Lung Cancer 90 (2015) 281–287

Contents lists available at ScienceDirect

Lung Cancer

journal homepage: www.elsevier.com/locate/lungcan

Validation of survival prognostic models for non-small-cell lung cancer in stage- and age-specific groups☆

Xiaofei Wang[a,*,1], Lin Gu[a,1], Ying Zhang[a,1], Daniel J. Sargent[b,2], William Richards[c], Apar Kishor Ganti[d,3], Jeffery Crawford[e,4], Harvey Jay Cohen[e,4], Thomas Stinchcombe[f,5], Everett Vokes[g,6], Herbert Pang[a,h,1,7]

[a] Department of Biostatistics & Bioinformatics and Alliance Statistics and Data Center, Duke University, Durham, NC, United States
[b] Alliance Statistics and Data Center, Mayo Clinic, Rochester, MN, United States
[c] Brigham and Women's Hospital, Boston, MA, United States
[d] Department of Internal Medicine, VA Nebraska Western Iowa Health Care System and University of Nebraska Medical Center, Lincoln, NE, United States
[e] Department of Medicine, Duke University Medical Center, Durham, NC, United States
[f] Department of Medicine, University of North Carolina, Chapel Hill, NC, United States
[g] Department of Medicine, University of Chicago, Chicago, IL, United States
[h] School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China

ARTICLE INFO

Article history:
Received 24 March 2015
Received in revised form 3 July 2015
Accepted 13 August 2015

Keywords:
Independent validation
Non-small-cell lung cancer (NSCLC)
Performance status
Prognostic models

ABSTRACT

Purpose: Prognostic models have been proposed to predict survival for non-small-cell lung cancer (NSCLC). It is important to evaluate whether these models perform better than performance status (PS) alone in stage- and age-specific subgroups.
Patients and methods: The validation cohort included 2060 stage I and 1611 stage IV NSCLC patients from 23CALGB studies. For stage I, Blanchon (B), Chansky (C) and Gail (G) models were evaluated along with the PS only model. For stage IV, Blanchon (B) and Mandrekar (M) models were compared with the PS only model. The *c*-index was used to assess the concordance between survival and risk scores. The *c*-index difference (*c*-difference) and the integrated discrimination improvement (IDI) were used to determine the improvement of these models over the PS only model.
Results: For stage I, B and PS have better survival separation. The *c*-index for B, PS, C and G are 0.61, 0.58, 0.57 and 0.52, respectively, and B performs significantly better than PS with *c*-difference = 0.034. For stage IV, B, M and PS have *c*-index 0.61, 0.64 and 0.60, respectively; B and M perform significantly better than PS with *c*-difference = 0.015 and 0.033, respectively.
Conclusion: Although some prognostic models have better concordance with survival than the PS only model, the absolute improvement is small. More accurate prognostic models should be developed; the inclusion of tumor genetic variants may improve prognostic models.

1. Introduction

Lung cancer is the leading cause of cancer death among both men and women in the United States [1]. Over 85% of lung cancer cases are non-small cell lung cancer (NSCLC) [2]. Performance status (PS) is a simple functional assessment based on daily physical activities. There are two mutually convertible scoring systems for PS: the Karnofsky score [3] and the ECOG score, also called Zubrod/WHO score [4]. While PS is criticized for its subjectivity [5] and lack of prediction for chemotherapy toxicity [6], PS has become a popular prognostic tool in practice [7] and is one of the most commonly used eligibility criteria and stratification factors in randomized trials [8].

☆ Presented in part in abstract form at the 104th Annual Meeting of the American Association for Cancer Research, Washington, DC, April 6–10, 2013.

* Corresponding author at: Duke University, 2424 Erwin Road, Suite 1102, Durham, NC 27705, United States.
E-mail address: xiaofei.wang@duke.edu (X. Wang).

[1] Supported by U10CA180882, CA033601 and NIA5R21AG042894.
[2] Supported by U10CA180882 and CA033601.
[3] Supported by CA180821 and VA Career Development Award.
[4] Supported by CA47577.
[5] Supported by CA47559 and NIA 5R21AG042894.
[6] Supported by CA41287 and NIA 5R21AG042894.
[7] Supported by NIA 5R21AG042894.

http://dx.doi.org/10.1016/j.lungcan.2015.08.007
0169-5002/

Table 1
Prognostic models investigated.

	B model	C model	G model	M model
Citation	Blanchon et al. [9]	Chansky et al. [10]	Gail et al. [11]	Mandrekar et al. [12]
Geography area	France	IASLC	LCSG	North America
Stage	Stage I–IV	Stage I–IIIA resected	Resected stage I	Advanced stage
Training *n*	2929	9137	392	782 from NCCTG
Validation *n*	1500	SEER 9221	na	433 from SWOG
Endpoint	OS	OS	Recurrence, OS	OS, TTP

OS: overall survival; TTP: time to progression; LCSG: Lung Cancer Study Group (LCSG); NCCTG: North Central Cancer Treatment Group; SWOG: Southwest Oncology Group.

Even with the success of PS, the lung cancer community has been keen on developing new, potentially more objective and accurate prognostic models. Four existing prognostic models for NSCLC patients are summarized in Table 1. It is not surprising that three of the four models include PS as one of the predictors. Blanchon et al. [9] developed a prognostic model (B model) using the data from 2979 NSCLC patients. The prognostic model was based on multivariate Cox regression modeling with baseline prognostic factors: age, histology, PS, sex and stage (I–IV). Chansky et al. [10] (C model) assembled data from 9137 surgically resected NSCLC patients from North America, and used both Cox regression and recursive partitioning and amalgamation analyses to identify risk factors, including age, sex and stage (I–IIIA), and provide classification of risk groups. A prognostic model established by Gail et al. [11] (G model) using 392 early stage NSCLC patients. The prognostic model was based on a Weibull survival model with histology, PS, TNM staging and post-operative infections (empyema, pneumonia or wound infections) as risk predictors. Mandrekar et al. [12] developed a prognostic model (M model) using 782 advanced NSCLC patients. This model incorporated patient characteristics, such as age, PS, sex, stage and body mass index (BMI), and pre-treatment laboratory values, such as hemoglobin (HGB) and white blood cell (WBC) count.

However, these models are not commonly used in clinical practice or clinical trial design. Compared to the PS only model, these models require additional prognostic factors and some factors (e.g., lab values) may not be readily available. Most of these models already include PS as a predictor in the statistical models, but the discriminative accuracies of these models relative to the PS only model have not yet been fully validated using data from independent studies.

Since 1988, the Cancer Leukemia Group B (CALGB) (now part of the Alliance for Clinical Trials in Oncology) has conducted phase II and III clinical trials with NSCLC patients. The goal of the present research was to evaluate the accuracy of the four externally developed prognostic models noted above, for predicting overall survival of stage I and stage IV NSCLC patients. The added value of these prognostic models in discriminating overall survival over the PS only model was a focus. Also, as these prognostic models were predominantly developed on young patients (defined as age < 70 years), we were also interested in evaluating the performance of these prognostic models among elderly patients (defined as age ≥ 70 years).

2. Patients and methods

2.1. Selection criteria

This study utilized data from the CALGB clinical trials to validate four published prognostic models in stage I and stage IV NSCLC. We identified all NSCLC studies targeting either stage I or IV patients conducted between 1988 and 2009 by the CALGB. Stage II and III patients were excluded because of the limited number of patients. Ongoing trials and recently closed trials with pending publications, trials with missing staging, diagnosis and histology information were excluded. The 23 CALGB studies [13–35] included in the validation analysis are listed in the Supplementary material Table S1. Of note none of the trials included selected patients based on epidermal growth factor receptor (*EGFR*) mutation or anaplastic lymphoma kinase (*ALK*) rearrangement status. The validation cohort consists of 2060 stage I (IA–IB) and 1611 stage IV NSCLC patients, among which 2453 patients were <70 years old and 1218 patients were ≥70 years old.

Table 2
Prognostic model risk score.

Prognostic model	Risk score
B model	$1x$(if age > 70)[b] + $1x$(if gender = male) + $3x$(if PS = 1) + $5x$(if PS = 2) + $8x$(if PS = 3) + $10x$(if PS = 4) + $2x$(if histology = large-cell) + $3x$(if stage = IIA–IIB) + $6x$(if stage = IIIA–IIIB) + $8x$(if stage = IV)
C model	log(1.35)x(if adenocarcinoma) + log(1.161)x(if squamous) + log(1.38159)x(if large cell) + log(1.353958)x(if adenosquamous) + log(1.21)x(if gender = male) + log(1.51)x(if age ≥ 70) + log(1.30)x(if stage = IB) + log(1.872)x(if stage = IIA) + log(2.4336)x(if stage = IIB) + log(3.553056)x(if stage = IIIA)
G model[a]	−8.971 + 0.679x(if T1N0 nonsquamous) + 0.103x(if T1N1 squamous) + 1.600x(if T1N1 nonsquamous) + 0.981x(if T2N0 squamous) + 1.282x(if T2N0 nonsquamous) + 0.450x(if PS ≥ 2)
M model	0.26 + 0 x(if PS = 0) + 0.48 x(if PS = 1) + 0.96 x(if PS = 2 or 3) + 0.60 x(if underweight) + 0x(if normal weight) + 0.11x(if overweight) − 0.11x(if obese) + 0x(if normal HGB) + 0.41x(if abnormal HGB) + 0x(if normal WBC) + 0.35x(if high WBC)

[a] Data on wound infection were not collected on 80% stage I NSCLC patients in the validation cohort. To evaluate the G model, we assumed all patients had no wound infection or equivalently the regression coefficient of infection is zero.
[b] (if x) is an indicator function which equals 1 if the statement 'x' is true and 0 otherwise.

2.2. Prognostic models

Calculations of the risk scores based on the four prognostic models are briefly summarized in Table 2. Details on variable definitions and risk score calculations can be found in the corresponding publications [9–12]. Risk groups generating potential separation of survival curves can be defined on mutually exclusive intervals of risk scores. Six risk groups for the B model are formed on the score ranges specified in Table 3 of Blanchon et al. [9]. Five risk groups for the C model are formed based on the survival tree in Fig. 2 of Chansky et al. [10]. Three risk groups for the G model are formed based on the rule in Table 9 of Gail et al. [11]. Risk group was not discussed in Mandrekar et al. [12]. For illustration, we created five