

Optimizing the G8 Screening Tool for Older Patients With Cancer: Diagnostic Performance and Validation of a Six-Item Version

CLAUDIA MARTINEZ-TAPIA,^a FLORENCE CANOUI-POITRINE,^{a,c} SYLVIE BASTUJI-GARIN,^{a,c,d} PIERRE SOUBEYRAN,^{e,h} SIMONE MATHOULIN-PELISSIER,^{i,j} CHRISTOPHE TOURNIGAND,^{b,e} ELENA PAILLAUD,^{a,f} MARIE LAURENT,^{a,f} ETIENNE AUDUREAU,^{a,c}
ON BEHALF OF THE ELCAPA STUDY GROUP

^aEA 7376 CEpiA (Clinical Epidemiology and Ageing Unit), IMRB, A-TVH DHU and ^bEC2M3 Unit, VIC DHU, Université Paris Est (UPE), Créteil, France; ^cPublic Health Department, ^dClinical Research Unit (URC Mondor), ^eDepartment of Medical Oncology, and ^fInternal Medicine and Geriatric Department, Assistance Publique Hôpitaux de Paris, Henri-Mondor Hospital, Créteil, France; ^gDepartment of Medical Oncology, Institut Bergonié, Bordeaux, France; ^hUniversité de Bordeaux, Bordeaux, France; ⁱINSERM U897, CIC-EC07, ISPED, Université de Bordeaux, France; ^jClinical and Epidemiological Research Unit, Institut Bergonié, Bordeaux, France

Disclosures of potential conflicts of interest may be found at the end of this article.

Key Words. Geriatric assessment • Cancer • Frailty • Sensitivity and specificity

ABSTRACT

Background. A multidimensional geriatric assessment (GA) is recommended in older cancer patients to inventory health problems and tailor treatment decisions accordingly but requires considerable time and human resources. The G8 is among the most sensitive screening tools for selecting patients warranting a full GA but has limited specificity. We sought to develop and validate an optimized version of the G8.

Patients and Methods. We used a prospective cohort of cancer patients aged ≥ 70 years referred to geriatricians for GA (2007–2012: $n = 729$ [training set]; 2012–2014: $n = 414$ [validation set]). Abnormal GA was defined as at least one impaired domain across seven validated tests. Multiple correspondence analysis, multivariate logistic regression, and bootstrapped internal validation were performed sequentially.

Results. The final model included six independent predictors for abnormal GA: weight loss, cognition/mood, performance

status, self-rated health status, polypharmacy (≥ 6 medications per day), and history of heart failure/coronary heart disease. For the original G8, sensitivity was 87.2% (95% confidence interval, 84.3–89.7), specificity 57.7% (47.3–67.7), and area under the receiver-operating characteristic curve (AUROC) 86.5% (83.5–89.6). The modified G8 had corresponding values of 89.2% (86.5–91.5), 79.0% (69.4–86.6), and 91.6% (89.3; 93.9), with higher AUROC values for all tumor sites and stable properties on the validation set.

Conclusion. A modified G8 screening tool exhibited better diagnostic performance with greater uniformity across cancer sites and required only six items. If these features are confirmed in other settings, the modified tool may facilitate selection for a full GA in older patients with cancer. *The Oncologist* 2016;21:188–195

Implications for Practice: Several screening tools have been developed to identify older patients with cancer likely to benefit from a complete geriatric assessment, but none combines appropriate sensitivity and specificity. Based on a large prospective cohort study, an optimized G8 tool was developed, combining a systematic statistical approach with expert judgment to ensure optimal discriminative power and clinical relevance. The improved screening tool achieves high sensitivity, high specificity, better homogeneity across cancer types, and greater parsimony with only six items needed, facilitating selection for a full geriatric assessment.

INTRODUCTION

A majority and growing number of cancers are diagnosed in patients aged 65 years or more [1]. Population aging in industrialized countries and the increase in cancer incidence with advancing age make the management of older patients with cancer a major public health challenge. Older patients vary widely regarding multiple aspects of

the aging process and health status. Individual treatment tailoring is therefore particularly important to ensure optimal efficacy and to minimize toxicity in this age group.

To assess the profile of each individual patient, a multidimensional geriatric assessment (GA) has been recommended since 2005 by the International Society of Geriatric Oncology

Correspondence: Etienne Audureau, M.D., Ph.D., Public Health Department, Hôpital Henri Mondor, 51 avenue du Maréchal de Lattre de Tassigny, 94010 Créteil, France. Telephone: 33-19813664; E-Mail: etienne.audureau@hmn.aphp.fr Received August 13, 2015; accepted for publication October 27, 2015; published Online First on January 13, 2016. ©AlphaMed Press 1083-7159/2016/\$20.00/0 <http://dx.doi.org/10.1634/theoncologist.2015-0326>

[2, 3]. The GA produces an inventory of health problems and an assessment of physical, psychosocial, and functional capabilities [4]. The GA has proved useful for characterizing health and functional impairments potentially associated with oncological outcomes [2, 5–7]. Nevertheless, it requires considerable time and human resources and is actually not required in all patients [6].

Screening tools have therefore been developed to discriminate between fit older patients who are likely to tolerate standard therapy and vulnerable patients who would benefit from specific treatment tailoring based on findings from a complete GA [8]. The ideal screening tool is easy to perform, requires little time, covers all the domains routinely assessed by geriatricians, and effectively separates fit from vulnerable patients. Although many tools have been described, none combines both appropriate sensitivity and specificity for predicting an abnormal GA [8, 9]. The G8, a screening tool specifically developed for older patients with cancer [10, 11], is among the most sensitive but lacks specificity [8, 9, 12]. Importantly, most screening instruments, including the G8, use items selected based on expert opinion and/or existing assessment tools (e.g., mini-nutritional assessment-short form [MNA-SF] in the G8) that were originally validated to detect outcomes other than GA impairment.

Here, our objectives were to evaluate the performance of the G8 in identifying older cancer patients likely to have abnormal GAs and to determine whether modifications to the G8 might improve this performance. We used a systematic statistical approach to simultaneously test the original G8 items and a set of additional, potentially relevant items in a large cohort of older patients with cancer.

MATERIALS AND METHODS

Patients

We analyzed data from the ELCAPA (ELderly CAncer PATients) study, a prospective open-cohort survey of consecutive patients aged 70 years or older with solid cancers or hematologic malignancies who were referred by an oncologist, surgeon, radiotherapist, or other specialist to one of two geriatric oncology clinics in teaching hospitals in the Paris urban area. A multidimensional GA was performed by geriatricians specialized in oncology. Then a multidisciplinary meeting was held to determine the best treatment strategy. For the present analysis (ELCAPA-07), we included patients for whom complete G8 and GA information were available in the electronic database. Patients enrolled from January 2007 to October 2012 were considered as the training set, whereas patients subsequently recruited until July 2014 were considered as the validation set. Informed consent was obtained from all study patients prior to inclusion. The study was approved by the appropriate ethics committee (Comité de Protection des Personnes [CPP] Ile-de-France I, France).

GA Reference Procedure

The GA consisted of a set of seven validated tests covering a variety of important health domains in older cancer patients [13] and consistent with the questionnaires and thresholds

used in the primary G8 validation study [11]. Abnormal GA was thus defined as an impaired score on at least one of the following tests: Activities of Daily Living (ADL $\leq 5/6$), Instrumental Activities of Daily Living (IADL $\leq 7/8$), Mini Mental State Examination (MMSE $\leq 23/30$), Mini Geriatric Depression Scale (Mini GDS ≥ 1), Mini Nutritional Assessment (MNA $\leq 23.5/30$), Cumulative Illness Rating Scale for Geriatrics (CIRS-G; at least one comorbidity grade 3 or 4), and Timed Up-and-Go test (> 20 seconds). Sensitivity analyses were performed to test the robustness of the results, by using an alternative cutoff of ≥ 2 impaired tests to define abnormal GA or by omitting the CIRS-G or ADL/IADL from the reference GA.

Candidate Items

We identified 22 candidate items for a modified G8, based on both the literature and clinical expertise, to avoid the overfitting to the training data set seen when item selection relies solely on statistical significance [14]. In addition to the G8 items (depression/dementia, body mass index [BMI], anorexia, weight loss, age, medications [> 3 per day], mobility, and self-rated health status), we selected 14 items routinely collected during geriatric evaluations and known to be clinically relevant for assessing older cancer patients: asthenia, incontinence, fall risk (single-leg stance time < 5 seconds), history of fall(s) in the past 6 months, Eastern Cooperative Oncology Group Performance Status (ECOG-PS), gender, living alone at home, metastatic status, and a selection of six comorbidities (diabetes, hypertension, heart failure and/or coronary heart disease [CHD], complete arrhythmia with atrial fibrillation [CAAF], chronic renal failure, and chronic respiratory failure).

Statistical Analysis

Training Set Analysis

First, we performed a descriptive univariate analysis to assess associations between candidate items and the reference GA, using χ^2 or Fisher's exact tests as appropriate and estimating the crude odds ratios (ORs) with their 95% confidence intervals. Receiver-operating curve (ROC) analysis was performed for the number of medications to assess alternative cutoffs to the one used in the original G8. Second, we used multiple correspondence analysis (MCA) to investigate correlations between candidate items and identify redundancies across conceptually close qualitative variables, thus helping to decide which variables should be combined, dichotomized, or omitted. Additional details are shown in supplemental online Figure 1.

Remaining candidate items were then entered into a multivariate logistic regression model, using a stepwise backward procedure to sequentially remove items based on a $p < .05$ level until the final model was obtained. Model discrimination was assessed by the area under the ROC curve (AUROC) and calibration by the Hosmer-Lemeshow χ^2 test. Regression coefficients were considered for use as weights to compute the final score. We rescaled (multiplied) and rounded them to the closest integer, using the algorithm described by Cole to find the optimal solution that both improved simplicity of use in the clinical setting and preserved initial model accuracy [15].

Table 1. Patient characteristics in the ELCAPA-07 cohort study

Characteristics	n (%)	
	Training set (n = 729)	Validation set (n = 414)
Male gender	387 (53.1)	197 (47.6)
Age in years, median (IQR)	80 (76–84)	81 (78–86)
Living alone at home	274 (37.6)	156 (37.7)
Cancer site		
Breast	137 (18.7)	72 (17.4)
Colorectal	131 (17.9)	73 (17.6)
Urinary tract	118 (16.1)	61 (14.7)
Liver or upper gastrointestinal tract	117 (16.0)	71 (17.1)
Prostate	99 (13.5)	32 (7.7)
Hematological	49 (6.7)	25 (6.0)
Other ^a	78 (10.7)	79 (19.1)
Metastasis	299 (41.0)	141 (34.1)
Number of medications per day, median (IQR)	6 (4–8)	6 (4–9)
ECOG-PS		
0 - Fully active	177 (24.3)	58 (14.0)
1 - Restricted activity but ambulatory	213 (29.2)	128 (30.9)
2 - Up >50% of waking hours	129 (17.7)	82 (19.8)
3 - Confined to bed >50% of the day	148 (20.3)	85 (20.5)
4 - Completely disabled	61 (8.4)	60 (14.5)
Abnormal geriatric assessment ^b	632 (86.7)	390 (94.2)
ADL ≤5	218 (29.9)	147 (35.5)
IADL ≤7	457 (62.7)	282 (68.1)
MMSE ≤23	193 (26.5)	124 (30.0)
Mini GDS ≥1	250 (34.3)	125 (30.2)
MNA ≤23.5	426 (58.4)	256 (61.8)
CIRS-G ≥1 comorbidity grade 3/4	414 (56.8)	282 (68.1)
TUG ≥20 seconds	304 (41.7)	113 (27.3)

^aTraining/validation set: unknown primary origin (n = 21/9), lung (n = 17/22), skin (n = 14/27), sarcoma (n = 9/4), brain (n = 5/6), gynecologic (n = 4/5), and other (n = 8/6).

^b≥1 impaired score in ADL, IADL, MMSE, Mini GDS, MNA, CIRS-G, and/or TUG.

Abbreviations: ADL, Activities of Daily Living; CIRS-G, Cumulative Illness Rating Scale for Geriatrics; ECOG-PS, Eastern Cooperative Oncology Group performance status; IADL, Instrumental ADL; IQR, interquartile range; Mini GDS, Mini Geriatric Depression Scale; MMSE, Mini Mental State Examination; MNA, Mini Nutritional Assessment; TUG, Timed Up-and-Go test.

We internally validated our model using bootstrapping procedures with 300 replications to estimate the amount of optimism in our measurement of model discrimination and to compute the bias-corrected AUROC accordingly [16]. The modified G8 was finally applied to the validation set population in which AUROC, sensitivity, specificity, positive predictive values (PPV), and negative predictive value (NPV) were calculated.

There were no missing data for G8 items or GA findings. Few data were missing for the 14 additional items: their proportion ranged from 0% to 4.9% (chronic renal failure) in the training set and 4.1% (health perception status) in the validation set. We imputed missing values using 10-fold multiple imputation by chained equations and combining the estimates using Rubin's rules [17]. Data were assumed to be missing at random, conditional on other predictors and on the outcome. All analyses were performed using Stata v12.1 (StataCorp, College Station, TX, <http://www.stata.com>) at the

two-tailed $p < .05$ level. This observational study is reported according to the STARD checklist for diagnostic accuracy studies.

RESULTS

Patient Characteristics

Between January 2007 and October 2012, 1,056 patients were included into the ELCAPA cohort (training set), of whom 729 had complete G8 data available at the time of our analysis. Between November 2012 and July 2014, 442 patients were included (validation set), of whom 414 had complete G8 data (supplemental online Fig. 2). Table 1 reports the general characteristics of the study populations. Overall, 632 patients (86.7%) had at least one impaired GA test in the training set (14.8% had 1 impaired test, 30.1% had 2 or 3 impaired tests, and 41.8% had 4 or more impaired tests) and 390 (94.2%) had at least one impaired GA test in the validation set (14.5%

Table 2. Univariate analysis of the ability of candidate items to predict impairment of the geriatric assessment (training set; $N = 729$)

Characteristics	Normal geriatric assessment (<i>n</i> = 97)	Impaired geriatric assessment (<i>n</i> = 632)	<i>p</i> value
Original G8 items			
Anorexia			
Absent	88 (90.7%)	296 (46.8%)	<.001
Moderate	8 (8.2%)	232 (36.7%)	
Severe	1 (1.0%)	104 (16.5%)	
Weight loss			
Absent	80 (82.5%)	223 (35.3%)	<.001
1–3 kg	14 (14.4%)	119 (18.8%)	
Does not know	1 (1.0%)	63 (10.0%)	
>3 kg	2 (2.1%)	227 (35.9%)	
BMI (kg/m ²)			
≥23	79 (81.4%)	425 (67.2%)	.06
21 ≤ BMI < 23	7 (7.2%)	97 (15.3%)	
19 ≤ BMI < 21	8 (8.2%)	67 (10.6%)	
<19	3 (3.1%)	43 (6.8%)	
Mobility			
Goes out	97 (100.0%)	409 (64.7%)	<.001
Able to get out of bed/chair but does not go out	0 (0.0%)	131 (20.7%)	
Bed or chair bound	0 (0.0%)	92 (14.6%)	
Dementia/ Depression			
Absent	84 (86.6%)	325 (51.4%)	<.001
Moderate	10 (10.3%)	245 (38.8%)	
Severe	3 (3.1%)	62 (9.8%)	
Prescription drugs (>3)	52 (53.6%)	499 (79.0%)	<.001
Self-rated health status			
Better	58 (59.8%)	174 (27.5%)	<.001
As good	30 (30.9%)	207 (32.8%)	
Does not know	8 (8.2%)	167 (26.4%)	
Not as good	1 (1.0%)	84 (13.3%)	
Age			
<80 years	58 (59.8%)	289 (45.7%)	.02
80–85 years	30 (30.9%)	227 (35.9%)	
>85 years	9 (9.3%)	116 (18.4%)	
New items			
Asthenia	45 (46.4%)	487 (77.1%)	<.001
Risk of fall ^a	26 (26.8%)	435 (68.8%)	<.001
ECOG-PS			
0	71 (73.2%)	106 (16.8%)	<.001
1	24 (24.7%)	189 (29.9%)	
2/3/4	2 (2.1%)	336 (53.2%)	
Urinary and/or fecal incontinence	6 (6.2%)	159 (25.2%)	<.001
Heart failure and/or coronary heart disease	6 (6.2%)	198 (31.3%)	<.001
Complete arrhythmia with atrial fibrillation	6 (6.2%)	126 (19.9%)	.0025
Hypertension	53 (54.6%)	431 (68.2%)	.008
Diabetes	10 (10.3%)	160 (25.3%)	.002
Chronic renal failure	37 (38.1%)	376 (59.5%)	<.001

^aSingle-leg stance, <5 seconds.

Abbreviations: BMI, body mass index; ECOG-PS, Eastern Cooperative Oncology Group performance status.

Table 3. Final multivariate logistic model for predicting impairment of the geriatric assessment (training set; $N = 729$)

Characteristics	Regression coefficient		Odds ratio		<i>p</i> value	Final weights ^a
	Estimate	95% CI	Estimate	95% CI		
No weight loss	0 (ref)	—	1 (ref)	—	—	0
Weight loss 1–3 kg	0.70	(0.00–1.40)	2.01	(1.00–4.06)	.052	2
Weight loss >3 kg/unknown	2.77	(1.56–3.97)	15.90	(4.75–53.23)	<.0001	10
Dementia/depression	0.84	(0.13–1.55)	2.32	(1.14–4.73)	.020	3
Drugs per day ≥ 6	0.64	(0.04–1.23)	1.89	(1.04–3.43)	.036	2
Lower self-rated health status	0.87	(0.06–1.69)	2.40	(1.06–5.43)	.036	3
ECOG-PS grade 0	0 (ref)	—	1 (ref)	—	—	0
ECOG-PS grade 1	1.15	(0.57–1.73)	3.16	(1.77–5.65)	<.0001	4
ECOG-PS grade 2, 3, or 4	3.31	(1.84–4.78)	27.39	(6.28–119.39)	<.0001	12
Heart failure and/or coronary heart disease	1.35	(0.42–2.27)	3.85	(1.53–9.71)	.004	5

^aCalculated by multiplying the regression coefficients by 3.5 and then rounding the results to integers.

Abbreviations: —, no data; CI, confidence interval; ECOG-PS, Eastern Cooperative Oncology Group performance status; ref, reference category.

had 1 impaired test, 35.3% had 2 or 3 impaired tests, and 44.4% had 4 or more impaired tests).

Univariate Analysis

Table 2 reports the main results for the original G8 items and additional candidate items. Of the G8 items, 7 were significantly associated with an abnormal full GA; the remaining item was BMI ($p = .06$). Of the 14 additional items, 12 were significantly associated with an abnormal full GA: asthenia, fall risk, ECOG-PS, incontinence, heart failure/CHD, CAAF, hypertension, diabetes, chronic renal failure (Table 2), fall(s) in the 6 past months ($p < .001$), metastasis ($p = .02$), and chronic respiratory failure ($p = .04$). No significant association was found for gender ($p > .10$), and living alone was significant only in patients with urinary tract cancers ($p = .001$; $p > .2$ for all other cancer sites). ROC analysis of the number of medications per day showed an optimal cutoff of ≥ 6 , as shown by the maximized Youden's index (33.0 [≥ 6] vs. 9.9 [> 3]) and minimal distance between the ROC curve and upper left corner.

Multiple Correspondence Analysis

MCA was conducted to assess the correlations between items (supplemental online Fig. 1). Because several variables were located in close proximity, only those items exhibiting the highest discriminative power and showing a graded distribution along the first MCA axis were kept for the multivariate analysis. These items were the ECOG-PS, weight loss, and fall risk. Thus, we omitted the G8 BMI, anorexia, mobility, and history of fall in the past 6 months.

Multivariate Model

Based on the previous steps, we selected 16 items for the multivariate analysis. A backward stepwise approach showed that six items were independently associated with an abnormal GA. We merged several categories when similar OR values were found across adjacent modalities or for certain modalities that almost perfectly predicted an abnormal GA (ECOG-PS, weight loss, dementia/depression, and self-rated health status), yielding the final model shown in Table 3. Model calibration was excellent ($\chi^2 = 69.8$; $p = .97$). The

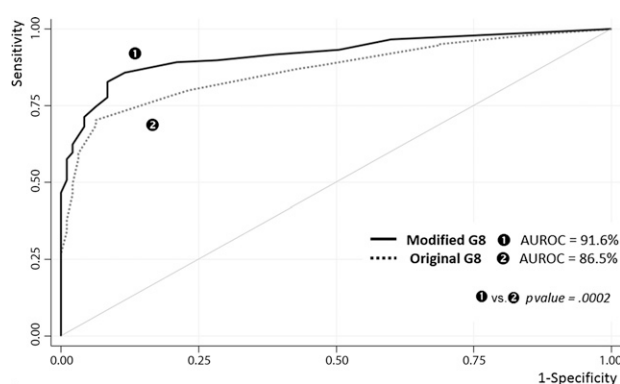


Figure 1. ROC curves for predicting impairment of the geriatric assessment, used as the reference test: original versus modified G8 questionnaire (training set; $n = 729$).

Abbreviations: AUROC, area under the ROC curve; ROC, receiver-operating characteristic.

regression coefficients were then rescaled and rounded to integers to provide weights suitable for use in clinical practice. Among tested multiplication coefficients, 3.5 proved optimal. The final six-item questionnaire is shown in supplemental online Table 1.

Diagnostic Performance of the Original versus Modified G8

Using the recommended score cutoff of ≤ 14 [11], the original G8 demonstrated the following indices: sensitivity, 87.2% (84.3–89.7); specificity, 57.7% (47.3–67.7); PPV, 93.1% (90.7–95.0); and NPV, 40.9% (32.6–49.6). For the modified G8, the cutoff of ≥ 6 of 35 points maximized sensitivity and yielded the following characteristics: sensitivity, 89.2% (86.5–91.5); specificity, 79.0% (69.4–86.6); PPV, 96.5% (94.7–97.9); and NPV, 52.8% (44.3–61.2). Using the higher cutoff of ≥ 7 of 35 produced the following values: sensitivity, 85.8% (82.8–88.5); specificity, 88.4% (80.2–94.1); NPV, 48.8% (41.2–56.6); and PPV, 98.0% (96.4–99.0). The AUROC was 86.5% (83.5–89.6) for the original G8 and 91.6% (89.3–93.9) for the modified G8 ($p = .0002$) (Fig. 1). When we analyzed each cancer site, we found that the modified G8 yielded consistently higher

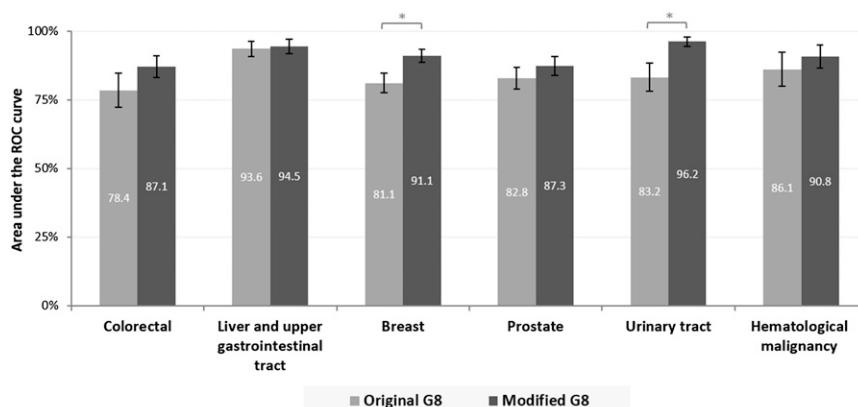


Figure 2. Area under the ROC curve by cancer site: original versus modified G8. *, $p < .05$. Abbreviation: ROC, receiver-operating characteristic.

AUROC values with greater uniformity (from 87.1% in colorectal cancers to 96.2% in urinary tract cancers) compared with the original G8 (from 78.4% in colorectal cancer to 93.6% in liver and upper gastrointestinal tract cancer) (Fig. 2).

Internal Validation and Sensitivity Analyses

The internal validation procedure showed no evidence of over-optimism (optimism = $0.89\% \pm 0.13\%$). The bias-corrected AUROC was 90.7% for the final prediction model. Sensitivity analyses found an AUROC of 91.4% (89.2–93.6) when the CIRS-G was removed from the GA, 90.4% (88.0–92.8) when the ADL/IADL was removed from the GA, and 90.3% (88.0–92.6) when an abnormal GA was defined as impairment of ≥ 2 tests. Applying the modified G8 to the validation set produced the following values: AUROC, 92.8% (88.4–97.2); at the ≥ 6 of 35 cutoff: sensitivity, 91.4% (88.0–94.1); specificity, 75.0% (53.3–90.2); PPV, 98.2% (96.0–99.3); NPV, 37.5% (24.0–52.6); at the ≥ 7 of 35 cutoff: sensitivity, 88.0% (84.1–91.2); specificity, 87.5% (67.6–97.3); PPV, 99.0% (97.2–99.8); and NPV, 33.3% (22.0–46.3). Model calibration was excellent ($\chi^2 = 3.12$; $p = .93$).

DISCUSSION

We developed a modified version of the G8 based on six simple items that are routinely collected by geriatricians. The modified G8 demonstrated a sensitivity of 89.2% and a specificity of 79.0% at the optimized cutoff of ≥ 6 of 35 points, with evidence of homogenous performance across tumor sites.

A wide variety of screening tools have been evaluated to identify patients likely to benefit from a complete GA. The G8, the Vulnerable Elders Survey-13 [18], and the Groningen Frailty Indicator [19] are among the most extensively assessed. Although none of the available tools is markedly better than the others [8], the G8 has the theoretical advantage of having been specifically developed for older patients with cancer, with a selection of items covering important domains in this population [11]. A recent review identified eight studies evaluating the ability of the G8 to predict an abnormal GA [8]. Sensitivity was usually high, with a range of 65%–92% (median, 85.5%), but specificity was lower, ranging from 3% to 75% (median, 59.5%). Similarly, in our study, the original G8 was 87.2% sensitive but only 57.7% specific. It is noticeable that the

G8 was derived from the MNA-SF questionnaire, because of its known high prognostic value in older patients [11]. The fact that the MNA-SF was not designed to specifically detect an abnormal GA probably explains the lack of specificity of the G8 as a screening instrument.

Our objective was to develop a variant of the G8 that would improve the identification of patients requiring a full GA. We used a systematic approach, as typically applied for developing clinical rules or prediction models. This approach consisted of the initial selection of candidate items based on clinical reasoning, multivariate analyses to identify items conveying independent information, internal validation based on bootstrapping techniques to prevent overfitting [14], and reassessment of the model performance on a validation set population. Special attention was given to weights computation to obtain an easy-to-use tool while limiting the loss of information inevitably associated with rounding [15]. Finally, we used multiple imputation at each step of model development to maintain an effective sample size and to control their potential influence on the final model [14].

The modified G8 has only six items yet covers multiple domains included in the GA, namely, nutritional status, mood or cognition, comorbidities, and polypharmacy, in addition to self-rated health status and a simplified version of the ECOG-PS. Interestingly, alterations in these items have been shown to predict adverse outcomes [12, 20–22]. For the number of medications per day, instead of the >3 cutoff used in the G8, we found that the ≥ 6 cutoff improved discrimination, in keeping with conclusions from a recent expert consensus conference [23]. Similarly to the original G8, most items in the modified G8 are subjective and therefore not well suited as criteria for individual diagnosis [11]. The assessment of a past history of heart failure/CHD should not be viewed as an abbreviated version of comorbidity assessment tools such as the CIRS-G but rather as a marker predicting an impaired GA when used in combination with the other items. Despite being associated with an abnormal GA in univariate analysis, several items (age, BMI, anorexia, and mobility) were omitted from the modified G8 because they had minimal independent discriminative power, given their close correlations with other variables, as visualized by MCA.

The modified G8 was robust to sensitivity analyses involving changes in the definition of GA/abnormal GA. This point is of particular interest, because various definitions have been used in previous studies assessing the performance of the G8 in the absence of a clearly defined reference GA [8, 9]. Moreover, the modified G8 showed homogeneity across tumor sites, including various solid tumors and hematological malignancies, whereas evidence for heterogeneity was previously reported for the original G8 [12, 24, 25]. However, despite those apparent benefits, these findings should still be interpreted cautiously in the absence of external validation studies, which are crucial to verify the stability of the model in other populations. Should the improved screening performance of the modified G8 be confirmed, this new tool may encourage the actual use of a two-step approach, in which the results of screening determine whether a full GA is performed [8]. High discriminative power is essential to avoid performing time-consuming unnecessary GAs (false positives) and to ensure that no patients requiring a GA are missed (false negatives). This last point is of major importance, given the consistently high prevalence of abnormal GA findings in several studies conducted in various settings (>80% [11, 12, 24]; 86.7% in our study).

Our study has several limitations. First, in keeping with our study objective, we confined our sample to patients for whom the GA and G8 items were available, which resulted in 327 patients being excluded from the original sample of 1,056 patients. However, we found no statistically significant differences between included and excluded patients regarding the main demographic and clinical features (age, gender, cancer type, and cancer spread; data not shown), suggesting minimal selection bias. Second, several potentially relevant variables were not entered in our database at the time of the analysis, including specific items from validated scales or details on the social environment. These variables deserve investigation in future studies.

CONCLUSION

We developed a modified G8 screening tool that exhibited better diagnostic performance across a variety of tumor sites and greater parsimony, with only six items instead of eight. Our work illustrates the usefulness of combining in-depth statistical analyses with expert judgment to ensure both optimal discriminative power and clinical relevance. Further research is needed to confirm the features of the

modified G8 in other populations and to measure its prognostic value and its impact on treatment decisions and health outcomes.

ACKNOWLEDGMENTS

The ELCAPA Study Group was composed of five geriatricians (P. Cailliet, M. Laurent, E. Liuu, E. Paillaud, H. Vincent), two oncologists (S. Culine, C. Tournigand), one radiation oncologist (J.L. Lagrange), three epidemiologists (F. Canoui-Poittrine, S. Bastuji-Garin, E. Audureau), one pharmacist (M. Carvahlo-Verlinde), one biostatistician (E. Guery), one clinical-research medical doctor (N. Reinald), and two clinical research assistants (C. Pomba, J. Francese).

This work was supported by a French National Cancer Institute grant to the Geriatric Oncologic Unit (Grant RINC4) and a special grant from Ligue contre le Cancer to C.M.T. (doctoral fellowship). The French National Cancer Institute and Ligue contre le Cancer had no role in the design, data collections, analysis, preparation and decision to submit the manuscript for publication. This study's findings were presented orally at the International Society of Geriatric Oncology annual meeting (October 23, 2014, in Lisbon, Portugal). We thank Antoinette Wolfe for editing the manuscript.

AUTHOR CONTRIBUTIONS

Conception/Design: Claudia Martinez-Tapia, Florence Canoui-Poittrine, Sylvie Bastuji-Garin, Pierre Soubeyran, Simone Mathoulin-Pelissier, Elena Paillaud, Etienne Audureau

Provision of study material or patients: Christophe Tournigand, Elena Paillaud, Marie Laurent

Collection and/or assembly of data: Christophe Tournigand, Elena Paillaud, Marie Laurent

Data analysis and interpretation: Claudia Martinez-Tapia, Florence Canoui-Poittrine, Sylvie Bastuji-Garin, Etienne Audureau

Manuscript writing: Claudia Martinez-Tapia, Florence Canoui-Poittrine, Sylvie Bastuji-Garin, Pierre Soubeyran, Simone Mathoulin-Pelissier, Christophe Tournigand, Elena Paillaud, Marie Laurent, Etienne Audureau

Final approval of manuscript: Claudia Martinez-Tapia, Florence Canoui-Poittrine, Sylvie Bastuji-Garin, Pierre Soubeyran, Simone Mathoulin-Pelissier, Christophe Tournigand, Elena Paillaud, Marie Laurent, Etienne Audureau

DISCLOSURES

Pierre Soubeyran: Teva, Celgene (C/A), Pierre Fabre Oncology (other). The other authors indicated no financial relationships.

(C/A) Consulting/advisory relationship; (RF) Research funding; (E) Employment; (ET) Expert testimony; (H) Honoraria received; (OI) Ownership interests; (IP) Intellectual property rights/inventor/patent holder; (SAB) Scientific advisory board

REFERENCES

- Howlander N, Noone A, Krapcho M et al. SEER Cancer Statistics Review, 1975–2011. Bethesda, MD: National Cancer Institute. Available at http://seer.cancer.gov/archive/csr/1975_2011/. Accessed August 1, 2015.
- Wildiers H, Heeren P, Puts M et al. International Society of Geriatric Oncology consensus on geriatric assessment in older patients with cancer. *J Clin Oncol* 2014;32:2595–2603.
- Extermann M, Aapro M, Bernabei R et al. Use of comprehensive geriatric assessment in older cancer patients: Recommendations from the task force on CGA of the International Society of Geriatric Oncology (SIOG). *Crit Rev Oncol Hematol* 2005;55:241–252.
- Rubenstein LZ, Stuck AE, Siu AL et al. Impacts of geriatric evaluation and management programs on defined outcomes: Overview of the evidence. *J Am Geriatr Soc* 1991;39:85–185.
- Puts MT, Santos B, Hardt J et al. An update on a systematic review of the use of geriatric assessment for older adults in oncology. *Ann Oncol* 2014;25:307–315.
- Cailliet P, Laurent M, Bastuji-Garin S et al. Optimal management of elderly cancer patients: Usefulness of the Comprehensive Geriatric Assessment. *Clin Interv Aging* 2014;9:1645–1660.
- Cailliet P, Canoui-Poittrine F, Vouriot J et al. Comprehensive geriatric assessment in the decision-making process in elderly patients with cancer: ELCAPA study. *J Clin Oncol* 2011;29:3636–3642.
- Decoster L, Van Puyvelde K, Mohile S et al. Screening tools for multidimensional health problems warranting a geriatric assessment in older cancer patients: An update on SIOG recommendations. *Ann Oncol* 2015;26:288–300.
- Hamaker ME, Jonker JM, de Rooij SE et al. Frailty screening methods for predicting outcome of a comprehensive geriatric assessment in elderly patients with cancer: A systematic review. *Lancet Oncol* 2012;13:e437–e444.
- Soubeyran P, Bellera C, Goyard J et al. Screening for vulnerability in older cancer patients: The

ONCODAGE Prospective Multicenter Cohort Study. *PLoS One* 2014;9:e115060.

11. Bellera CA, Rainfray M, Mathoulin-Pélissier S et al. Screening older cancer patients: First evaluation of the G-8 geriatric screening tool. *Ann Oncol* 2012;23:2166–2172.

12. Hamaker ME, Mitrovic M, Stauder R. The G8 screening tool detects relevant geriatric impairments and predicts survival in elderly patients with a haematological malignancy. *Ann Hematol* 2014;93:1031–1040.

13. Extermann M, Hurria A. Comprehensive geriatric assessment for older patients with cancer. *J Clin Oncol* 2007;25:1824–1831.

14. Hendriksen JM, Geersing GJ, Moons KG et al. Diagnostic and prognostic prediction models. *J Thromb Haemost* 2013;11(suppl 1):129–141.

15. Cole TJ. Scaling and rounding regression coefficients to integers. *J R Stat Soc Ser C Appl Stat* 1993;42:261–268.

16. Steyerberg E. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*. New York: Springer, 2008.

17. Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. New York, NY: John Wiley & Sons, 1987.

18. Augschoell J, Kemmler G, Hamaker ME et al. PPT and VES-13 in elderly patients with cancer: Evaluation in multidimensional geriatric assessment and prediction of survival. *J Geriatr Oncol* 2014;5:415–421.

19. Baitar A, Van Fraeyenhove F, Vandebroek A et al. Evaluation of the Groningen Frailty Indicator and the G8 questionnaire as screening tools for frailty in older patients with cancer. *J Geriatr Oncol* 2013;4:32–38.

20. Aaldriks AA, Giltay EJ, le Cessie S et al. Prognostic value of geriatric assessment in older patients with advanced breast cancer receiving chemotherapy. *Breast* 2013;22:753–760.

21. Kristjansson SR, Nesbakken A, Jordhøy MS et al. Comprehensive geriatric assessment can predict complications in elderly patients after elective surgery for colorectal cancer: A prospective

observational cohort study. *Crit Rev Oncol Hematol* 2010;76:208–217.

22. Soubeyran P, Fonck M, Blanc-Bisson C et al. Predictors of early death risk in older patients treated with first-line chemotherapy for cancer. *J Clin Oncol* 2012;30:1829–1834.

23. Petit-Monegera A, Rainfray M, Soubeyran P et al. Amélioration des performances d'un test de dépistage gériatrique à huit questions visant à détecter les patients âgés fragiles en cancérologie. *Rev Epidemiol Sante Publique* 2014;62: S118.

24. Velghe A, Petrovic M, De Buyser S et al. Validation of the G8 screening tool in older patients with aggressive haematological malignancies. *Eur J Oncol Nursing* 2014;18:645–648.

25. Liuu E, Canoui-Poittrine F, Tournigand C et al. Accuracy of the G-8 geriatric-oncology screening tool for identifying vulnerable elderly patients with cancer according to tumour site: The ELCAPA-02 study. *J Geriatr Oncol* 2014;5: 11–19.



CME

See <http://www.TheOncologist.com> for supplemental material available online.

This article is available for continuing medical education credit at CME.TheOncologist.com.

For Further Reading:

Caroline Mariano, Grant Williams, Allison Deal et al. Geriatric Assessment of Older Adults With Cancer During Unplanned Hospitalizations: An Opportunity in Disguise. *The Oncologist* 2015;20:767–772.

Implications for Practice:

Geriatric assessment (GA) is an important tool in the management of older cancer patients; however, its primary clinical use has been in the outpatient setting. During an unplanned hospitalization, patients are extremely frail and are most likely to benefit from GA. This study demonstrates that hospitalized older adults with cancer have high levels of functional deficits on GA. These deficits are under-recognized and poorly managed by hospital-based clinicians in a tertiary care setting. Incorporation of GA measures during a hospital stay is a way to improve outcomes in this population.

Supplemental Table for:
Optimizing the G8 Screening Tool for Older Patients with Cancer: Diagnostic Performance and Validation of a 6-item Version
Etienne Audureau et al.

Table S1. Final modified G8: the ELCAPA-07 cohort study

Items	Score
1 Weight loss during the past 3 months	
>3 kg / patient does not know	10
1-3 kg	2
No weight loss	0
2 Neuropsychological problems	
Mild / severe dementia or depression	3
No neuropsychological problems	0
3 Takes at least six drugs per day	
Yes	2
No	0
4 Compared to other people of the same age, how does the patient rate his or her health status?	
Not as good / patient does not know	3
As good or better	0
5 Performance Status (PS)	
PS 2, 3, or 4: Ambulatory but unable to carry out any work activities / Confined to bed >50% / Disabled	12
PS 1: Restricted in physically strenuous activity but ambulatory and able to carry out work of a light or sedentary nature	4
PS 0: Fully active	0
6 Past history of heart failure or coronary artery disease	
Yes	5
No	0
Total	___ / 35

The total score is the sum of the scores on each of the 6 items.
A total score ≥ 6 is considered abnormal and warrants a full geriatric assessment in the two-step approach.

Supplemental Table for:

Optimizing the G8 Screening Tool for Older Patients with Cancer: Diagnostic Performance and Validation of a 6-item Version

Etienne Audureau et al.