FULL PAPER

BJC

British Journal of Cancer (2017) 117, 744–751 | doi: 10.1038/bjc.2017.232

Keywords: risk score; prognostic model; survival; lung cancer; non-small-cell lung cancer

Lung cancer prognostic index: a risk score to predict overall survival after the diagnosis of non-small-cell lung cancer

Marliese Alexander*[,1,2], Rory Wolfe[1], David Ball[3,4], Matthew Conron[5], Robert G Stirling[6], Benjamin Solomon[4,7], Michael MacManus[3,4], Ann Officer[8], Sameer Karnam[5], Kate Burbury[9] and Sue M Evans[1]

*[1]Department of Epidemiology and Preventive Medicine Monash University, 99 Commercial Road, Melbourne, VIC 3004, Australia; [2]Pharmacy Department Peter MacCallum Cancer Centre, 305 Grattan Street, Melbourne, VIC 3000, Australia; [3]Department of Radiation Oncology Peter MacCallum Cancer Centre, 305 Grattan Street, Melbourne, VIC 3000, Australia; [4]Sir Peter MacCallum Department of Oncology, The University of Melbourne, Parkville, VIC 3010, Australia; [5]Department Respiratory and Sleep Medicine St Vincent's Hospital, 41 Victoria Parade, Fitzroy, VIC 3065, Australia; [6]Department of Allergy Immunology and Respiratory Medicine, Alfred Hospital, 55 Commercial Road, Melbourne, VIC 3004, Australia; [7]Department of Medical Oncology Peter MacCallum Cancer Centre, 305 Grattan Street, Melbourne, VIC 3000, Australia; [8]Departments of Radiation Oncology and Medical Oncology Peter MacCallum Cancer Centre, 305 Grattan Street, Melbourne, VIC 3000, Australia and [9]Department of Haematology Peter MacCallum Cancer Centre, 305 Grattan Street, Melbourne, VIC 3000, Australia

Introduction: Non-small-cell lung cancer outcomes are poor but heterogeneous, even within stage groups. To improve prognostic precision we aimed to develop and validate a simple prognostic model using patient and disease variables.

Methods: Prospective registry and study data were analysed using Cox proportional hazards regression to derive a prognostic model (hospital 1, $n=695$), which was subsequently tested (Harrell's c-statistic for discrimination and Cox–Snell residuals for calibration) in two independent validation cohorts (hospital 2, $n=479$ and hospital 3, $n=284$).

Results: The derived Lung Cancer Prognostic Index (LCPI) included stage, histology, mutation status, performance status, weight loss, smoking history, respiratory comorbidity, sex, and age. Two-year overall survival rates according to LCPI in the derivation and two validation cohorts, respectively, were 84, 77, and 68% (LCPI 1: score $\leqslant 9$); 61, 61, and 42% (LCPI 2: score 10–13); 33, 32, and 14% (LCPI 3: score 14–16); 7, 16, and 5% (LCPI 4: score $\geqslant 15$). Discrimination (c-statistic) was 0.74 for the derivation cohort, 0.72 and 0.71 for the two validation cohorts.

Conclusions: The LCPI contributes additional prognostic information, which may be used to counsel patients, guide trial eligibility or design, or standardise mortality risk for epidemiological analyses.

While there are several recognised factors that predict outcome for patients diagnosed with non-small-cell lung cancer (NSCLC), our ability to risk-stratify individual patients at the time of their diagnosis is limited (Mahar *et al*, 2015; Detterbeck *et al*, 2016).

Lung cancer staging enables clustering of tumours of similar anatomic extent (Goldstraw *et al*, 2007). Staging informs therapy and broadly predicts outcome, however heterogeneity within stage groups suggests individual patient, disease and/or treatment factors influence survival outcomes. In a recent paper by authors on the International Association for the Study of Lung Cancer (IASLC) staging project, fundamental differences between stage classification and prognosis prediction are eloquently discussed (Detterbeck

*Correspondence: M Alexander; E-mail: Marliese.Alexander@petermac.org

Received 26 February 2017; revised 22 May 2017; accepted 13 June 2017; published online 20 July 2017

et al, 2016). The authors highlight the multiple and variable factors that influence prognosis *vs* the deliberately static and singular aim of staging.

Lung cancer prognostication remains rudimentary and to date no single model has demonstrated superior performance, clinical utility, or widespread global uptake. This may in part be due to limited ability to routinely measure proposed variables, or to perceived small incremental survival prediction gain beyond stratification by stage alone. A systematic review of 32 lung cancer prognostic models published from 1996 to 2015 described poor study design and inconsistent analytic approaches as major limitations for meaningful clinical utility (Mahar *et al*, 2015). Critically, while most tools were subject to some form of internal validation, many used inappropriate methods (subject to over-optimism), and/or reported inappropriate statistical measures. External validation was performed on just one-third of studies reviewed.

The objectives of this study were to identify baseline patient and disease variables associated with overall survival (OS) in patients with newly diagnosed NSCLC, and to derive and subsequently validate a simple and generalisable prognostic model.

MATERIALS AND METHODS

Population. Patients with a new histological diagnosis of NSCLC from 2012 to 2014, with prospectively collected data at the time of diagnosis (baseline) and follow-up data were eligible for inclusion. Treatment was clinician directed including positron emission tomography (PET) staging for patients with potentially curable disease and/or participation in an interventional study.

The derivation cohort (D1) included patients from the Peter MacCallum Cancer Centre Thoracic Malignancies Cohort (TMC); an ethics committee approved single institution prospective observational study. Consenting patients were followed from diagnosis, thrice monthly until death or loss to follow-up.

The validation cohorts (V1 and V2) included patients from two other Australian metropolitan tertiary referral centres, identified from participation in the Victorian Lung Cancer Registry (VLCR) (Stirling *et al*, 2014). From 2012 the VLCR, an ethics committee approved clinical quality registry, prospectively followed patients at 6 and 12 months following diagnosis, and until death or loss to follow-up.

Ethics approval for this study was obtained from the Peter MacCallum Cancer Centre (study no. 14/107) and Monash University (study no. CF15/727–2015000333) ethics committees. As all sites were located in Melbourne, cohorts were cross-checked and duplicated patients excluded from V1 or V2 as relevant.

Data. Data were prospectively collected according to predefined criteria specified for the TMC and VLCR, with the exception of mutation status, which in the validation cohorts was collected retrospectively from patient medical records.

Variables considered for the survival prediction model were: TNM stage according to 7th edition UICC staging criteria (Sobin *et al*, 2009); histological subtype (adenocarcinoma, squamous cell carcinoma, large cell carcinoma, NSCLC not otherwise specified (NOS)); mutation status (epidermal growth factor receptor (EGFR), anaplastic lymphoma kinase (ALK) and KRAS); comorbidities according to the Simplified Comorbidity Score including aggregate score and individual parameters of tobacco consumption, diabetes mellitus, renal insufficiency, respiratory comorbidity (history of tuberculosis, history of pleural effusion or pneumonia, asthma, pulmonary embolism, hypoxaemia <60 mm Hg, and/or chronic obstructive pulmonary disease inducing a FEV1 <1.5 l), cardiovascular comorbidity, neoplastic comorbidity, and alcoholism (Colinet *et al*, 2005); Eastern Cooperative Oncology Group (ECOG) performance status (PS) (<2 *vs* $\geqslant 2$) (Oken *et al*, 1982); weight loss within 3 months of diagnosis (0–10, 11–15, $>15\%$); smoking history (current, past, never), smoking magnitude (pack years), sex, and age.

Variables which were dichotomised included: PS to reflect inclusion thresholds for most interventional studies and to limit subjectivity between lower (0–1) and upper (2–3) categories, weight loss (0–10 *vs* $>10\%$) and smoking history (ever *vs* never) based on previous analyses demonstrating no incremental benefit from additional stratification (Alexander *et al*, 2016). Patient status, and where applicable date of death, for D1 was established thrice monthly at attendance for review or phone call to patient or general practitioner, and for V1 and V2, 3–6 monthly phone follow-up. Patients from all cohorts were linked with the Victorian Register for Births, Deaths and Marriages accurate to 31 December 2015, which collects death notification and cause of death.

Statistical analyses. Survival probability using the Kaplan–Meier method was estimated from date of histological diagnosis until death, with living patients censored at the earlier of 31 December 2015 or last study follow-up.

The prognostic influence of variables was assessed by univariate and multivariate Cox proportional hazards regression for patients in D1 with all data fields complete. Preselected variables (TNM stage group) and variables demonstrating an association with OS in univariate analyses ($P<0.2$) were assessed in multivariate models. Variables retaining an association ($P<0.05$) were included in the final model. A secondary model excluding PS (subjective variable) and respiratory comorbidity (not routinely or consistently reported) was also evaluated.

Points were allocated to each variable based on the strength of association (the logarithm of the hazard ratio, HR) in the final model to create a simple scoring system for the full model (Lung Cancer Prognostic Index (LCPI)) and for the model excluding PS and respiratory comorbidity (modified LCPI (m-LCPI)). Interactions were explored and sensitivity analyses undertaken to determine composition and points weighting. For illustrative purposes that would be meaningful for clinicians and patients, points were grouped; the number of groups and the threshold for point cut-offs to define the groups were subjectively determined by the ability to detect significantly different survival outcomes for a maximal number of groups while ensuring reasonably narrow confidence intervals (CI). The LCPI and m-LCPI were applied to D1, V1, and V2 to assess internal (D1) and external validity. Missing data from V1 and V2 were resolved using multiple imputation by chained equations using 50 imputations (van Buuren, 2012). Graphical illustrations of survival curves used the 50th imputation. Predictive performance of a Cox model including only tumour stage was also assessed and compared against the LCPI to ascertain any additive benefit of other included variables. Discrimination ability was assessed using Harrell's c-statistic, and Gonen and Heller's c-statistic with values from multiply imputed data combined according to Rubin's rules (Gönen and Heller, 2005; Marshall *et al*, 2009). Reporting the median for c-statistics was considered (Marshall *et al*, 2009), but this proposed approach lacked theoretical justification and we found mean and median values to be almost identical. Harrell's c is most commonly reported in other studies (facilitating comparison) but lacks a variance measure. Gonen and Heller's c is most appropriate for censored data, is consequently lower in value than Harrell's c, but addresses any bias concerns about the use of Harrell's c (Gönen and Heller, 2005). Model calibration was assessed by plotting Cox–Snell residuals against the cumulative hazard function (Lee and Wang, 2003). All tests were two-sided using a 5% significance level and corresponding 95% CI calculated. All analyses were performed using Stata 14.0 software (StataCorp LP, College Station, TX, USA).

RESULTS

Study population. Analyses included 1458 newly diagnosed NSCLC patients from three independent cohorts, Figure 1 and Table 1. Median follow-up for D1, V1, and V2, respectively, was 27 months (range 0–56), 29 months (range 0–51), and 27 months (range 0–47). During follow-up 880 (60%) patients died and 2-year OS rates for D1, V1, and V2, respectively, were 42% (95% CI 38–46), 54% (95% CI 50–59), and 31% (95% CI 26–37).

Model development – cohort D1. By univariate analyses, age (continuous variable), male sex, comorbidity (renal), stage (IIIA, IIIB, and IV), mutation status (EGFR and ALK but not KRAS), histology (adenocarcinoma), weight loss $>10\%$, and ECOG PS$\geqslant 2$ predicted OS, Supplementary Table 1. Patients not tested for mutation were more likely to be early stage ($P<0.01$) and of non-adenocarcinoma histology ($P<0.01$). Among patients with stage IIIB/IV adenocarcinoma, those not tested for mutation were more likely to be aged $\geqslant$75years ($P=0.01$), and to have worse prognosis than mutation negative patients (HR 1.54, 95% CI 1.07–2.23).

In multivariate analyses (Table 2), OS was similar for patients with EGFR and ALK mutations (hence combined as actionable mutation). KRAS mutation, with no specific drug target, conferred no survival benefit and differences between patients not tested and negative for mutation were not large after adjusting for age (HR 1.41, 95% CI 0.96–2.03). These groups were combined for future analyses (no proven actionable mutation). After adjusting for mutation, there was no survival difference between adenocarcinoma and non-adenocarcinoma histology, but NOS remained predictive of poorer prognosis.

Using the LCPI and m-LCPI scoring systems (Table 3), each one-point increase was associated with >2-fold increased mortality risk (HR 2.21, 95% CI 2.07–2.37 and HR 2.08, 95% CI 1.95–2.22, respectively). The ability of the LCPI to stratify patients according to prognostic group was maintained (and superior to stage alone) in subgroup analyses for early stage patients likely receiving curative surgery or curative chemoradiotherapy (stage I–IIIA) and for advanced stage patients likely receiving non-curative chemotherapy, radiotherapy, or molecular therapies (stage IIIB–IV) – Supplementary Figure 1.

Model performance and validation – cohorts V1 and V2. In the validation cohorts data were missing mostly for weight loss, PS, and respiratory comorbidity, Table 1. In V1, 40% of patents had all data available, 33% were missing one variable, 11% two variables, 9% three variables, and 7% four variables. In V2, 81% of patents had all data available, 11% were missing one variable, 3% two variables, 4% three variables, and 1% four variables.

Survival by cohort and LCPI is presented in Figure 2 and Table 4 (m-LCPI results in Supplementary Figure 2 and Table 2). Overall, the LCPI and m-LCPI provided greater discrimination than stage alone (Harrell's c 0.73, 0.72, and 0.68, respectively), Supplementary Figure 3. Model discrimination ability (Harrell's c) of the LCPI and m-LCPI, respectively, were 0.74 and 0.72 for D1, 0.72 and 0.73 for V1, and 0.71 and 0.68 for V2 (Table 4; Supplementary Table 2). Discrimination on external validation was similar when undertaking analyses for all cases with multiple imputation for missing data (Figure 2) and for complete cases only (Supplementary Figure 4). For all cohorts, Cox–Snell residuals closely aligned with the plotted LCPI cumulative hazard function up to 2 or 3 years, Supplementary Figure 5.

DISCUSSION

The proposed LCPI is a compound scoring system, including established and novel variables, for the prediction of survival following NSCLC diagnosis. The LCPI is simple and generalisable, using data easily and routinely obtained during diagnostic evaluations. Importantly, the model was developed from prospectively collected data and we provide external validation to demonstrate superior survival prediction compared to stage alone (current best practice), and consistent predictive performance across two independent cohorts, c-statistic (Harrell's c) 0.73 and 0.68.

Development of a single model for all stage groups and histological subtypes aimed to provide broad clinical utility and encourage routine application, while maintaining relevance for individual disease stages. Increasing LCPI was correlated with advancing stage; however the LCPI was able to better stratify patients by using other prognostic factors. For example, across our three cohorts up to 22% of patients with metastatic disease scored low-risk LCPI 1–2 and 37% of patients with stage I–II disease scored high-risk LCPI 3–4. Assessment of LCPI calibration on external validation demonstrated good model fit with deviation from the plotted cumulative function observed only at longer survival times, not unexpected in models with censored data.

With fewer prognostic determinants, performance of the m-LCPI was inconsistent; similar to the LCPI for V1, but comparatively lower both in D1 and V2. Notably, in V2 m-LCPI performance was the same as by stage alone (Harrell's c 0.68), suggesting no additional benefit of the modified model omitting PS and respiratory comorbidity in this cohort. Predictive ability of m-LCPI was linked to disease stage; V1 had significantly more early-stage patients (46% stage I–II compared to 23% in D1 and 31% in V2), for whom poor PS and respiratory comorbidity were less common and influence on survival prediction less relevant. Moreover, although maintaining performance in V1, m-LCPI had reduced ability to stratify patients based on the presence/absence of other risk factors. For example, in V1 91% of stage I patients were classified in m-LCPI 1 *vs* 76% in LCPI 1, and only 5% of stage III patients were classified in m-LPCI 4 *vs* 17% in LCPI 4.

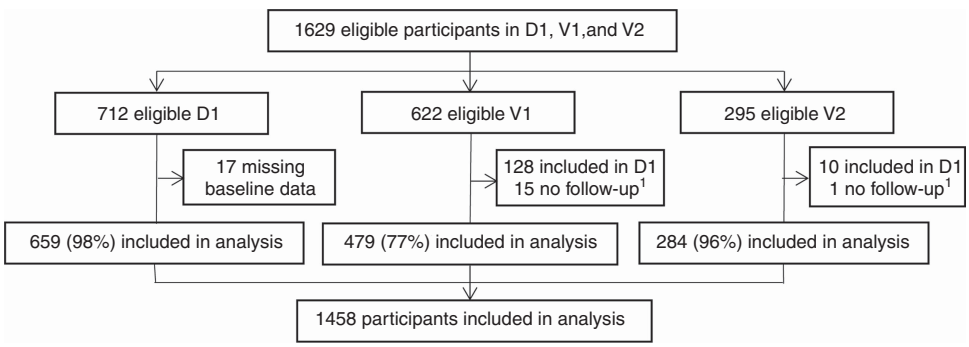

Figure 1. Patient enrolment flow diagram. 1 No hospital follow-up beyond diagnosis and unable to be matched to state death registry (interstate patients). D1, derivation cohort; V1, validation cohort 1; V2, validation cohort 2.

Table 1. Characteristics of patients in the derivation cohort (D1) and validation cohorts (V1 and V2)

	D1, $n=695$		V1, $n=479$			V2, $n=284$		
Characteristic	***n***	**%[a]**	***n***	**%[a]**	***P*-value[b]**	***n***	**%[a]**	***P*-value[b]**
Deceased	448	64.5	233	48.6	0.04	199	70.1	0.26
Male sex	434	62.4	280	58.4	0.13	165	58.1	0.55
Median age (range), years	67 (28–93)		70 (21–89)		<0.01	71 (34–93)		<0.01
Stage group (UICC 7th edition)					0.06			0.04
I	112	16.1	154	32.2		61	21.5	
II	70	10.1	67	14.0		26	9.2	
IIIA	112	16.1	66	13.8		41	14.4	
IIIB	70	10.1	30	6.3		21	7.4	
IV	331	44.6	156	32.6		109	38.4	
Unknown	0	0.0	6	1.3		26	9.2	
Histological diagnosis[c]					0.49			0.76
Adenocarcinoma	442	63.6	318	66.4		156	54.9	
Squamous cell carcinoma	180	25.9	116	24.2		64	22.5	
Large cell	14	2.1	20	4.2		10	3.5	
NOS	59	8.8	25	5.2		54	19.0	
Mutation, all patients					0.05			<0.01
KRAS mutation positive	71	10.4						
Actionable mutation positive	119	17.1	59	12.3		6	2.1	
EGFR positive	84	70.6	54	11.2		6	2.1	
ALK positive	35	29.4	5	1.4		0	0.0	
Actionable mutation, 3B/IV adenocarcinoma					<0.01			<0.01
Positive	84	30.1	26	13.5		3	3.9	
Negative	152	54.5	151	78.7		21	27.3	
Not tested	43	15.4	15	7.8		53	68.8	
Tobacco					0.06			<0.01
Current	128	18.4	103	24.8		86	32.0	
Past	454	65.3	258	62.0		159	59.1	
Never	113	16.3	55	13.2		24	8.9	
Unknown	0	0.0	63	13.2		15	5.3	
Weight loss					0.76			<0.01
0–10%	563	81.0	305	80.3		139	53.7	
>10%	132	19.0	75	19.7		120	46.3	
Unknown	0	0.0	99	20.7		25	8.9	
ECOG performance status					0.46			0.20
ECOG 0	130	18.7	101	41.7		113	39.8	
ECOG 1	429	61.7	99	40.9		80	28.2	
ECOG ⩾2	136	19.6	42	17.4		67	25.8	
Unknown	0	0.0	237	49.5		24	8.5	
Respiratory comorbidity					0.05			<0.01
Yes	241	34.7	156	47.0		48	16.9	
No	454	65.3	225	53.0		236	83.1	
Unknown	0	0.0	98	20.5		0	0.0	

Abbreviations: ALK = anaplastic lymphoma kinase; ECOG = Eastern Cooperative Oncology Group; EGFR = epidermal growth factor receptor; NOS = not otherwise specified.
[a]% incidence for known cases (i.e., excludes unknown category).
[b]P-value for the comparison of known cases in validation cohort vs derivation cohort.
[c]Mixed histological subtypes classified according to dominant subtype.

Application of alternatively proposed prognostic models in our cohorts to directly compare predictive performance with the LCPI (i.e., fitting models to our data), was limited by inconsistencies between available and required data, a major influence on the global uptake of any proposed model. Accordingly, we discuss LCPI performance relative to other models based on comparative performance in external validation. While many models have been published, comparisons were limited consequent to a common lack of external validation among proposed prognostic models, most of which have been derived from retrospective data analysis (Mahar *et al*, 2015).

Lung Cancer Prognostic Index discrimination performance on external validation (Harrell's c 0.72 and 0.71) was similar to a Surveillance, Epidemiology, and End Results Program (SEER) model (Harrell's c 0.69–0.72), but overcomes several limitations (Putila *et al*, 2011). Surveillance, Epidemiology, and End Results Program model derivation on almost two decades of data (1998–2006) achieved an impressive sample size (>230 000) but introduced heterogeneity in terms of available diagnostic modalities, for example, PET staging (Gregory *et al*, 2012), and treatment algorithms, for example, molecular therapies (Sequist *et al*, 2007; Solomon *et al*, 2014). Equally, the SEER model (including tumour stage and grade, age, race, and sex), has not been tested on a more recent cohort, which now routinely incorporates PET and targeted therapies where appropriate. A reflection of this, mutation testing was performed more frequently in our three cohorts compared to a 2010 SEER cohort (72 *vs* 23% for stage IV adenocarcinoma) (Enewold and Thomas, 2016). The role of tumour grade was considered differently in the SEER and our models, deliberately omitted from our cohorts due to reporting subjectivity and limited application in real-world clinical settings (Stang *et al*, 2006). Additionally, the SEER model includes only

Table 2. Predictors of survival in the derivation cohort by multivariate Cox proportional hazards regression

	LCPI				m-LCPI			
	HR	Log HR	95% CI	*P*-value[a]	HR	Log HR	95% CI	*P*-value[a]
Stage group								
I	1.0	0.00			1.0	0.00		
II	1.57	0.45	0.96–2.58	0.09	1.56	0.44	0.96–2.56	0.08
IIIA	3.06	1.12	2.05–4.57	<0.01	2.67	0.98	1.79–3,97	<0.01
IIIB	4.45	1.49	2.91–6.84	<0.01	4.06	1.40	2.66–6.21	<0.01
IV	7.98	2.08	5.55–11.52	<0.01	7.13	1.96	5.01–10.22	<0.01
NSCLC NOS	1.91	0.65	1.48–2.65	<0.01	1.99	0.69	1.47–2.67	<0.01
No proven actionable mutation[b]	1.91	0.65	1.31–2.53	<0.01	1.94	0.66	1.40–2.69	<0.01
ECOG performance status ⩾2	1.78	0.57	1.40–2.23	<0.01	Excluded	Excluded	Excluded	Excluded
Ever Smoker	1.65	0.50	1.19–2.35	<0.01	1.84	0.61	1.30–2.55	<0.01
Respiratory comorbidity	1.46	0.38	1.11–1.68	<0.01	Excluded	Excluded	Excluded	Excluded
Weight loss >10%	1.42	0.35	1.13–1.79	<0.01	1.64	0.50	1.31–2.05	<0.01
Male sex	1.36	0.31	1.11–1.68	0.01	1.29	0.25	1.05–1.59	0.02
Age (per 20 years aged >50)	1.25	0.22	1.05–1.50	<0.01	1.38	0.32	1.16–1.64	<0.01

Abbreviations: ALK = anaplastic lymphoma kinase; CI = confidence interval; ECOG, Eastern Cooperative Oncology Group; EGFR = epidermal growth factor receptor; HR = hazard ratio; LCPI = lung cancer prognostic index; m-LCPI = modified lung cancer prognostic index; NOS = not otherwise specified; NSCLC = non-small-cell lung cancer.
[a]*P*-value for the comparison of overall survival for patients with and without the specified factor, or for stage group as compared with that for patients in the reference group.
[b]Includes EGFR/ALK negative, KRAS positive, mutation not tested.

Table 3. Weighted scores for predictive model for overall survival in newly diagnosed non-small-cell lung cancer

	LCPI points	m-LCPI
Stage group		
I	0	0
II	2	2
IIIA	5	4
IIIB	7	6
IV	9	8
NSCLC NOS	3	3
No proven actionable mutation[a]	3	3
ECOG performance status ⩾2	3	Excluded
Ever smoker	2	2
Respiratory comorbidity	2	2
Weight loss >10%	2	Excluded
Male sex	1	1
Age group 50 or less		
⩽50 years	0	0
51–70	1	1
71–90	2	2
⩾91 years	3	3
LCPI Group	LCPI score	m-LCPI score
LCPI 1	⩽9	⩽8
LCPI 2	10–13	9–11
LCPI 3	14–16	12–14
LCPI 4	⩾17	⩾15

Abbreviations: ALK = anaplastic lymphoma kinase; ECOG = Eastern Cooperative Oncology Group; EGFR = epidermal growth factor receptor; LCPI = Lung Cancer Prognostic Index; m-LCPI = modified lung cancer prognostic index; NOS = not otherwise specified; NSCLC = non-small-cell lung cancer.
[a]Includes EGFR/ALK negative, KRAS positive, mutation not tested.

adenocarcinoma and squamous cell carcinoma, and as such would not be applicable for >10% of our study cohort and the general NSCLC population.

A model published in 2015 using data from 1997 to 2008 also included all stage groups but used individual TNM status rather than stage groups for prediction, justified on the basis that periodic staging classification updates would negate the value of a model including stage groups (Schild *et al*, 2015). This model has not yet undergone prospective external validation. One advantage of using stage group is the ability to include patients where TNM staging is incomplete, for example, documentation of metastasis (common in large epidemiological data sets). In addition to TNM status, that model included quality-of-life score, age, sex, PS, and smoking status. Mutation status, a relevant prognostic indicator for 5–10% of patients harbouring actionable mutations with demonstrated survival benefits (Sequist *et al*, 2007; Solomon *et al*, 2014; Tan *et al*, 2017), was a notable exclusion as for other models.

Global prognostic models (all stages, all treatments) are advantageous for use in clinical settings and for application to large data sets for epidemiologic analyses. Predictive performance in sub-populations remains important, however stage and treatment-specific prognostic models with homogenous populations would be expected to outperform generalised models such as the LCPI. As depicted in Supplementary Figure 1, application of the LCPI to early-stage (I–IIIA) and advanced stage (IIIB–IV) subgroups demonstrated an unsurprising reduction in model performance (LCPI c-statistics: overall: 0.73, stage I–IIIA: 0.69, stage IIIB–IV: 0.67). However the model maintained ability to stratify by survival groups and superior predictive ability compared to stage alone (stage group c-statistics: overall: 0.65, stage I–IIIA: 0.62, stage IIIB–IV: 0.52). The aforementioned SEER model also reported subgroup analyses (by treatment rather than stage) with c-statistic best for early stage and reducing for advanced stages: single modality surgery: 0.77, surgery with radiation: 0.69, single modality radiotherapy: 0.63, treatment unknown: 0.60 (Putila *et al*, 2011). Subgroup analyses were not reported for the discussed 2015 global prognostic model (Schild *et al*, 2015).

Withstanding stated disparities in model design and intent, compared with an externally validated baseline variable model for advanced NSCLC (stage IIIB–IV) (Mandrekar *et al*, 2006), the LCPI performed favourably. That model, including stage, PS, body mass index, haemoglobin level, and white cell count, achieved external validation c-statistic 0.65, whereas the LCPI achieved 0.72 and 0.71 in our two validation cohorts. Although developed using retrospective data and tested in small validation cohorts ($n = 65$, $n = 36$), a model for inoperable NSCLC (stage I–IIIB) treated with radical chemoradiotherapy demonstrated impressive discriminative ability, superior to the LCPI;

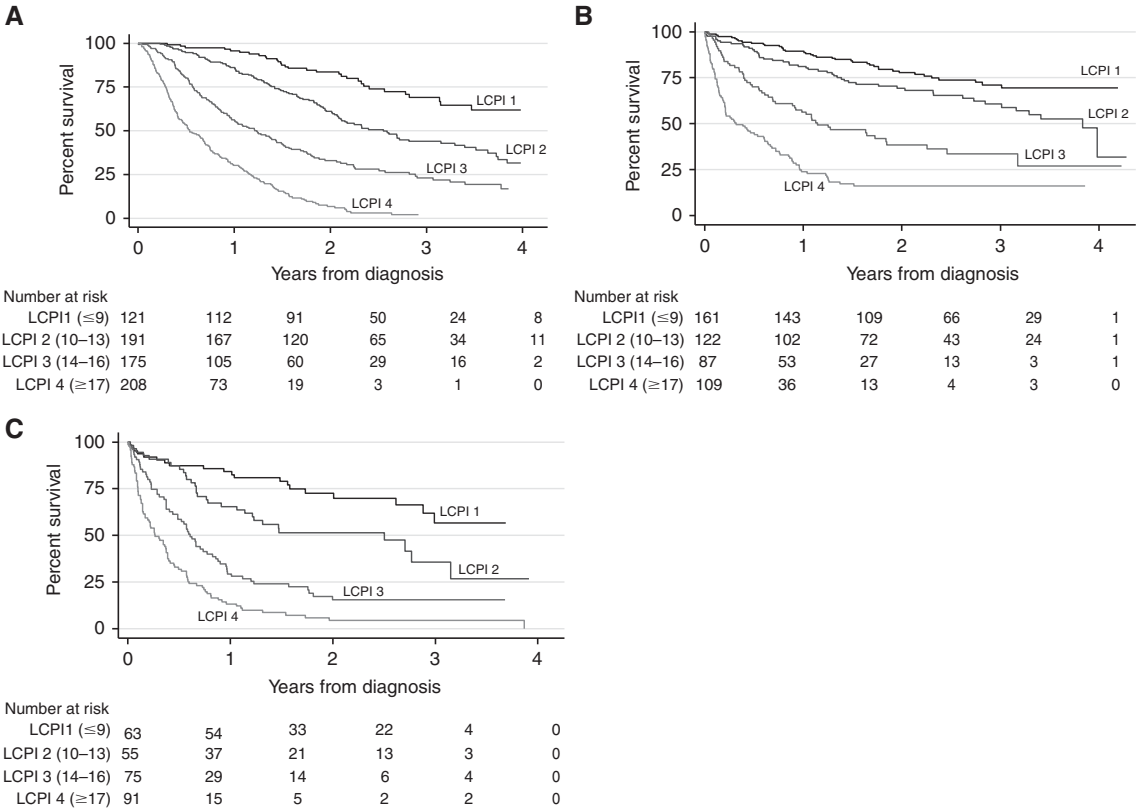

Figure 2. Overall survival curves according to Lung Cancer Prognostic Index (LCPI) in (**A**) derivation cohort, (**B**) validation cohort 1, and (**C**) validation cohort 2.

Table 4. Overall survival according to cohort and LCPI

Cohort description	All patients No. (%)	Stage I No. (%)	Stage II No. (%)	Stage IIIA No. (%)	Stage IIIB No. (%)	Stage IV No. (%)	Mortality HR (95% CI)[a]	1-year OS % (95% CI)	2-year OS % (95% CI)	Median OS, month, (95 % CI)	*P*-value[b]
Derivation Cohort: Harrell's c-statistic 0.74, Gonen and Heller's c-statistic 0.71 (95% CI 0.69–0.73)											
All patients	695 (100.0)	112 (16.1)	70 (10.1)	112 (16.1)	70 (10.1)	331 (47.6)	—	63.5 (59.8–67.0)	42.1 (38.2–45.9)	18.4 (16.7–21.0)	—
LCPI 1 (score ⩽9)	121 (17.4)	68 (60.7)	31 (44.3)	13 (11.6)	2 (2.9)	7 (2.1)	Reference	95.7 (90.0–98.2)	83.7 (75.3–89.4)	NR (41.6–NR)	—
LCPI 2 (score 10–13)	191 (27.5)	41 (36.6)	27 (38.6)	44 (39.3)	13 (18.6)	66 (19.9)	2.4 (1.6–3.6)	86.2 (80.4–90.4)	60.7 (53.1–67.5)	30.6 (25.0–39.5)	<0.01
LCPI 3 (score 14–16)	175 (25.2)	3 (2.7)	9 (12.9)	39 (34.8)	27 (38.6)	97 (29.3)	4.9 (3.3–7.3)	56.0 (48.2–63.0)	32.9 (25.8–40.3)	14.8 (11.4–17.8)	<0.01
LCPI 4 (score ⩾17)	208 (29.9)	0 (0)	3 (4.3)	16 (14.3)	28 (40.0)	161 (48.6)	12.9 (8.8–19.1)	29.6 (23.4–36.0)	6.7 (3.5–11.2)	6.3 (4.9–8.2)	<0.01
Validation Cohort 1: Harrell's c-statistic 0.72, Gonen and Heller's c-statistic 0.69 (95% CI 0.66–0.71)[c]											
All patients	479 (100)	155 (32.4)	67 (14.0)	66 (13.8)	30 (6.3)	161 (33.6)	—	66.4 (61.9–70.4)	54.4 (49.6–58.9)	33.0 (23.4–46.0)	—
LCPI 1 (score ⩽9)	161 (33.6)	119 (76.7)	27 (40.3)	10 (15.2)	4 (13.3)	1 (0.6)	Reference	88.2 (82.8–92.0)	76.9 (70.0–82.5)	NR (NR)	—
LCPI 2 (score 10–13)	122 (25.5)	36 (23.2)	30 (44.8)	28 (42.4)	7 (23.3)	21 (13.4)	1.9 (1.2–3.0)	76.1 (67.1–82.9)	61.3 (51.3–69.8)	47.8 (32.8–NR)	0.01
LCPI 3 (score 14–16)	87 (18.1)	0 (0)	8 (11.9)	22 (33.3)	12 (40.0)	45 (28.0)	3.6 (2.3–5.8)	45.2 (34.4–55.5)	32.0 (21.8–42.7)	13.2 (8.3–22.1)	0.01
LCPI 4 (score ⩾17)	109 (22.8)	0 (0)	2 (3.0)	6 (9.1)	7 (23.3)	94 (58.4)	8.1 (5.4–12.3)	24.4 (16.0–33.9)	15.7 (8.9–24.3)	4.7 (2.5–7.3)	<0.01
Validation Cohort 2: Harrell's c-statistic 0.71, Gonen and Heller's C-statistic 0.69 (95% CI 0.66–0.72)[c]											
All patients	284 (100)	67 (23.6)	27 (9.5)	48 (16.9)	23 (8.1)	119 (41.9)	—	43.3 (37.5–49.0)	31.2 (25.7–36.9)	9.1 (7.3–11.6)	—
LCPI 1 (score ⩽9)	63 (22.2)	59 (88.1)	4 (14.8)	0 (0)	0 (0)	0 (0)	Reference	79.5 (68.7–86.9)	68.3 (55.8–77.9)	NR (33.2–NR)	—
LCPI 2 (score 10–13)	55 (19.4)	5 (7.5)	15 (55.6)	25 (52.1)	5 (4.2)	5 (4.2)	2.1 (1.1–3.9)	53.7 (39.6–65.9)	42.1 (28.8–54.9)	30.0 (13.7–NR)	0.02
LCPI 3 (score 14–16)	75 (26.4)	2 (3.0)	8 (29.6)	16 (33.3)	8 (34.8)	41 (34.5)	4.9 (2.9–8.3)	27.4 (17.8–37.9)	13.8 (6.5–23.1)	6.9 (4.5–8.8)	<0.01
LCPI 4 (score ⩾17)	91 (32.0)	1 (1.5)	0 (0)	7 (14.6)	10 (43.5)	73 (61.3)	8.7 (5.2–14.7)	13.9 (7.4–22.5)	5.1 (1.5–12.2)	3.7 (2.3–4.7)	<0.01

Abbreviations: CI = confidence interval; HR = hazard ratio; LCPI = lung cancer prognostic index; NR = not reached; OS = overall survival.
[a]Hazard ratio for death *vs* LCPI 1.
[b]*P*-value for the overall survival difference compared to lower LCPI group (LCPI 12–15 *vs* LCPI 0–11, LCPI 16–18 *vs* LCPI 12–15, and LCPI 19–29 *vs* LCPI 16–18).
[c]Analyses for validation cohorts includes imputed data for missing variables. One and 2-year OS are presented for the 50th data set.

external validation c-statistics 0.75 and 0.76 (Dehing-Oberije *et al*, 2009). Like the LCPI, that model included sex and PS, but differed by inclusion of forced expiratory volume (*vs* respiratory comorbidity), and nodal status/tumour volume (*vs* stage group). Forced expiratory volume and tumour volume are not routinely assessed in non-curative settings rendering model of limited generalisability.

Lung Cancer Prognostic Index variables are clinically relevant and can be obtained during the diagnostic evaluation of patients presenting with newly diagnosed NSCLC. Advancing age and male sex, as in other studies of early and advanced NSCLC (Wakelee *et al*, 2006; Agarwal *et al*, 2010), conferred increased mortality risk. Although both have been linked to comorbidity burden impacting treatment deliverability and survival (Tammemagi *et al*, 2004), discounting respiratory disease, we found no residual association between comorbidity (including aggregate scores) and survival. With variable prognostic ability across NSCLC cohorts (Alexander *et al*, 2016), possibly related to data reliability (Alexander *et al*, 2017, in press), we strongly advocate prospective assessment of

respiratory comorbidity (and indeed all other varibles) against predefined critieria. Performance status was a strong predictor of OS in this and other studies, including comparative survival estimation with complex multivariate models (Wang *et al*, 2015), yet subjectivity may still be percieved as a limitation. We found exclusion of respiratory comorbidity and PS (m-LCPI) reduced model predictive performance.

Inclusion of mutation status within our model is novel and provided increased stratification for patients with advanced disease. Mutation status is a globally relevant prognostic factor with available targeted treatment options, and reflex EGFR and ALK testing for advanced stage non-squamous carcinomas in patients suitable for therapy has received international endorsement (Leighl *et al*, 2014). Currently including only EFGR and ALK, the LCPI may be expanded in the future to accommodate newly identified targetable mutations. Pursuit of tissue for mutation testing is encouraged in older patients and those presenting with poor PS, particularly where linked to lung cancer burden, with adverse survival outcomes for patients not tested *vs* negative for mutation attributed to these factors in multivariate analyses. Conversely, patient factors did not explain the adverse prognosis of patients with NSCLC NOS in our study, a finding supported by other studies (Righi *et al*, 2014; Tane *et al*, 2014).

The three cohorts in this study were representative of heterogonous real-world patient populations. There were differences across the cohorts, possibly reflecting referral patterns of the tertiary referral specialist cancer centre (D1), compared to major metropolitan general public hospitals (V1 and V2). Notably our population is biased to metropolitan centres and there was overlap in the population catchment areas. However, the ability of the LCPI to retain predictive performance across institutions and cohorts supports generalisability of the model. Reasons for the comparatively poorer 2-year OS rates for V2 are unclear; further validation or recalibration may be required to address this uncertainty. Longer-term analyses are ongoing and will allow assessment of prediction beyond 3–4 years after diagnosis. Missing data were evident in both validation cohorts but overcome by imputation, and reassuringly the LCPI maintained prognostic performance in both imputed and complete case analyses. Further prospective validation, with close to perfect data capture as for our derivation cohort, is warranted. Data linkage to the state-based death registry was implemented to ensure accurate and complete mortality data. However, it is possible a small number of patient deaths were excluded due to interstate migration. While LCPI discrimination ability exceeds current best practice (stage alone), further improvements may be gained by more comprehensive purpose-defined data sets. However, the pursuit of perfect prediction must be balanced against the need for relevant real-time assessment.

Baseline and treatment-based prognostic models have fundamental differences in terms of derivation and application. Treatment undoubtedly impacts survival; however treatment detail, particularly as patients' cycle through various combination strategies, is not always easy to capture. As such, its contribution to static and simple models may not be feasible. While baseline models do incorporate predictors of treatment, treatment strategy models may offer additional prognostication, particularly as alternate treatment pathways emerge and in the setting of recurrent or progressive NSCLC where prior and future treatments may impact outcome. Key laboratory parameters that have shown some interest (Ay *et al*, 2012; Khorana *et al*, 2013; Gu *et al*, 2015) were not included in LCPI due to limited data availability across the three cohorts, and indeed may not be feasible for use in real-world routine care. Similar consideration was given to inclusion of genomic information, which may in time contribute to prognostication, however, since there is currently a lack of globally standardised routine testing, it was not considered appropriate to include in this model intended for routine clinical use.

Importantly, the validated LCPI from our study has current real-world relevance for routine care contributing additional prognostic information, which may be utilised in conjunction with validated tools and evidence-based patient management guidelines. Specifically, the LCPI could be used to counsel patients, guide treatment decisions and clinical trial eligibility, inform comparative effectiveness analyses, adjust for confounders and case-mix when benchmarking quality of care across institutions, and to standardise mortality risk for epidemiological analyses. We present the m-LCPI as a pragmatic alternate model, which may be utilised where data availability is limited, although this model's varied consistency and potential lack of additive value over stage alone in some cohorts must be considered. Both the LCPI and m-LCPI are simple to use and include data that can easily be obtained during the diagnostic evaluation of patients with NSCLC.

ACKNOWLEDGEMENTS

We would like to thank hospital and registry staff for their contributions to study and registry enrolment and data acquisition. In particular we acknowledge the contribution of Professor Catriona McLean, director of anatomical pathology Alfred Health, for assisting in the ascertainment of results for mutation testing. The funding for the study was obtained from Australian Post Graduate Award Monash University.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Agarwal M, Brahmanday G, Chmielewski GW, Welsh RJ, Ravikrishnan KP (2010) Age, tumor size, type of surgery, and gender predict survival in early stage (stage I and II) non-small cell lung cancer after surgical resection. *Lung Cancer* **68**(3): 398–402.

Alexander M, Evans SM, Stirling RG, Wolfe R, Officer A, Macmanus M, Solomon B, Burbury KL, Ball DL (2016) The influence of comorbidity and the simplified comorbidity score on overall survival in non-small cell lung cancer-a Prospective Cohort Study. *J Thorac Oncol* **11**(5): 748–757.

Alexander M, Evans SM, Wolfe R, Ball DL, Burbury KL (2017) Risks associated with use of medical record and administrative data to derive prognostic models. *Med J Aust* in press.

Ay C, Dunkler D, Pirker R, Thaler J, Quehenberger P, Wagner O, Zielinski C, Pabinger I (2012) High D-dimer levels are associated with poor prognosis in cancer patients. *Haematologica* **97**(8): 1158–1164.

Colinet B, Jacot W, Bertrand D, Lacombe S, Bozonnat MC, Daures JP, Pujol JL (2005) A new simplified comorbidity score as a prognostic factor in non-small-cell lung cancer patients: description and comparison with the Charlson's index. *Br J Cancer* **93**(10): 1098–1105.

Dehing-Oberije C, Yu S, De Ruysscher D, Meersschout S, Van Beek K, Lievens Y, Van Meerbeeck J, De Neve W, Rao B, Van Der Weide H, Lambin P (2009) Development and external validation of prognostic model for 2-year survival of non-small-cell lung cancer patients treated with chemoradiotherapy. *Int J Radiat Oncol Biol Phys* **74**(2): 355–362.

Detterbeck F, Chansky K, Groome P, Bolejack V, Crowley J, Shemanski L, Kennedy C, Krasnik M, Peake M, Rami-Porta R (2016) The IASLC Lung Cancer Staging Project: methodology and validation used in the development of proposals for revision of the stage classification of NSCLC in the forthcoming (Eighth) edition of the TNM Classification of Lung Cancer. *J Thorac Oncol* **11**(9): 1433–1446.

Enewold L, Thomas A (2016) Real-world patterns of EGFR testing and treatment with erlotinib for non-small cell lung cancer in the United States. *PLoS One* **11**(6): e0156728.

Goldstraw P, Crowley J, Chansky K, Giroux DJ, Groome PA, Rami-Porta R, Postmus PE, Rusch V, Sobin L (2007) The IASLC Lung Cancer Staging Project: proposals for the revision of the TNM stage groupings in the

forthcoming (seventh) edition of the TNM Classification of malignant tumours. *J Thorac Oncol* **2**(8): 706–714.
Gönen M, Heller G (2005) Concordance probability and discriminatory power in proportional hazards regression. *Biometrika* **92**: 965–970.
Gregory DL, Hicks RJ, Hogg A, Binns DS, Shum PL, Milner A, Link E, Ball DL, MacManus MP (2012) Effect of PET/CT on management of patients with non-small cell lung cancer: results of a prospective study with 5-year survival data. *J Nucl Med* **53**: 1007–1015.
Gu XB, Tian T, Tian XJ, Zhang XJ (2015) Prognostic significance of neutrophil-to-lymphocyte ratio in non-small cell lung cancer: a meta-analysis. *Sci Rep* **5**: 12493.
Khorana AA, Francis C, Culakova E, Lyman GH (2013) Risk score for venous thromboembolism (VTE) in ambulatory patients: impact of cancer type and risk score on VTE incidence, early mortality and progression-free survival. *Blood* **122**: 3597–3597.
Lee ET, Wang JW (2003) *Statistical Methods for Survival Data Analysis*. 3rd edn, (John Wiley & Sons Inc: Hoboken, New Jersey.
Leighl NB, Rekhtman N, Biermann WA, Huang J, Mino-Kenudson M, Ramalingam SS, West H, Whitlock S, Somerfield MR (2014) Molecular testing for selection of patients with lung cancer for epidermal growth factor receptor and anaplastic lymphoma kinase tyrosine kinase inhibitors: American Society of Clinical Oncology Endorsement of the College of American Pathologists/International Association for the Study of Lung Cancer/Association for Molecular Pathology Guideline. *J Clin Oncol* **32**: 3673–3679.
Mahar AL, Compton C, Mcshane LM, Halabi S, Asamura H, Rami-Porta R, Groome PA (2015) Refining prognosis in lung cancer: a Report on the Quality and Relevance of Clinical Prognostic Tools. *J Thorac Oncol* **10**(11): 1576–1589.
Mandrekar SJ, Schild SE, Hillman SL, Allen KL, Marks RS, Mailliard JA, Krook JE, Maksymiuk AW, Chansky K, Kelly K, Adjei AA, Jett JR (2006) A prognostic model for advanced stage non-small cell lung cancer. Pooled analysis of North Central Cancer Treatment Group trials. *Cancer* **107**(4): 781–792.
Marshall A, Altman DG, Holder RL, Royston P (2009) Combining estimates of interest in prognostic modelling studies after multiple imputation: current practice and guidelines. *BMC Med Res Methodol* **9**: 57.
Oken MM, Creech RH, Tormey DC, Horton J, Davis TE, Mcfadden ET, Carbone PP (1982) Toxicity and response criteria of the Eastern Cooperative Oncology Group. *Am J Clin Oncol* **5**(6): 649–655.
Putila J, Remick SC, Guo NL (2011) Combining clinical, pathological, and demographic factors refines prognosis of lung cancer: a population-based study. *PLoS One* **6**: e17493.
Righi L, Vavala T, Rapa I, Vatrano S, Giorcelli J, Rossi G, Capelletto E, Novello S, Scagliotti GV, Papotti M (2014) Impact of non-small-cell lung cancer-not otherwise specified immunophenotyping on treatment outcome. *J Thorac Oncol* **9**(10): 1540–1546.
Schild SE, Tan AD, Wampfler JA, Ross HJ, Yang P, Sloan JA (2015) A new scoring system for predicting survival in patients with non-small cell lung cancer. *Cancer Med* **4**: 1334–1343.
Sequist LV, Joshi VA, Jänne PA, Muzikansky A, Fidias P, Meyerson M, Haber DA, Kucherlapati R, Johnson BE, Lynch TJ (2007) Response to treatment and survival of patients with non-small cell lung cancer undergoing somatic EGFR mutation testing. *Oncologist* **12**(1): 90–98.
Sobin LH, Gospodarowicz M, Wittekind C (2009) International Union against Cancer (UICC). *TNM Classification of Malignant Tumours*. 7th edn, (Wiley-Blackwell: Hoboken, New Jersey.
Solomon BJ, Mok T, Kim DW, Wu YL, Nakagawa K, Mekhail T, Felip E, Cappuzzo F, Paolini J, Usari T, Iyer S, Reisman A, Wilner KD, Tursi J, Blackhall F (2014) First-line Crizotinib versus chemotherapy in ALK-positive lung cancer. *N Engl J Med* **371**: 2167–2177.
Stang A, Pohlabeln H, Müller KM, Jahn I, Giersiepen K, Jöckel KH (2006) Diagnostic agreement in the histopathological evaluation of lung cancer tissue in a population-based case-control study. *Lung Cancer* **52**(1): 29–36.
Stirling RG, Evans SM, Mclaughlin P, Senthuren M, Millar J, Gooi J, Irving L, Mitchell P, Haydon A, Ruben J, Conron M, Leong T, Watkins N, Mcneil JJ (2014) The Victorian Lung Cancer Registry pilot: improving the quality of lung cancer care through the use of a disease quality registry. *Lung* **192**(5): 749–758.
Tammemagi CM, Neslund-Dudas C, Simoff M, Kvale P (2004) In lung cancer patients, age, race-ethnicity, gender and smoking predict adverse comorbidity, which in turn predicts treatment and survival. *J Clin Epidemiol* **57**(6): 597–609.
Tan L, Alexander M, Officer A, MacManus M, Mileshkin L, Jennens R, Herath D, De Boer R, Fox S, Ball D, Solomon B (2017) Survival difference according to mutation status in a prospective cohort study of Australian patients with metastatic non-small cell lung carcinoma (NSCLC). *Intern Med J*; e-pub ahead of print 24 May 2017; doi:10.1111/imj.13491.
Tane S, Nishio W, Ogawa H, Hokka D, Tane K, Tanaka Y, Tauchi S, Uchino K, Sakai Y, Ohbayashi C, Yoshimura M, Maniwa Y (2014) Clinical significance of the 'not otherwise specified' subtype in candidates for resectable non-small cell lung cancer. *Oncol Lett* **8**: 1017–1024.
van Buuren S (2012) Flexible imputation of missing data. *Chapman & Hall/CRC Interdisciplinary Statistics*. CRC Press: Hoboken, New Jersey.
Wakelee HA, Wang W, Schiller JH, Langer CJ, Sandler AB, Belani CP, Johnson DH (2006) Survival differences by sex for patients with advanced non-small cell lung cancer on Eastern Cooperative Oncology Group Trial 1594. *J Thorac Oncol* **1**(5): 441–446.
Wang X, Gu L, Zhang Y, Sargent DJ, Richards W, Ganti AK, Crawford J, Cohen HJ, Stinchcombe T, Vokes E, Pang H (2015) Validation of survival prognostic models for non-small-cell lung cancer in stage- and age-specific groups. *Lung Cancer* **90**(2): 281–287.

Supplementary Information accompanies this paper on British Journal of Cancer website (http://www.nature.com/bjc)