57)	11%	92% (64-100)	53% (43-63)	20%	98%
			0.14 (0.02-0.96)					
HFmrEF (n=28)			2.78 (0.96-8.03)	11%	67% (9-99)	76% (55-91)	25%	95%
			0.44 (0.09-2.21)					
HFpEF (n=30)			1.63 (0.81-3.28)	13%	75% (19-99)	54% (33-73)	20%	93%
			0.46 (0.08-2.63)					
NYHA I/II (n=139)			1.94 (1.27-2.97)	8%	73% (39-94)	63% (54-71)	14%	96%
			0.44 (0.16-1.16)					
NYHA III/IV (n=35)			1.44 (1.12-1.87)	26%	100% (66-100)	31% (14-52)	33%	100%

.2   .25 .5   1   2   4   8

Figure 2 Performance of the SQ in predicting 1-year mortality among patients in ambulatory HF clinics. HFmrEF, heart failure with mildly reduced ejection fraction; HFpEF, heart failure with preserved ejection fraction; HFrEF, heart failure with reduced ejection fraction; LR, likelihood ratio; NPV, negative predictive value; NYHA, New York Heart Association; PPV, positive predictive value; SQ, Surprise Question.

by women presenting with HF when they are older and have more comorbidities.[23] However, the most recent evidence suggests that age-adjusted mortality is similar between sexes.[26]

Among the subgroups, the SQ's best performance was for those aged under 70 years. The main difference between its performance for those older and younger than 70 years was specificity. An age over 70 years likely contributes to a +SQ response from cardiologists, leading to an increase in the proportion of false positives; thus, a reduction in specificity.

The best sensitivity was for the group of patients with reduced LVEF. With more evidence of effective therapies to reduce morbidity and mortality, this type of HF is the most studied and best understood.[27] In addition, evidence has shown that for patients with LVEF regardless of age, the lower the LVEF, the higher the mortality.[28] This might explain why the performance of an intuitive prognostic tool is better for this type of HF. The worst performance was among patients with preserved LVEF. This type of HF is not well understood, and there is no evidence of pharmacological therapy that decreases mortality for patients with preserved LVEF.[27 29] An intuitive prediction of mortality for this group of patients is especially complex because the relationship between LVEF and mortality is U shaped.[16 28] Above certain LVEF values, age-adjusted mortality increases, which is comparable with patients with LVEF between 30% and 35%.[16 28]

As for NYHA classifications, all patients classified at NYHA III/IV who died had a +SQ, which led to a sensitivity of 100%. However, for patients at NYHA III/IV functional class who survived 1 year, the majority also had a +SQ. Perhaps due to previous knowledge that

mortality increases with increasing NYHA functional class, cardiologists are more likely to assign +SQ to patients with HF in more advanced stages of NYHA,[14] which is similar to what happens with older patients with HF.

Comparison with previous studies
Previously, the SQ's performance predicting 1-year mortality had been evaluated for patients with HF in emergency[13] and inpatient settings.[12] In the emergency department, the SQ has a sensitivity of 79% and a specificity of 57% when answered by emergency physicians.[13] For hospitalised patients, the SQ has a sensitivity of 85% and a specificity of 59% when answered by cardiologists.[12] We found that the SQ's sensitivity (85%) and specificity (57%) for outpatient settings are equal to inpatient settings. Since the SQ's sensitivity and specificity were the same for decompensated (inpatients) and stable patients (outpatients), it suggests that the SQ's performance predicting 1-year mortality for patients with HF does not vary significantly. However, to compare the SQ's performance across settings and its interpretation in clinical practice for individual patients, we would have to compare LRs. No studies assessing the SQ's performance for populations with HF reported LRs. However, since LRs are calculated using sensitivity and specificity, we estimated them using other studies' reported test sensitivities and specificities. The +LR and −LR were similar across settings and represent minimal increases in the likelihood of dying within 1 year.
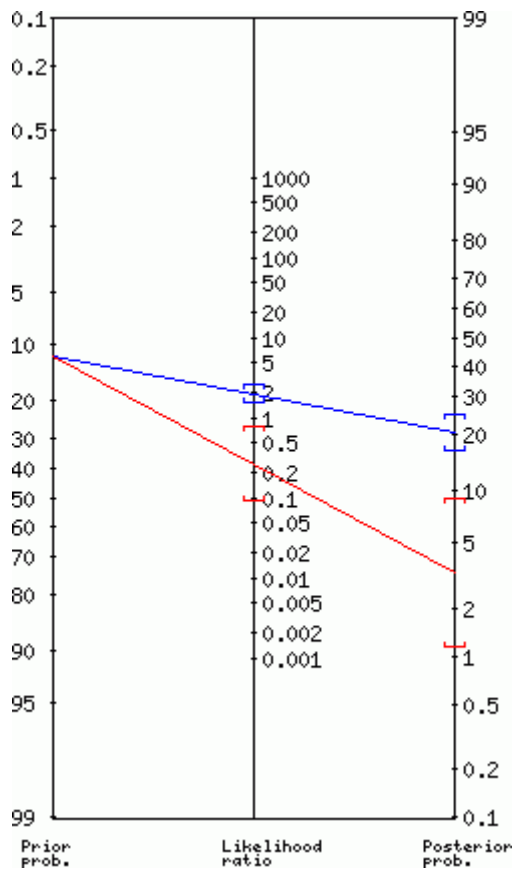
Figure 3 Fagan's nomogram for the overall sample. Based on a pretest probability of dying within 1 year of 12%, the blue line shows a post-test probability of a patient with a positive Surprise Question (+SQ) dying within a year of 20% (95% CI 17% to 25%) according to the positive likelihood ratio (+LR) of 1.98. A +SQ increases the probability of a patient dying within 1 year by 8 percentage points. The red line shows a post-test probability of a patient with a negative SQ (−SQ) dying within a year of 3% (95% CI 1% to 9%) according to the negative LR (−LR) of 0.22. A −SQ decreases the probability of a patient dying within 1 year by 9 percentage points.

Interpretations for clinical practice

When there is high test sensitivity, fewer false negatives occur, which increases the chance that patients with HF in need of ACP or PC will receive these services. Since determining sensitivity was the main criterion to evaluate the SQ's performance, we consider it is an acceptable prognostic screening tool for patients with HF. Yet, there remains a significant percentage of patients with HF within the last year of their life who might be left out (15% of the whole sample) and might benefit from ACP or PC. Both in this population and in the literature, the SQ predicts 1-year mortality well among patients with HF.[12 13] However, there are several aspects of the tool that users should be aware of before applying it in clinical practice.

First, the SQ's performance depends on the clinical expertise and experience of the staff using it, as well as their knowledge of the patient with HF. For example, Straw *et al* conducted a study that evaluated the SQ's

performance among an inpatient population with HF.[12] They showed that the SQ's sensitivity decreased from 85% when cardiologists used the tool to 75% when physicians in training used it; and from 90% when HF nurses used the SQ to 66% among non-specialist nurses. Second, screening strategies such as the SQ do not afford assessment of the complexity of patients' needs or the level of training that professionals should have to address ACP discussions or provide PC .

Finally, a major limitation of this screening tool is that it potentially excludes patients who will survive longer than 1 year but still would benefit from ACP or PC. Even in scenarios when the SQ predicts mortality well, using life expectancy as the sole criterion for assessing the need for ACP or PC is limiting. We consider the SQ can be used as a screening tool to initiate ACP or refer to PC for patients with a life expectancy of less than 1 year. However, we also consider the parallel use of needs assessment tools for patients with life expectancy of more than 1 year. For example, two recent systematic reviews of available tools to assess PC needs in patients with HF concluded that the Needs Assessment Tool: Progressive Disease-Heart Failure (NAT: PD-HF) was the most appropriate tool to determine the unmet needs of patients with HF.[30 31]

The NAT: PD-HF offers an alternative solution to several previously discussed points: it is not based on survival prognosis or severity factors, but it comprehensively evaluates different spheres.

The NAT: PD-HF is made of questions to determine a patient's physical and psychological symptoms, daily life activity limitations, spiritual concerns, financial or legal concerns, and health-related information needs. Although the NAT: PD-HF does not focus on PC needs or referrals for specialised PC services, it does assess patients' unmet needs and matches those needs with appropriate services, including specialised PC and other services. Finally, the NAT: PD-HF assesses the patients' and the caregivers' needs, including the caregiver's ability to take care of the patient.[32] However, in clinical situations where there is not enough time to gather answers to the NAT: PD-HF's comprehensive question sets, there is enough time for the clinician to ask themselves the singular SQ, which is better than no needs screening at all.

Strengths and limitations of this study

The low proportion of patients classified as NYHA III/IV might be a limit of the generalisability of our results. The risk of mortality increases with a higher NYHA classification.[33] Among our population, 80% of the patients were classified as NYHA I or II. As expected, the mortality rate was low compared with what has been reported in other HF clinics where mortality is around 30%.[12 13] However, as we conducted different subgroup analyses, including analysis according to NYHA functional class, we provide different analyses

that can be assessed according to each HF clinic population.

As most encounters between healthcare personnel and patients occur in ambulatory settings, a strength of this study is the contribution of the SQ's psychometric profile for ambulatory patients with HF. In addition, when screening for limited life expectancy and the need to initiate ACP, since the outpatient population are more stable patients, ACP needs may be overlooked. The systematic use of a tool such as the SQ could help identify patients with HF eligible for end-of-life care without being limited by having ACP discussion or making decisions within unstable medical contexts. Furthermore, to support better generalisation, this study provides substantial data regarding subcategories.

CONCLUSION

The SQ showed a good sensitivity predicting 1-year mortality for patients in ambulatory HF clinics. However, the likelihood of dying within 1 year increases little when having a +SQ. We suggest that the SQ can be used as a starting point to identify patients who might benefit from having an ACP discussion or a referral to PC due to limited life expectancy toward a patient's end of life.

ORCID iDs
Valentina Gonzalez-Jaramillo http://orcid.org/0000-0002-2470-7022
Alicia Krikorian http://orcid.org/0000-0003-2118-5692

REFERENCES

1 Bekelman DB, Havranek EP, Becker DM, *et al*. Symptoms, depression, and quality of life in patients with heart failure. *J Card Fail* 2007;13:643–8.
2 Mbakwem A, Aina F, Amadi C. *Expert Opinion*-depression in patients with Heart failure: is enough being done? *Card Fail Rev* 2016;2:110–2.
3 Schichtel M, Wee B, Perera R, *et al*. The effect of advance care planning on heart failure: a systematic review and meta-analysis. *J Gen Intern Med* 2020;35:874–84.
4 Hamano J, Morita T, Inoue S, *et al*. Surprise questions for survival prediction in patients with advanced cancer: a multicenter prospective cohort study. *Oncologist* 2015;20:839–44.
5 Moroni M, Zocchi D, Bolognesi D, *et al*. The 'surprise' question in advanced cancer patients: A prospective study among general practitioners. *Palliat Med* 2014;28:959–64.
6 Moss AH, Lunney JR, Culp S, *et al*. Prognostic significance of the "surprise" question in cancer patients. *J Palliat Med* 2010;13:837–40.
7 Downar J, Goldman R, Pinto R, *et al*. The "surprise question" for predicting death in seriously ill patients: a systematic review and meta-analysis. *Can Med Assoc J* 2017;189:E484–93.
8 Moss AH, Ganjoo J, Sharma S, *et al*. Utility of the "surprise" question to identify dialysis patients with high mortality. *Clin J Am Soc Nephrol* 2008;3:1379–84.
9 Tripp D, Janis J, Jarrett B, *et al*. How well does the surprise question predict 1-year mortality for patients admitted with COPD? *J Gen Intern Med* 2021. doi:10.1007/s11606-020-06512-8. [Epub ahead of print: 06 Jan 2021].
10 Downar J, Goldman R, Pinto R, *et al*. The "surprise question" for predicting death in seriously ill patients: a systematic review and meta-analysis. *CMAJ* 2017;189:E484–93.
11 White N, Kupeli N, Vickerstaff V, *et al*. How accurate is the 'Surprise Question' at identifying patients at the end of life? A systematic review and meta-analysis. *BMC Med* 2017;15:139.
12 Straw S, Byrom R, Gierula J, *et al*. Predicting one-year mortality in heart failure using the 'Surprise Question': a prospective pilot study. *Eur J Heart Fail* 2019;21:227–34.
13 Aaronson EL, George N, Ouchi K, *et al*. The surprise question can be used to identify heart failure patients in the emergency department who would benefit from palliative care. *J Pain Symptom Manage* 2019;57:944–51.
14 Ahmed A, Aronow WS, Fleg JL. Higher New York heart association classes and increased mortality and hospitalization in patients with heart failure and preserved left ventricular function. *Am Heart J* 2006;151:444–50.
15 Taylor CJ, Ordóñez-Mena JM, Roalfe AK, *et al*. Trends in survival after a diagnosis of heart failure in the United Kingdom 2000-2017: population based cohort study. *BMJ* 2019;364:l223.
16 Curtis JP, Sokol SI, Wang Y, *et al*. The association of left ventricular ejection fraction, mortality, and cause of death in stable outpatients with heart failure. *J Am Coll Cardiol* 2003;42:736–42.
17 Wehner GJ, Jing L, Haggerty CM, *et al*. Routinely reported ejection fraction and mortality in clinical practice: where does the nadir of risk lie? *Eur Heart J* 2020;41:1249–57.
18 von Elm E, Altman DG, Egger M, *et al*. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol* 2008;61:344–9.

19 Gilder S. World Medical association meets in Helsinki. *Br Med J* 1964;2:299–300.
20 Parikh R, Parikh S, Arun E, *et al*. Likelihood ratios: clinical application in day-to-day practice. *Indian J Ophthalmol* 2009;57:217–21.
21 Fagan TJ. Letter: nomogram for Bayes theorem. *N Engl J Med* 1975;293:257. doi:10.1056/NEJM197507312930513
22 Bozkurt B, Coats AJS, Tsutsui H, *et al*. Universal definition and classification of heart failure: a report of the heart failure Society of America, heart failure association of the European Society of cardiology, Japanese heart failure Society and writing Committee of the universal definition of heart failure: endorsed by the Canadian heart failure Society, heart failure association of India, cardiac Society of Australia and New Zealand, and Chinese heart failure association. *Eur J Heart Fail* 2021;23:352–80. doi:10.1002/ejhf.2115
23 Lam CSP, Arnott C, Beale AL, *et al*. Sex differences in heart failure. *Eur Heart J* 2019;40:3859–68.
24 Woodward M. Cardiovascular disease and the female disadvantage. *Int J Environ Res Public Health* 2019;16. doi:10.3390/ijerph16071165. [Epub ahead of print: 01 Apr 2019].
25 Ho KK, Anderson KM, Kannel WB, *et al*. Survival after the onset of congestive heart failure in Framingham heart study subjects. *Circulation* 1993;88:107–15.
26 Taylor CJ, Ordóñez-Mena JM, Jones NR, *et al*. National trends in heart failure mortality in men and women, United Kingdom, 2000-2017. *Eur J Heart Fail* 2021;23:3–12.
27 Groenewegen A, Rutten FH, Mosterd A, *et al*. Epidemiology of heart failure. *Eur J Heart Fail* 2020;22:1342–56.
28 Wehner GJ, Jing L, Haggerty CM, *et al*. Routinely reported ejection fraction and mortality in clinical practice: where does the nadir of risk lie? *Eur Heart J* 2020;41:1249–57.
29 Ponikowski P, Voors AA, Anker SD, *et al*. 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.
30 Remawi BN, Gadoud A, Murphy IMJ, *et al*. Palliative care needs-assessment and measurement tools used in patients with heart failure: a systematic mixed-studies review with narrative synthesis. *Heart Fail Rev* 2021;26:137–55.
31 Ament SM, Couwenberg IM, Boyne JJ, *et al*. Tools to help healthcare professionals recognize palliative care needs in patients with advanced heart failure: a systematic review. *Palliat Med* 2021;35:45–58.
32 Waller A, Girgis A, Davidson PM, *et al*. Facilitating needs-based support and palliative care for people with chronic heart failure: preliminary evidence for the acceptability, inter-rater reliability, and validity of a needs assessment tool. *J Pain Symptom Manage* 2013;45:912–25.
33 Dokainish H, Teo K, Zhu J, *et al*. Global mortality variations in patients with heart failure: results from the International congestive heart failure (INTER-CHF) prospective cohort study. *The Lancet Global Health* 2017;5:e665–72.

Contents

Supplementary Table 1. Two by two table

<table>
<tr><td></td><td></td><td colspan="2">**1-year mortality**</td><td>**Total**</td></tr>
<tr><td rowspan="4">**Surprise question**</td><td></td><td>+</td><td>-</td><td></td></tr>
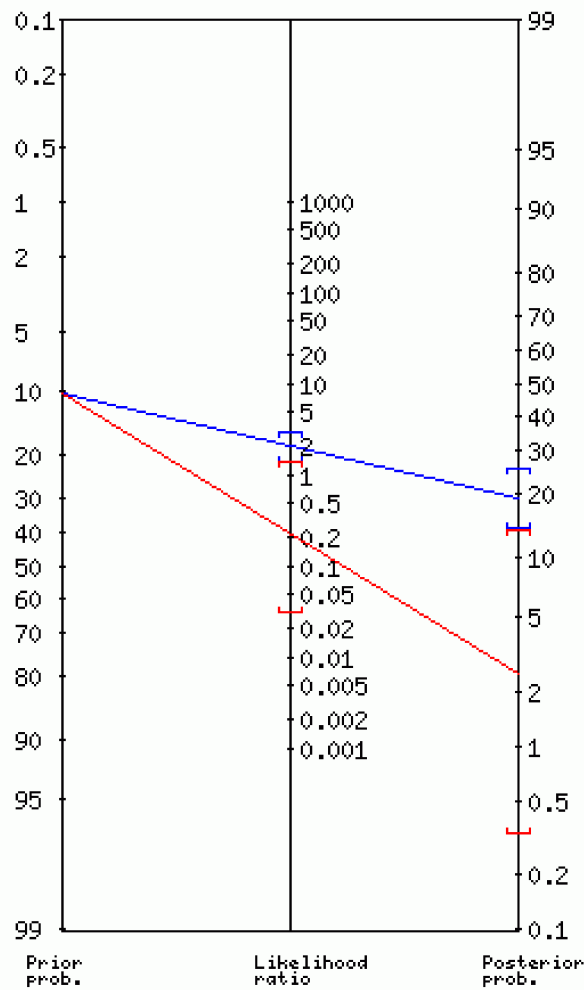<tr><td>+</td><td>17
a</td><td>66
b</td><td>83
a+b</td></tr>
<tr><td>-</td><td>3
c</td><td>88
d</td><td>91
c+d</td></tr>
<tr><td>**Total**</td><td>20
a+c</td><td>154
b+d</td><td>174</td></tr>
</table>

Supplementary Table 2. Performance of the SQ to predict 1-year mortality among ambulatory patients with HF

Sensitivity	0.85 a/(a+c)	Specificity	0.57 d/(d+b)
Positive Likelihood Ratio (+LR)	1.98 Sensitivity/1-Specificity	Negative Likelihood Ratio (-LR)	0.26 1-Sensitivity/Specificity
Positive Predictive Value (PPV)	0.20 a/(a+b)	Negative Predictive Value (NPV)	0.97 d/(c+d)

2

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)
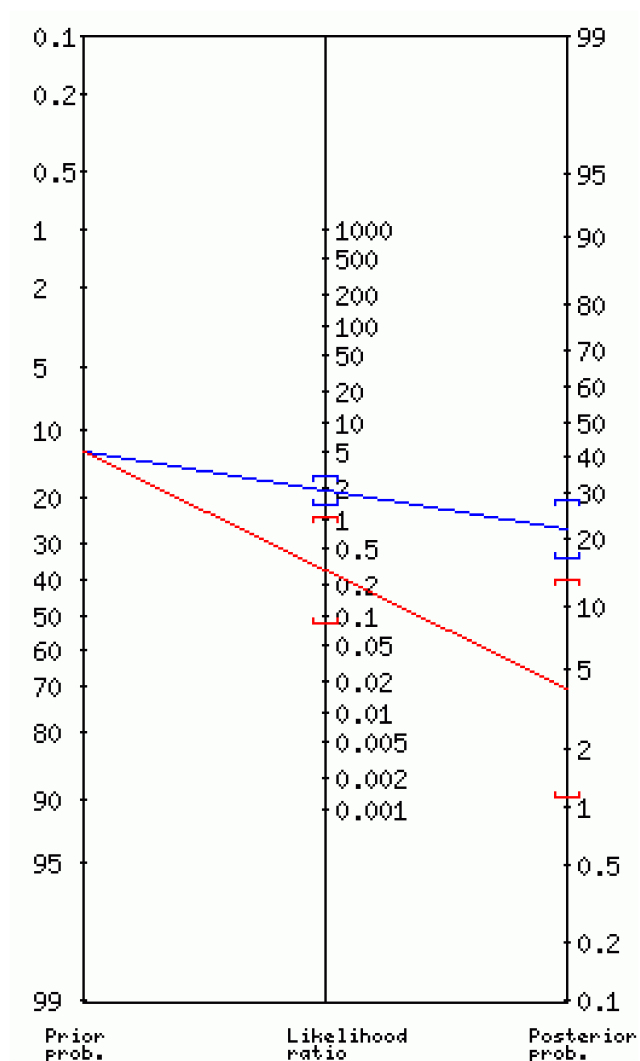
BMJ Support Palliat Care

Supplementary Figure 1. Fagan´s nomogram for women

Based on a pre-test probability of dying within 1 year of 10%, the blue line shows a post-test probability of a patient with a positive Surprise Question (+SQ) dying within 1 year of 19% (95% CI, 14%-25%) according to the +LR of 2.04. A +SQ increases the probability of a patient dying within 1 year by 9 percentage points.
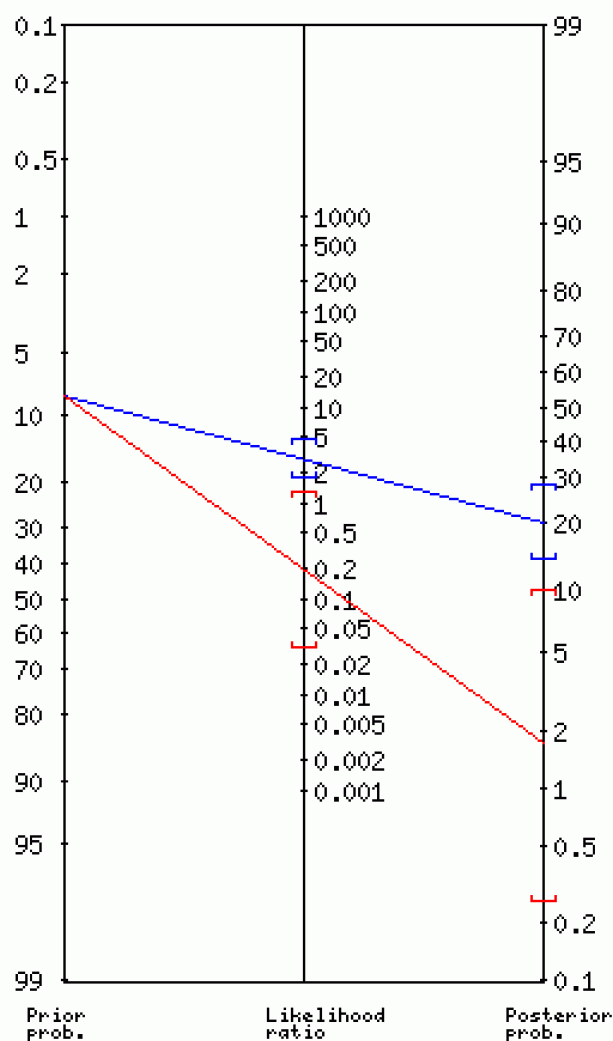
The red line shows a post-test probability of a patient with a negative SQ (−SQ) dying within one year of 2% (95% CI, 0%-14%) according to the negative likelihood ratio (−LR) of 0.22. A -SQ decreases the probability of a patient dying within 1 year by 8 percentage points.

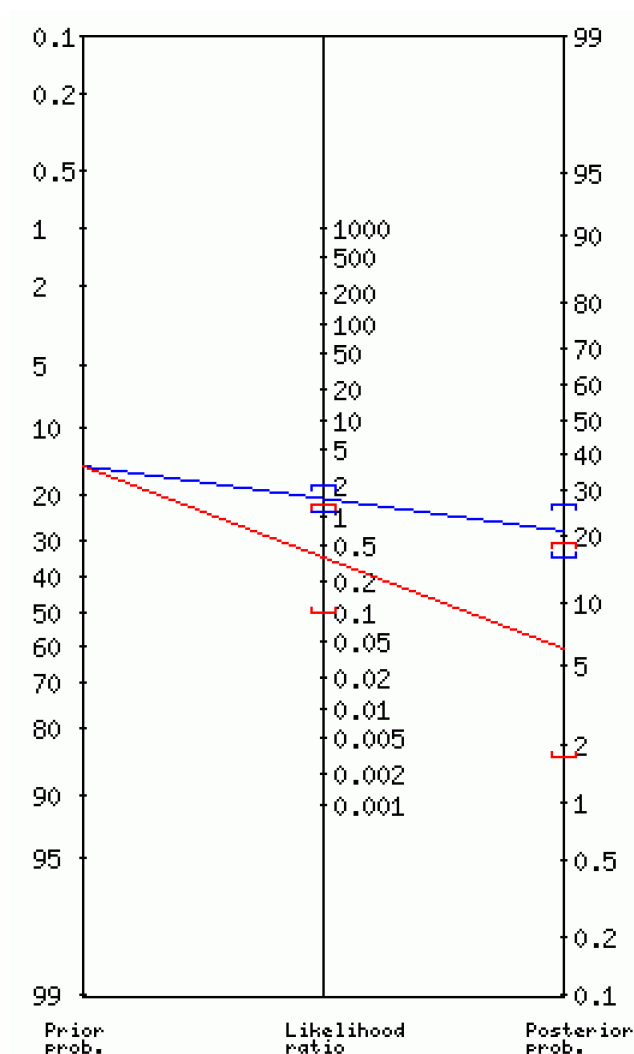**Supplementary Figure 2.** Fagan´s nomogram for men

Based on a pre-test probability of dying within 1 year of 13%, the blue line shows a post-test probability of a patient with a positive Surprise Question (+SQ) dying within a year of 22% (95% CI, 16%-28%) according to a positive likelihood ratio (+LR) of 1.94. A +SQ increases the probability of a patient dying within 1 year by 9 percentage points. The red line shows a post-test probability of a patient with a negative SQ (–SQ) dying within a year of 4% (95% CI, 1%-13%) according to the negative LR (–LR) of 0.29. A -SQ decreases the probability of a patient dying within 1 year by 9 percentage points.

4

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

BMJ Support Palliat Care

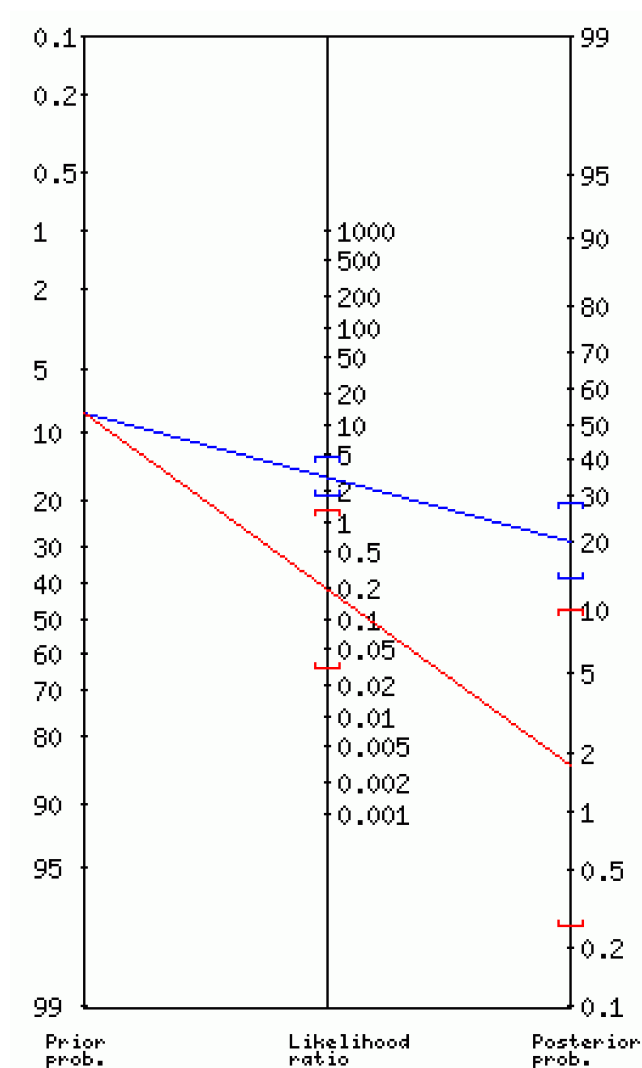**Supplementary Figure 3.** Fagan´s nomogram for those aged under 70

Based on a pre-test probability of dying within 1 year of 8%, the blue line shows a post-test probability of a patient with a positive Surprise Question (+SQ) dying within 1 year of 20% (95% CI, 14%-28%) according to the positive likelihood ratio (+LR) of 2.86. A +SQ increases the probability of a patient dying within 1 year by 12 percentage points. The red line shows a post-test probability of a patient with a negative SQ (−SQ) dying within a year of 2% (95% CI, 0%-10%) according to the negative LR (−LR) of 0.20. A -SQ decreases the probability of a patient dying within 1 year by 6 percentage points.

5

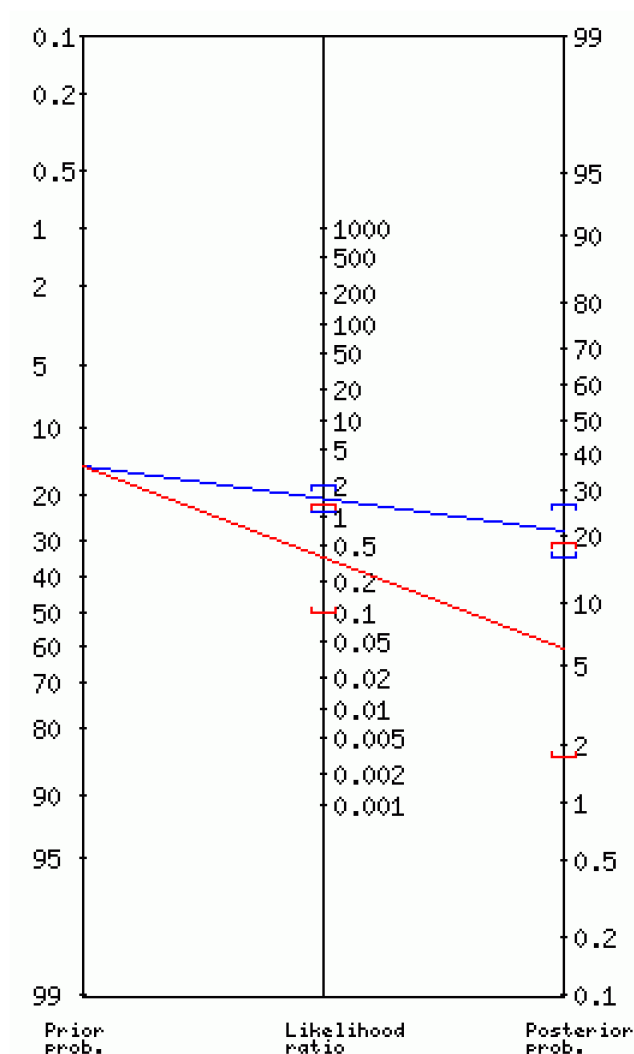**Supplementary Figure 4.** Fagan´s nomogram for those aged 70 or older

Based on a pre-test probability of dying within 1 year of 15%, the blue line shows a post-test probability of a patient with a positive Surprise Question (+SQ) dying within 1 year of 21% (95% CI, 16%-26%) according to the positive likelihood ratio (+LR) of 1.49. A +SQ increases the probability of a patient dying within 1 year by 6 percentage points. The red line shows a post-test probability of a patient with a negative SQ (–SQ) dying within a year of 6% (95% CI, 2%-19%) according to the negative LR (–LR) of 0.36. A -SQ decreases the probability of a patient dying within 1 year by 9 percentage points.

6

Supplementary Figure 5. Fagan´s nomogram for those with reduced ejection fraction (HFrEF)
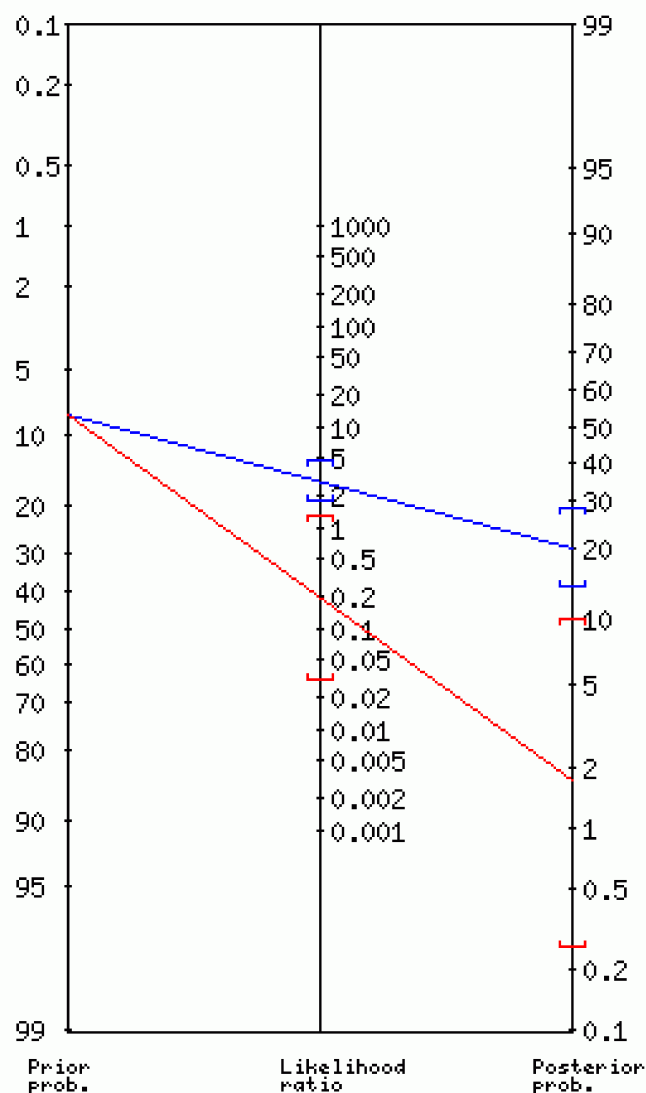

Based on a pre-test probability of dying within 1 year of 11%, the blue line shows a post-test probability of a patient with a positive Surprise Question (+SQ) dying within 1 year of 20% (95% CI, 16%-24%) according to the positive likelihood ratio (+LR) of 1.98. A +SQ increases the probability of a patient dying within 1 year by 9 percentage points. The red line shows a post-test probability of a patient with a negative SQ (–SQ) dying within a year of 2% (95% CI, 0%-11%) according to the negative LR (–LR) of 0.14. A -SQ decreases the probability of a patient dying within 1 year by 9 percentage points.

7

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

BMJ Support Palliat Care

Supplementary Figure 6. Fagan´s nomogram for those with mildly reduced ejection fraction (HFmrEF)
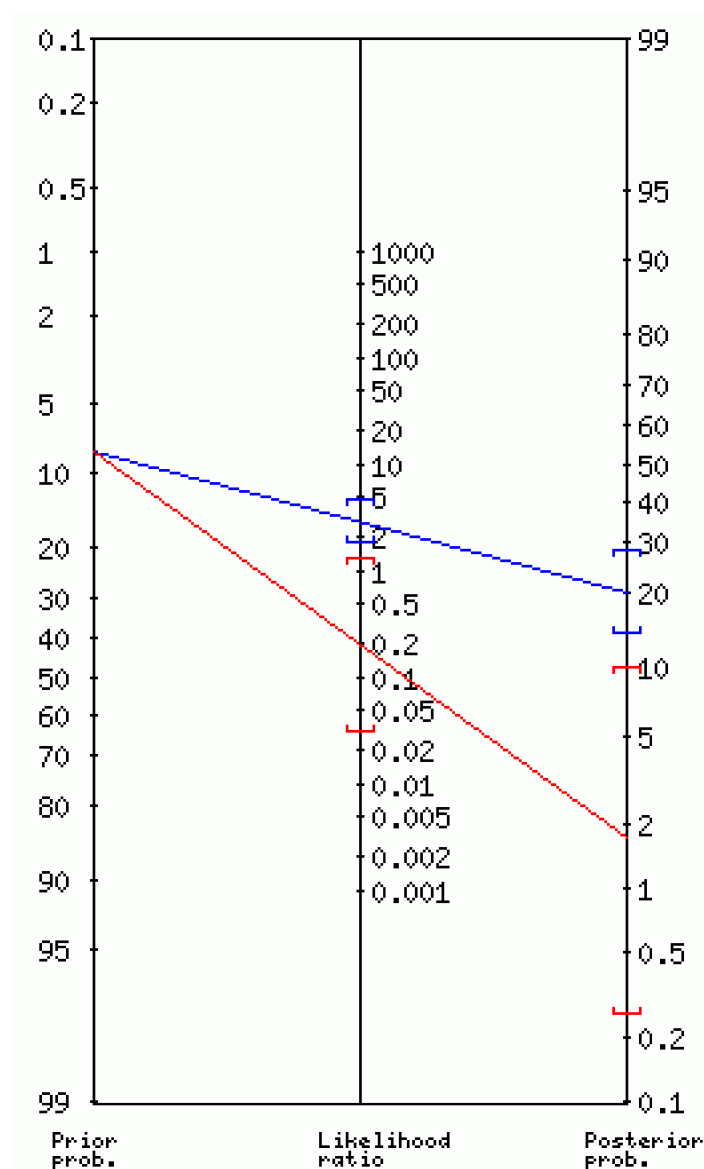

Based on a pre-test probability of dying within 1 year of 11%, the blue line shows a post-test probability of a patient with a positive Surprise Question (+SQ) dying within 1 year of 25% (95% CI, 10%-49%) according to the positive likelihood ratio (+LR) of 2.78. A +SQ increases the probability of a patient dying within 1 year by 14 percentage points. The red line shows a post-test probability of a patient with a negative SQ (–SQ) dying within a year of 5% (95% CI, 1%-21%) according to the negative LR (–LR) of 0.44. A -SQ decreases the probability of a patient dying within 1 year by 6 percentage points.

Supplementary Figure 7. Fagan´s nomogram for those with preserved ejection fraction (HFpEF)

Based on a pre-test probability of dying within 1 year of 13%, the blue line shows a post-test probability of a patient with a positive Surprise Question (+SQ) dying within 1 year of 20% (95% CI, 11%-34%) according to the positive likelihood ratio (+LR) of 1.63. A +SQ increases the probability of a patient dying within 1 year by 7 percentage points. The red line shows a post-test probability of a patient with a negative SQ (–SQ) dying within a year of 7% (95% CI, 1%-29%) according to the negative LR (–LR) of 0.46. A -SQ decreases the probability of a patient dying within 1 year by 6 percentage points.

9

Supplementary Figure 8. Fagan´s nomogram for those at NYHA class I/II

Based on a pre-test probability of dying within 1 year of 8%, the blue line shows a post-test probability of a patient with a positive Surprise Question (+SQ) dying within 1 year of 14% (95% CI, 10%-20%) according to the positive likelihood ratio (+LR) of 1.94. A +SQ increases the probability of a patient dying within 1 year by 6 percentage points. The red line shows a post-test probability of a patient with a negative SQ (–SQ) dying within a year of 4% (95% CI, 1%-9%) according to the negative LR (–LR) of 0.44. A -SQ decreases the probability of a patient dying within 1 year by 4 percentage points.

10