BMJ Open

Predicting 7-year mortality for use with evidence-based guidelines for Prostate-Specific Antigen (PSA) testing: findings from a large prospective study of 123 697 Australian men

Grace Joshy,[1] Emily Banks,[1,2] Anthony Lowe,[3] Rory Wolfe,[4] Leonie Tickle,[5] Bruce Armstrong,[6] Mark Clements[7]

To cite: Joshy G, Banks E, Lowe A, *et al*. Predicting 7-year mortality for use with evidence-based guidelines for Prostate-Specific Antigen (PSA) testing: findings from a large prospective study of 123 697 Australian men. *BMJ Open* 2018;**8**:e022613. doi:10.1136/bmjopen-2018-022613

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2018-022613).

Received 8 March 2018
Revised 16 August 2018
Accepted 5 October 2018

For numbered affiliations see end of article.

Correspondence to
Dr Grace Joshy;
grace.joshy@anu.edu.au

ABSTRACT

Objectives To develop and validate a prediction model for short-term mortality in Australian men aged ≥45years, using age and self-reported health variables, for use when implementing the Australian Clinical Practice Guidelines for Prostate-Specific Antigen (PSA) Testing and Early Management of Test-Detected Prostate Cancer. Implementation of one of the Guideline recommendations requires an estimate of 7-year mortality.

Design Prospective cohort study using questionnaire data linked to mortality data.

Setting Men aged ≥45years randomly sampled from the general population of New South Wales, Australia, participating in the 45 and Up Study.

Participants 123 697 men who completed the baseline postal questionnaire (distributed from 1 January 2006 to 31 December 2008) and gave informed consent for follow-up through linkage of their data to population health databases.

Primary outcome measures The primary outcome was all-cause mortality.

Results 12 160 died during follow-up (median=5.9 years). Following age-adjustment, self-reported health was the strongest predictor of all-cause mortality (C-index: 0.827; 95% CI 0.824 to 0.831). Three prediction models for all-cause mortality were validated, with predictors: Model-1: age group and self-rated health; Model-2: variables common to the 45 and Up Study and the Australian Health Survey and subselected using stepwise regression and Model-3: all variables selected using stepwise regression. Final predictions calibrated well with observed all-cause mortality rates. The 90th percentile for the 7-year mortality risks ranged from 1.92% to 83.94% for ages 45–85 years.

Conclusions We developed prediction scores for short-term mortality using age and self-reported health measures and validated the scores against national mortality rates. Along with age, simple measures such as self-rated health, which can be easily obtained without physical examination, were strong predictors of all-cause mortality in the 45 and Up Study. Seven-year mortality risk estimates from Model-3 suggest that the impact of the mortality risk prediction tool on men's decision making would be small in the recommended age (50–69 years) for PSA testing, but it may discourage testing at older ages.

Strengths and limitations of this study

- The cohort study used for estimating all-cause mortality risk, the 45 and Up Study, is large and population-based and enabled the use of up to 40 self-reported potential predictor variables.
- We externally validated models built in a cohort recruited from residents of the Australia's largest state (NSW) by combining them with Australian national estimates of the distribution of the risk-predictive factors and predicting national mortality rates.
- The ability to link survey data from study participants to administrative records allowed virtually complete ascertainment of all-cause mortality.
- The study was limited by its use of self-reported individual-level characteristics only; it did not take into account potentially important geographical or social clustering variables when predicting mortality.
- We have developed an easily useable and credible approach to estimating mortality risk over 7 years, which requires input of age and self-reported health data only, for use with evidence-based guidelines, that Australian men can take into account when making their decision whether or not to commence Prostate-Specific Antigen testing for early diagnosis of prostate cancer.

INTRODUCTION

Although age is a very strong predictor of mortality in middle aged and elderly people, the number of chronic diseases a person has contributes important additional heterogeneity to life expectancy.[1] Knowledge of a person's life expectancy or fixed-term risk may be required for implementation of evidence-based guidelines for prevention or treatment of chronic disease.[2 3] The 2016 Australian Clinical Practice Guidelines for Prostate-Specific Antigen (PSA) Testing and Early Management of Test-Detected Prostate Cancer[4] recommend that men unlikely to live

a further 7 years should not be tested because they may not live long enough to benefit from earlier diagnosis of prostate cancer in which case the adverse effects of treatment for PSA detected cancer would unnecessarily reduce their quality of life.[5 6] Without a readily available life expectancy estimation tool, clinicians would find it difficult to properly apply this recommendation. Moreover, a guideline that recommends the giving of 'evidence-based decisional support to men considering whether or not to have a PSA test' suggests that men themselves should take life expectancy into account when deciding whether to start, or to continue, PSA testing. While contrary to The Guidelines, it has not been unusual for Australian men to continue PSA testing into their 80s and even 90s.[7] Our goal, therefore, was to develop an accurate and easily useable tool to predict male life expectancy for use when implementing The Guidelines.

While several relevant mortality prediction tools are available for clinical practice, they have been developed mainly for assessment in the elderly,[8 9] make only short-term (6-month) mortality predictions[10 11] or have been developed for specific high-risk populations.[12–17] Although methods to update existing clinical prediction algorithms have been proposed,[18] predictors used in existing models may not be available in datasets from other populations, particularly when, as in this case, use of self-reported health variables that can be easily obtained without physical examination would be desirable. Given these constraints and that no published models included a substantial subset of the questions in questionnaires from the 45 and Up Study (the largest Australian cohort study and potential source of a directly relevant mortality risk model and predictors) and the Australian Health Survey (which can provide normative data on the prevalence of risk predictors), we decided to develop a new model based exclusively on Australian data rather than to respecify an existing model. Thus, we set out to develop and validate a prediction model for short-term mortality in Australian men aged 45 and older using age and self-reported health variables, and to apply the model to predict 7-year risk of death in men. The Guidelines recommend PSA testing from 50 years of age, or 45 years for men at high risk, to 69 years.

METHODS

Data sources and study population

The Sax Institute's 45 and Up Study is an Australian cohort study of 267 153 men and women aged 45 and over randomly sampled from the general population of New South Wales (NSW), Australia's most populous state. Individuals joined the study by completing a baseline postal questionnaire (distributed from 1 January 2006 to 31 December 2008) and giving informed consent for follow-up through repeated data collection and linkage of their data to population health databases. The study methods are described in detail elsewhere.[19] Questionnaire data from study participants have been linked probabilistically to deaths and dates of death by the NSW Centre for Health Record Linkage (http://www.cherel.org.au/). Dates of death were ascertained from the date of recruitment up to 18 June 2014. After excluding 460 (0.17%) participants with invalid age or date of recruitment, data from 123 697 men were available for analysis. The study was restricted to men only since the aim is to aid men who have a limited life expectancy in making decisions about commencing or continuing PSA testing.

Community and participant involvement

This study was commissioned by the Prostate Cancer Foundation of Australia, a broad-based consumer organisation and the peak national body for prostate cancer in Australia. The research was conceived in part in response to feedback received during the public consultation for the Australian Clinical Practice Guidelines for PSA Testing and Early Management of Test-Detected Prostate Cancer. Strong representation was made by consumers that many men in their 70s are in excellent health and would therefore be expected to live long enough to benefit from PSA testing.

Participants in the 45 and Up Study receive an annual newsletter informing them of the study progress and projects underway using the study data. The study website is also kept up to date regarding research underway and events. There is no individual feedback to study participants regarding findings or results specific to that individual.

Prediction model development

The primary outcome was all-cause mortality

Model development followed a similar approach to those of Lee *et al*,[20] Schonberg *et al*[21] and Ganna and Ingelsson,[22] which had used data items similar to those of the 45 and Up Study and, in the case of Ganna and Ingelsson, only used self-reported data. We extended their approach, however, by assessing external validity, recalibrating the prediction models (by averaging, or standardising, over the distributions of the risk factors in the population) and using mortality rate forecasts for the risk predictions.

Baseline questionnaire data from the 45 and Up Study included self-reported information on sociodemographic and behavioural risk factors, medical and surgical history, functional capacity, self-rated health, quality of life and psychological distress. We selected 40 variables to be possible predictors of mortality in the initial stage of model development. They were baseline age (years), body mass index (BMI), smoking status (never/past/current), alcohol consumption (up to 14 or >14 alcoholic drinks per week), needing assistance with daily tasks because of long-term illness or disability (yes/no), perceived overall health (five-point scale), perceived overall quality of life (five-point scale), Medical Outcomes Study Physical Functioning (MOS-PF) score,[23 24] each of the 10 items contributing to the MOS-PF score, Kessler-10 (K10) score for psychological distress, each of the 10 items contributing to the K10 score, each of eight self-reported

chronic conditions ('Has a doctor ever told you that you have…' cancer other than melanoma and skin cancer, heart disease, blood clot, stroke, high blood pressure, diabetes, asthma, Parkinson's disease), sufficient physical activity (150 min of physical activity in five or more sessions a week), treatment for osteoporosis and treatment for arthritis (see online supplementary appendix 1 and online supplementary table 1 for details). Self-rated health and quality of life were based on questions, 'In general, how would you rate your overall health/quality of life?', followed by options of excellent, very good, good, fair and poor. The questions on the physical functioning scale asked participants about whether they are limited in their ability to perform vigorous and moderate physical activities and tasks such as: lifting shopping, climbing stairs, walking, bending, kneeling or stooping and bathing or dressing. Five response options were provided for each of the 10 items contributing to the K10 score: 'none of the time', 'a little of the time', 'some of the time', 'most of the time' or 'all of the time'. Correlations between predictor variables were assessed using Spearman's correlation coefficient (online supplementary table 2). High correlations (correlation coefficients below −0.7 or above 0.7) were observed for closely related variables, for example, between the MOS-PF item scores and the summary score. In such instances, only the relevant item scores were included in stepwise regression, not the summary score.

We studied the association of each variable with male all-cause mortality using Cox proportional hazard models with age as the underlying timescale (online supplementary table 3). A p<0.0001 from a test based on Schoenfeld residuals was used to identify proportionality violations,[22] which were taken as indicators of interaction with age. To assess the age-adjusted association of predictors with mortality, we used univariate analysis with time-in-study as the timescale. Age was added as a covariate and an interaction with log(age-44) (which makes 45 years the baseline age) was included if indicated. Discrimination was assessed by using Harrell's concordance (C)-index,[25] which is a generalisation to survival analysis of the area under the receiver operating characteristic curve. All C-indices reported include the effects of age and the examined covariates.

We imputed missing data using multiple imputation by chained equations, with five imputed datasets.[26] The imputation model included the Nelson-Aalen estimate of cumulative hazard, the event indicator, selected predictors and auxiliary variables (private health insurance, index of relative socioeconomic disadvantage population quintile, region of residence, highest educational qualification and annual household income). See online supplementary tables 4 and 5 for details. The proportion of missing values in any one auxiliary variable was less than 4%. For multi-item variables K10 and MOS-PF, individual items were imputed and the aggregate scores recalculated after appropriate transformation using Rubin's rule.[27] We assessed whether the imputations were acceptable by comparison of plots of the distribution of recorded and imputed values for all measurements.

We used Cox proportional hazards regression to develop three prediction models for all-cause mortality in males: Model 1 using only age group, self-rated overall health and their interaction; Model 2 by the backward stepwise procedure (with a Wald χ^2 p=0.25 level for entry into the model and p=0.15 level to remain in the model) from all variables common to the 45 and Up Study and the Australian Health Survey (11 variables including age and six interaction terms with log(age-44)) and Model 3 (the full model) by the backward stepwise procedure from 21 variables and 18 interaction terms with age that were the most strongly associated with mortality (based on univariate analysis adjusted for age) and not highly correlated with one another (Spearman's rank correlation coefficient between −0.7 and 0.7). For variables with an intrinsic degree of severity (eg, limitation in walking 100 m, 0.5 km and 1 km), only the item with the highest C-index value was included. The 21 variables were age group, self-rated health, smoking, alcohol consumption, needing help with daily tasks, sufficient physical activity, BMI, limitation in kneeling or stooping, limitation in climbing several flights of stairs, limitation in walking 1 km, limitation in lifting or carrying shopping, limitation in moderate activities, limitation in bathing or dressing, felt everything was an effort, felt tired out for no good reason, heart disease, stroke, blood clot, Parkinson's disease, cancer and treatment for osteoporosis. Seven of these variables had no missing values (age group, heart disease, stroke, blood clot, Parkinson's disease, cancer, treatment for osteoporosis); the remaining variables had fewer than 9.8% missing values (median 6.5% (IQR 3.1%–8.3%)). The prediction models were validated internally using 10-fold cross-validation.

External validation and calibration

A fully independent external validation to assess generalisability could not be pursued as a suitable external database was not available. As an alternative form of external validation, we compared hazard predictions from the 45 and Up Study weighted by covariate patterns from the Australian Health Survey (supplemented by data from the 45 and Up Study) with the observed male mortality rates for Australia in 2012. The intercept term for the hazard predictions for the 45 and Up Study was calculated by $\hat{a} = \log\left(-\log\left(\hat{S}_0\left(5\right)\right)/5\right)$, where $\hat{S}_0\left(5\right)$ is the Breslow estimator of survival at 5 years from the fitted Cox proportional hazards regression models. The average hazard for a given age group a was calculated from

$$\hat{h}\left(a\right) = E_X\left(\hat{h}\left(a|X\right)\right) = \sum_i w\left(a\right)_i \exp\left(\hat{\alpha} + \hat{\theta}_a + \sum_j \hat{\beta}_j x_{ij}\right)$$

where $\hat{\beta}$ are the estimated coefficients from the Cox proportional hazards regression model, i is an index for strata of covariates, j is an index for covariates (which may depend on age a), $\hat{\theta}_a$ is the log HR between age group

a and age group 45-49 years and $w(a)_i$ are age-specific weights for the proportion of the age-specific population with a given covariate pattern. The weights were calculated using two approaches. For Model 1, we used the proportion within an age group that had a particular category of self-rated overall health. For Model 2 and Model 3, for each age group, we simulated 2000 values from a multivariate Normal distribution based on the covariate correlation matrix and then categorised the covariates using the age-specific proportions for each covariate from the Australian Health Survey (supplemented by data from the 45 and Up Study). The categories were calculated from the simulated values using the Normal quantile function. We compared $\hat{h}(a)$ with the observed mortality rate $h_{obs}(a)$ in 2012,[28] abridged version.[29]

For *recalibration*, we calculated age-specific intercept terms

$$\hat{\gamma}_a = \log\left(\frac{1}{\sum_i w(a)_i \exp\left(\sum_j \hat{\beta}_j x_{ij}\right)}\right)$$

for each age group a. Then the recalibrated HRs for age a and covariates x are

$$\hat{HR}_c(a,x) = \exp\left(\hat{\gamma}_a + \sum_j \hat{\beta}_j x_j\right).$$

Note that $E_X\left(\hat{HR}_c(a,X)\right) = 1$.

Applying recalibrated estimates to life tables

Mortality rates were projected using the Booth-Maindonald-Smith variant[30] of the Lee-Carter method[31] fitted to male period mortality data for single years of age from 45 to 94 and age category 95+ years, and calendar years 1970 to 2011.[32] The upper age band of 95 years and over was chosen to minimise the impact of variable mortality rates at the oldest ages, and the 1970 to 2011 period was selected based on statistical goodness-of-fit criteria[30] and ensuring a minimum length of fitting period.

Cohort life expectancies were calculated from forecast mortality rates for each cohort defined by year of birth with the life tables constructed using standard methods.[33]

We used the recalibrated HRs with life tables to calculate survival to time t as follows:

$$\hat{S}(t|a,x) = S(t|a)^{\hat{HR}_c(a,x)},$$

where $S(t|a)$ is survival from a life table from age a to age $a+t$. For Model 3, we calculated age-specific deciles of the predicted 7-year survival based on the covariate distributions from the Australian Health Survey. We also presented the variation in age-specific 7-year survival using a violin plot, where the density of the risk predictions forms the left and right sides, and a boxplot forms the centre for each distribution.

Software

All imputation and prediction modelling was carried out using SAS software V.9.4.[34] External validation and calibration were undertaken using R software. All forecasts were implemented using the 'demography' package[35] on the Comprehensive R Archive Network. Finally, the recalibrated models were implemented in Javascript and HTML.

RESULTS

Of the 123 697 men aged 45 years and over, followed up for a median 5.9 years (IQR 5.7–6.4), 12 160 died during follow-up. Median follow-up times were 3.6 years for those who died and 5.9 years for those who did not die. In age-adjusted analyses, self-reported health was the strongest predictor of all-cause mortality among men (C-index 0.827 (95% CI 0.824 to 0.831)); other strong predictors included physical functioning limitations (items on MOS-PF scale: limitation in walking 1 km, moderate activities, climbing several flight of stairs, lifting, walking 0.5 km, climbing one flight of stairs, walking 100 m, vigorous activities, dressing and bending), quality of life, needing help with daily tasks, cancer, smoking, feeling tired, feeling everything was an effort, diabetes, heart disease, BMI, sufficient physical activity, stroke, treatment for osteoporosis, blood clot and Parkinson's disease (online supplementary table 6).

The coefficient estimates from the three models were used to create risk prediction scores. After stepwise regression, Model 2 included 7 variables and five interaction terms with age and Model 3 included 19 variables and eight interaction terms with age. Variables included in the scores are listed in table 1. Internal validation of the prediction scores indicated that each had good discrimination ability: the C-indices were, respectively, 0.828, 0.838 and 0.848 for Model 1, Model 2 and Model 3, compared with 0.803 for a model with only age. Further, when we compared the observed age-specific all-cause mortality rates for Australian males in 2012 with the model-predicted rates weighted by the relevant population exposure distributions (external validation), the predicted rates followed the pattern of the observed rates (figure 1) and the age-specific ratios of predicted to observed rates varied within relatively narrow ranges: Model 1 varied from 0.89 to 1.15, Model 2 varied from 0.91 to 1.18 and Model 3 varied from 0.82 to 1.03 (figure 2).

The three models led to different ranges in the prediction of short-term survival. For example, the range in the percentage of men remaining alive after 7 years is broader under Model 3 than Model 1. This difference is not unexpected given the simplicity of Model 1 compared with Model 3. The variation in age-specific 7-year survival under Model 3 is shown in figure 3.

Characteristics of the study population used in prediction modelling and the model estimates are presented in online supplementary tables 7 and table 1, respectively. The final estimates were recalibrated to predict the observed population rates. Prototypes for mortality prediction tools based on calibrated estimates from the three models are provided as source codes for HTML

Table 1 Coefficient estimates from three prediction models of mortality in men

		Model 3		Model 2		Model 1	
		Estimate	**SE**	**Estimate**	**SE**	**Estimate**	**SE**
Age, years	45–49	Ref		Ref		Ref	
	50–54	0.2044	0.1133	0.1695	0.1127	0.2752	0.1100
	55–59	0.3808	0.1263	0.2918	0.1253	0.4524	0.1183
	60–64	0.7408	0.1382	0.6545	0.1371	0.8600	0.1273
	65–69	1.1468	0.1494	1.0639	0.1481	1.2866	0.1363
	70–74	1.5714	0.1596	1.4989	0.1582	1.7396	0.1449
	75–79	2.0699	0.1685	2.0004	0.1671	2.2640	0.1526
	80–84	2.4602	0.1758	2.4021	0.1744	2.6947	0.1591
	85+	3.0051	0.1848	2.9968	0.1832	3.3333	0.1674
Self- rated health	Excellent	Ref		Ref		Ref	
	Very good	−1.4000	0.2634	−1.4502	0.2648	−1.5301	0.2603
	Good	−0.8637	0.2520	−0.8978	0.2498	−0.7036	0.2344
	Fair	0.2468	0.2907	0.4762	0.2768	1.1554	0.2364
	Poor	0.3589	0.3779	1.0413	0.3568	2.3425	0.2805
Age*Self-rated health	Excellent	Ref		Ref		Ref	
	Very good	0.4760	0.0802	0.5030	0.0805	0.5411	0.0793
	Good	0.3921	0.0769	0.4463	0.0761	0.4337	0.0718
	Fair	0.1259	0.0875	0.1627	0.0834	0.0625	0.0724
	Poor	0.1973	0.1126	0.1600	0.1068	−0.0463	0.0852
Smoking	Current smoker	1.5826	0.2056	1.5070	0.2086		
	Past smoker	0.5663	0.1634	0.5869	0.1649		
	Never smoker	Ref					
Age*Smoking	Current smoker	−0.2957	0.0644	−0.2662	0.0653		
	Past smoker	−0.0924	0.0474	−0.0924	0.0479		
	Never smoker	Ref		Ref			

Continued

Table 1 Continued

		Model 3		Model 2		Model 1	
		Estimate	**SE**	**Estimate**	**SE**	**Estimate**	**SE**
BMI	15–<18.5	0.2633	0.5291	0.4742	0.5060		
	18.5–<20	0.6451	0.4110	0.6830	0.4001		
	20–<22.5	−0.3112	0.2554	−0.2073	0.2545		
	22.5–<25	Ref		Ref			
	25–<27.5	−0.9042	0.2228	−0.9782	0.2267		
	27.5–<30	−0.5645	0.2312	−0.6830	0.2312		
	30–<35	−0.9336	0.2458	−1.1367	0.2517		
	8 35–50	−1.0104	0.3296	−1.3081	0.3325		
Age*BMI	15–<18.5	0.0876	0.1528	0.0438	0.1456		
	18.5–<20	−0.0808	0.1182	−0.0840	0.1152		
	20–<22.5	0.1575	0.0734	0.1307	0.0732		
	22.5–<25	Ref		Ref			
	25–<27.5	0.2203	0.0647	0.2437	0.0659		
	27.5–<30	0.1109	0.0681	0.1543	0.0682		
	30–<35	0.2142	0.0736	0.2910	0.0756		
	35–50	0.2906	0.1017	0.4011	0.1027		
Diabetes	No	Ref		Ref			
	Yes	1.3570	0.1964	1.4087	0.1976		
Age*Diabetes	No	Ref		Ref			
	Yes	−0.3308	0.0576	−0.3425	0.0579		
Heart disease	No	Ref		Ref			
	Yes	1.0315	0.1888	1.2094	0.1895		
Age*Heart disease	No	Ref		Ref			
	Yes	−0.2662	0.0545	−0.3088	0.0547		
Stroke	No	Ref		Ref			
	Yes	0.0528	0.0308	0.1362	0.0304		
K10 item (felt everything was an effort)	None of the time	Ref		Ref			
	A little of the time	−0.2895	0.1925	−0.3779	0.1957		
	Some of the time	0.0431	0.2484	0.0066	0.2457		
	Most of the time	−0.4929	0.3474	−0.4779	0.3531		
	All of the time	0.1338	0.3821	0.1414	0.3790		

Continued

Table 1 Continued

		Model 3		Model 2		Model 1	
		Estimate	**SE**	**Estimate**	**SE**	**Estimate**	**SE**
Age*K10 item (felt everything was an effort)	None of the time	Ref		Ref			
	A little of the time	0.1031	0.0568	0.1419	0.0577		
	Some of the time	0.0119	0.0733	0.0537	0.0729		
	Most of the time	0.1866	0.1039	0.2323	0.1070		
	All of the time	0.0028	0.1125	0.0473	0.1123		
Sufficient physical activity	0 No	0.1852	0.0209	0.3153	0.0195		
	1 Yes	Ref		Ref			
Treatment for osteoporosis	0 No	Ref		Ref			
	1 Yes	0.0903	0.0401	0.1525	0.0400		
K10 item (felt tired out for no good reason)	None of the time	Ref		Ref			
	A little of the time	−0.0836	0.0269	−0.0622	0.0266		
	Some of the time	−0.0695	0.0309	−0.0126	0.0302		
	Most of the time	−0.0577	0.0470	0.0293	0.0467		
	All of the time	0.0563	0.0683	0.1536	0.0683		
Cancer	0. No	Ref					
	1. Yes	3.2073	0.1715				
Age*Cancer	0. No	Ref					
	1. Yes	−0.7939	0.0498				
Parkinsons	0. No	Ref					
	1. Yes	0.2701	0.0576				
Needing help for disability	0. No	Ref					
	1. Yes	0.2780	0.0323				
MOS-PF item (Limitation in walking 1 km)	Yes, limited a lot	0.3787	0.0382				
	Yes, limited a little	0.2194	0.0317				
	No, not limited at all	Ref					
MOS-PF item (Limitation in kneeling or stooping)	Yes, limited a lot	−0.5421	0.2291				
	Yes, limited a little	−0.0295	0.1868				
	No, not limited at all	Ref					
Age*MOS-PF item (Limitation in kneeling or stooping)	Yes, limited a lot	0.0847	0.0662				
	Yes, limited a little	−0.0306	0.0544				
	No, not limited at all	Ref					

Continued

Table 1 Continued

		Model 3		Model 2		Model 1	
		Estimate	**SE**	**Estimate**	**SE**	**Estimate**	**SE**
MOS-PF item (Limitation in bathing or dressing)	Yes, limited a lot	0.1514	0.0471				
	Yes, limited a little	−0.0352	0.0330				
	No, not limited at all	Ref					
MOS-PF item (Limitation in moderate activities)	Yes, limited a lot	0.2607	0.0413				
	Yes, limited a little	0.1566	0.0301				
	No, not limited at all	Ref					
MOS-PF item (Limitation in climbing several flight of stairs)	Yes, limited a lot	0.2575	0.0402				
	Yes, limited a little	0.1581	0.0314				
	No, not limited at all	Ref					

Predictors corresponding to interaction with age were constructed using log(age-44).
Model 1: Age group, self-rated overall health and their interaction.
Model 2: Age group, self-rated overall health*, smoking*, BMI*, diabetes*, history of heart disease*, history of stroke; (removed in stepwise regression: felt everything was an effort*, felt tired out for no good reason, recent treatment for osteoporosis, sufficient physical activity).
Model 3: Age group, self-rated overall health*, smoking status*, BMI*, history of diabetes*, history of heart disease*, history of stroke, history of cancer*, history of Parkinson's disease, needing help for daily tasks, sufficient physical activity, felt everything was an effort*, felt tired out for no good reason, limitation in kneeling or stooping*, limitation in bathing or dressing, limitation in moderate activities, limitation in climbing several flights of stairs, recent treatment for osteoporosis; (removed in stepwise regression: limitation in walking 1 km, limitation in lifting or carrying shopping, history of blood clot).
The AIC values for final prediction models were 260066.758, 258639.148 and 257099.688 respectively for M1, M2 and M3. Similarly, BIC values were 260185.248, 259031.658 and 257610.694 respectively for M1, M2 and M3.
*Includes interaction term with age.
AIC, Akaike information criterion; BMI, body mass index; MOS-PF, Medical Outcomes Study Physical Functioning.

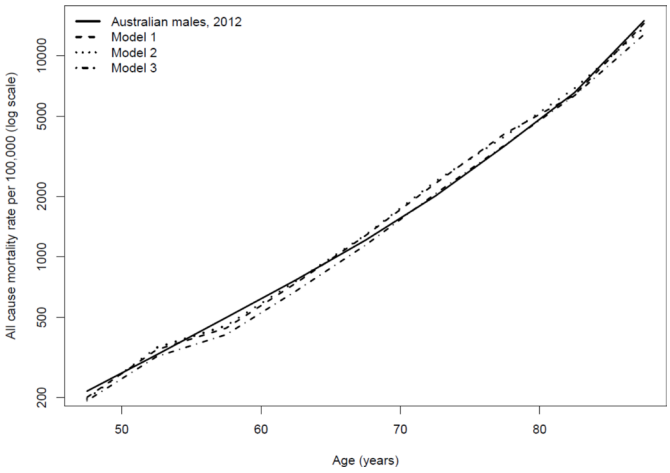

Figure 1 Comparison of predicted mortality rates with observed all-cause mortality rates in Australian men. *Model 1*: model with age, self-rated health and interaction term; *Model 2*: reduced model with variables common to the 45 and Up Study and the Australian Health Survey; *Model 3*: full model with stepwise selected variables and relevant interaction terms.

with embedded JavaScript (three online supplementary files M1_men_html, M2_men_html, M3_men_html).

An indication of the utility of these predictions in PSA testing decision making can be obtained from table 2, which shows, from Model 3, the deciles of the estimated risks of death within 7 years from calendar year 2018 for a simulated population with covariate distributions from the Australian Health Survey.

Both the risk of death and the spread of the distribution increase with age, especially above age 65 years. These numbers suggest that few men in The Guidelines' recommended age-range for PSA testing in Australia (50–69 years) would have greater than a 50% risk of death within 7 years. The highest decile of estimated 7-year risk of death at 70 years of age was only 19.0%; and at the recommended age for beginning testing, 50 years, the highest decile of risk of death within 7 years was 2.4%. It is important to note that the Guidelines' upper age limit for testing was not based on the all-cause mortality risk after this age, but on evidence that the harms of testing, particularly from overdiagnosis, exceed the benefits of testing after 69 years of age.

DISCUSSION

We used a large-scale prospective cohort study to investigate the associations of 40 potential predictor variables with all-cause mortality in middle aged and older Australian men and validated three prediction models for all-cause mortality using data from the Australian Health Survey to provide the population prevalence of predictor variables.

In age-adjusted analyses, self-reported health was found to be the strongest predictor of all-cause male mortality. Other strong predictors included physical functioning

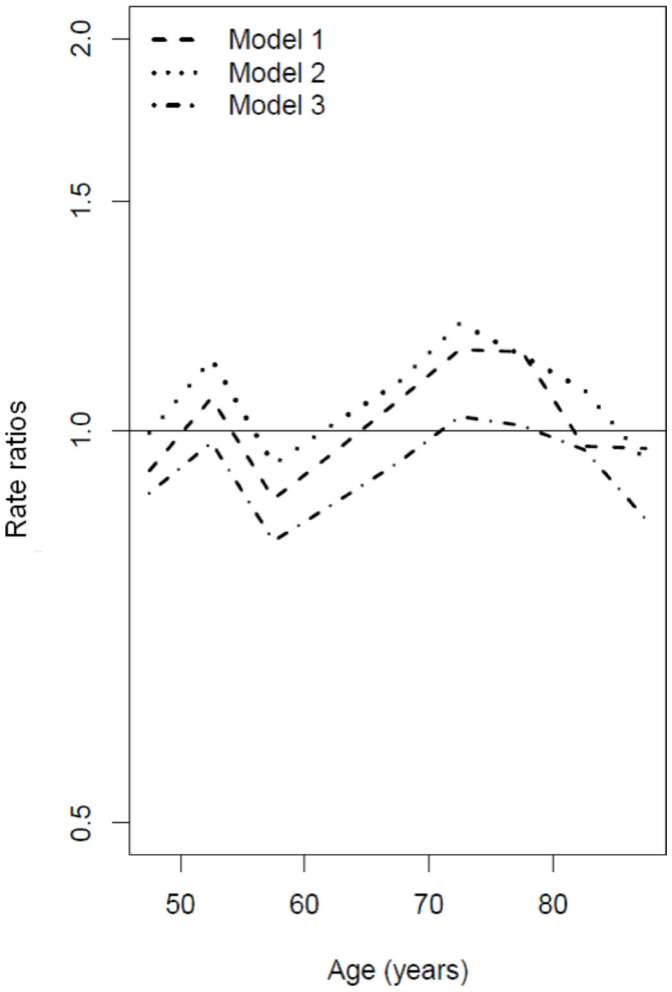

Figure 2 Predicted mortality rates relative to observed mortality rates in Australian men in 2012. *Model 1:* model with age, self-rated health and interaction term; *Model 2:* reduced model with variables common to the 45 and Up Study and the Australian Health Survey; *Model 3:* full model with stepwise selected variables and relevant interaction terms.

limitations, quality of life, lifestyle/behavioural factors (smoking, BMI, sufficient physical activity), comorbidities and measures of psychological distress. The models showed good discrimination as expressed using the C-index and appeared accurate when compared with all-cause mortality rates for Australian males in 2012. The prediction scores developed have been used in combination with Australian life tables to predict mortality risks.

Previous mortality prediction studies[20–22 36 37] among middle aged and older adults vary in terms of study population, predictors assessed and model development methods. In addition to age and sex, self-rated health and physical function difficulty have been identified as strong predictors of all-cause mortality in all studies with self-reported health information. In the absence of self-reported health information, Tan *et al*[37] showed that sex-specific models that used age and Elixhauser comorbidities could accurately predict life expectancy and risk of death at 5–10 years in a cohort of over 1.13 million US Medicare beneficiaries 66–90 years of age.

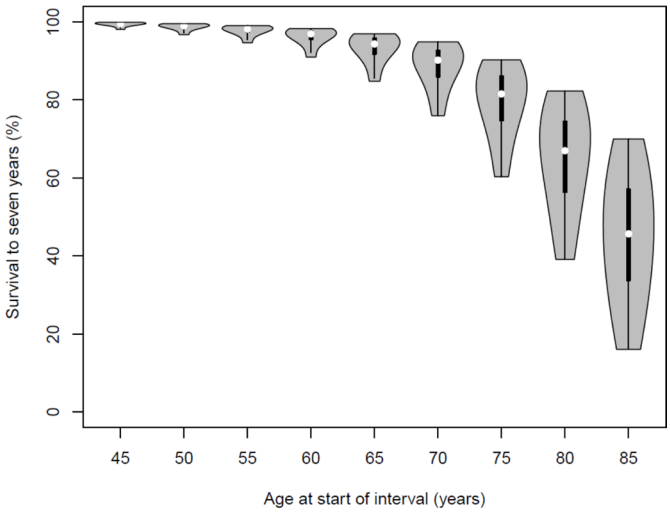

Figure 3 Violin plot of 7-year survival by age under Model 3. In the violin plot for variation in age-specific 7-year survival, the density of the risk predictions forms the left and right sides, and a boxplot forms the centre for each distribution.

A study based on participants aged 37–70 years in the UK Biobank[22] found that self-reported health was the strongest predictor of all-cause mortality in men in general (C-index 0.74 (95% CI 0.73–0.75)). The strength of this simple self-reported predictor of mortality in these data and in our data begs the question: Why propose use of a more complex model? We have no empirical data with which to answer but hypothesise men's confidence in a mortality prediction tool is likely to be greater the more of their health experience that the prediction questions cover. Our Model 1 may be perceived as too simple to provide credible information.

This study has the strength of being large and population-based, with independent and virtually complete data on the outcome, all-cause mortality. The study ascertained a range of predictors from self-reported questionnaire items; this approach has been used in other studies developing mortality prediction models.[20–22] An external form of validation was used to assess generalisability; we predicted national mortality rates from the risk model for a non-representative sample of NSW men (men in the 45 and Up Study) with national estimates of the exposure distributions.

Potential limitations of the study include, first, emigration from the study population over the follow-up period, which is unknown but we expect it to have been small because the period was short. Second, from a methodological perspective, we assume that the correlation between predictors is constant across age groups; it would be useful to investigate whether accounting for age-specific correlations would affect the predictions. Third, as is true for any such prediction model,[38] omitted individual-level and group-level heterogeneity will not be accounted for in the predictions, particularly from Model 1, which only includes age and self-reported health. Fourth, the predictors may not be causal and are not exhaustive. Moreover, most of these predictors are individual-level characteristics, which, as emphasised by Mackenzie and colleagues,[39] does not preclude the importance of place or group on health outcomes. Fifth, care should be taken when interpreting risk predictions for groups with few outcomes modelled with interaction terms. As an example, Model 1 predicts similar or lower risks for younger men with 'very good health' compared with younger men with 'excellent health', where there are few events.

We used a form of external validation, where we predicted national mortality rates from the risk model from a non-representative cohort from NSW combined with national estimates of the distribution of the risk factors. An alternative approach, which would have been preferred, is to use internal-external validation, where a large sample size dataset is split by study centre or by calendar time.[40] As this is not a multicentre study and the study recruitment was over a relatively short time period, meaningful non-random splits could not be made by place or time. The application of internal-external validation to this setting is an open avenue for future research.

Table 2 Estimated risks of death from any cause within 7 years at single years of age from 45 to 85 years

Percentiles of risk	Age at risk, years								
	45	50	55	60	65	70	75	80	85
10%	0.11	0.45	0.94	1.68	3.05	5.12	9.76	17.71	30.04
20%	0.15	0.55	1.11	1.95	3.58	6.18	12.10	21.52	37.01
30%	0.21	0.66	1.29	2.25	4.16	7.27	13.87	25.40	42.87
40%	0.28	0.81	1.51	2.61	4.93	8.44	15.81	28.98	49.13
50%	0.40	0.95	1.77	3.03	5.61	9.91	18.42	32.98	54.43
60%	0.54	1.16	2.13	3.64	6.71	11.66	21.68	37.76	60.33
70%	0.74	1.50	2.59	4.51	8.26	14.13	25.32	43.75	66.32
80%	0.99	2.04	3.40	5.83	10.48	17.90	30.62	50.77	74.37
90%	1.92	3.27	5.32	9.02	15.22	24.05	39.70	60.89	83.94

The estimated risks of death under Model 3, within 7 years from calendar year 2018 for a simulated population with covariate distributions from the Australian Health Survey.

We have developed an easily useable and credible approach to estimating mortality risk over 7 years, which requires input of self-reported health data only, for use by Australian men considering PSA testing so that they can take it into account with evidence-based guidelines when making their decision whether or not to be tested. A simple online tool is being developed to give men and their clinicians easy access to the mortality risk prediction. Output from the model when applied to the general male population suggests its impact on men's decision making will be small in the recommended age range for testing, 50–69 years, but it may discourage testing beyond this age. Its use and performance in practice should be evaluated.

Author affiliations
[1]National Centre for Epidemiology and Population Health, Research School of Population Health, Australian National University, Canberra, Australian Capital Territory, Australia
[2]Sax Institute, Haymarket, New South Wales, Australia
[3]Menzies Health Institute Queensland, Griffith University, Brisbane, Queensland, Australia
[4]School of Public Health and Preventive Medicine, Monash University, Melbourne, Victoria, Australia
[5]Department of Applied Finance and Actuarial Studies, Macquarie University, Sydney, New South Wales, Australia
[6]School of Population and Global Health, University of Western Australia, Perth, Western Australia, Australia
[7]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

Acknowledgements This research was completed using data collected through the 45 and Up Study (http://www.saxinstitute.org.au). The 45 and Up Study is managed by the Sax Institute in collaboration with major partner Cancer Council NSW and partners: the National Heart Foundation of Australia (NSW Division); NSW Ministry of Health; NSW Government Family & Community Services—Carers, Ageing and Disability Inclusion to NSW Government Family & Community Services—Ageing, Carers and the Disability Council NSW and the Australian Red Cross Blood Service. We thank the many thousands of people participating in the 45 and Up Study.

Contributors BA, AL and EB conceived the project. GJ and MC wrote the analysis plan. GJ and MC conducted the analysis and drafted the initial version of the manuscript, with methodological input from RW and LT. All authors contributed to a review of the analysis plan, interpretation of results and revisions of the manuscript.

Funding This research was cofunded by Actuaries Institute and Prostate Cancer Foundation of Australia. A VICBiostat visiting fellowship awarded to author GJ supported the statistical analysis.

Competing interests None declared.

Patient consent Not required.

Ethics approval NSW Population and Health Services Research Ethics Committee, the University of NSW Human Research Ethics Committee and the Australian National University Human Research Ethics Committee.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Information about data access and governance policies is available at: https://www.saxinstitute.org.au/our-work/45-up-study/for-researchers/.

REFERENCES

1. Lansdorp-Vogelaar I, Gulati R, Mariotto AB, *et al*. Personalizing age of cancer screening cessation based on comorbid conditions: model estimates of harms and benefits. *Ann Intern Med* 2014;161:104–12.
2. Pollack CE, Blackford AL, Schoenborn NL, *et al*. Comparing prognostic tools for cancer screening: considerations for clinical practice and performance assessment. *J Am Geriatr Soc* 2016;64:1032–8.
3. Kotwal AA, Mohile SG, Dale W. Remaining Life Expectancy Measurement and PSA Screening of Older Men. *J Geriatr Oncol* 2012;3:196–204.
4. Prostate Cancer Foundation of Australia and Cancer Council Australia PSA Testing Guidelines Expert Advisory Panel. *Clinical practice guidelines for PSA testing and early management of test-detected prostate cancer*. Sydney: Prostate Cancer Foundation of Australia and Cancer Council Australia, 2016.
5. Carter HB, Albertsen PC, Barry MJ, *et al*. Early detection of prostate cancer: AUA Guideline. 2013 https://www.auanet.org/education/guidelines/prostate-cancer-detection.cfm
6. American Cancer Society. Prostate cancer prevention and early detection. http://www.cancer.org/acs/groups/cid/documents/webcontent/003182-pdf.pdf
7. Carmichael LK, Goldsbury DE, O'Connell DL. Prostate cancer screening for men aged 75 to 84 years in New South Wales. *Aust N Z J Public Health* 2013;37:492–4.
8. Warnier RM, van Rossum E, van Velthuijsen E, *et al*. Validity, reliability and feasibility of tools to identify frail older patients in inpatient hospital care: a systematic review. *J Nutr Health Aging* 2016;20:218–30.
9. Yourman LC, Lee SJ, Schonberg MA, *et al*. Prognostic indices for older adults: a systematic review. *JAMA* 2012;307:182–92.
10. Duarte CW, Black AW, Murray K, *et al*. Validation of the Patient-Reported Outcome Mortality Prediction Tool (PROMPT). *J Pain Symptom Manage* 2015;50:241–7.
11. Eagle KA, Lim MJ, Dabbous OH, *et al*. A validated prediction model for all forms of acute coronary syndrome: estimating the risk of 6-month postdischarge death in an international registry. *JAMA* 2004;291:2727–33.
12. Mattishent K, Kwok CS, Mahtani A, *et al*. Prognostic indices for early mortality in ischaemic stroke - meta-analysis. *Acta Neurol Scand* 2016;133:41–8.
13. Aydede SK, Komenda P, Djurdjev O, *et al*. Chronic kidney disease and support provided by home care services: a systematic review. *BMC Nephrol* 2014;15:118.
14. Rodrigues G, Bauman G, Palma D, *et al*. Systematic review of brain metastases prognostic indices. *Pract Radiat Oncol* 2013;3:101–6.
15. Dijk WD, Bemt L, Haak-Rongen S, *et al*. Multidimensional prognostic indices for use in COPD patient care. A systematic review. *Respir Res* 2011;12:151.
16. Gravante G, Garcea G, Ong SL, *et al*. Prediction of mortality in acute pancreatitis: a systematic review of the published evidence. *Pancreatology* 2009;9:601–14.
17. Kent M, Penson DF, Albertsen PC, *et al*. Successful external validation of a model to predict other cause mortality in localized prostate cancer. *BMC Med* 2016;14:25.
18. Su TL, Jaki T, Hickey GL, *et al*. A review of statistical updating methods for clinical prediction models. *Stat Methods Med Res* 2018;27.
19. Banks E, Redman S, Jorm L, *et al*. Cohort profile: the 45 and Up Study. *Int J Epidemiol* 2008;37:941–7.
20. Lee SJ, Lindquist K, Segal MR, *et al*. Development and validation of a prognostic index for 4-year mortality in older adults. *JAMA* 2006;295:801–8.
21. Schonberg MA, Davis RB, McCarthy EP, *et al*. Index to predict 5-year mortality of community-dwelling adults aged 65 and older using data from the National Health Interview Survey. *J Gen Intern Med* 2009;24:1115–22.
22. Ganna A, Ingelsson E. 5 year mortality predictors in 498,103 UK Biobank participants: a prospective population-based study. *Lancet* 2015;386:533–40.
23. Hays R, Sherbourne D, Mazel R. *User's manual for the medical outcomes study (MOS) core: measures of health-related quality of life*. Santa Monica, California: RAND, 1995.
24. Stewart A, Kamberg C. *Measuring functioning and well-being: the Medical Outcomes Study Approach*. Durham, North Carolina: Duke University Press, 1992.
25. Wolbers M, Blanche P, Koller MT, *et al*. Concordance for prognostic models with competing risks. *Biostatistics* 2014;15:526–39.
26. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med* 2011;30:377–99.

27. Marshall A, Altman DG, Holder RL, *et al*. Combining estimates of interest in prognostic modelling studies after multiple imputation: current practice and guidelines. *BMC Med Res Methodol* 2009;9:57.
28. Australian Government Actuary. *Australian Life Tables 2010-12 Commonwealth of Australia*: Australian Government Actuary, 2014.
29. AIHW. *General Record of Incidence of Mortality (GRIM) books - All Causes Combined*: AIHW, 2012.
30. Booth H, Maindonald J, Smith L. Applying Lee-Carter under conditions of variable mortality decline. *Popul Stud* 2002;56:325–36.
31. Lee RD, Carter LR. Modelling and forecasting U.S. mortality. *Journal of the American Statistical Association* 1992;87:659–71.
32. Human Mortality Database. (data downloaded on 3 August 2015): University of California, Berkeley (U.S.A.) and Max Planck Institute for Demographic Research (Germany). 2010. Available at www.mortality.org or www.humanmortality.de
33. Chiang CL. *The life table and its applications*: R.E. Krieger Publishing Company, 1984.
34. *SAS Version 9.4 [program]*. NC, USA: SAS Institute Inc, 2013.
35. Demography: Forecasting mortality, fertility, migration and population data. With contributions from Heather Booth, Leonie Tickle and John Maindonald, R package version 1.18. 2015 https://cran.r-project.org/web/packages/demography/
36. Goldman N, Glei DA, Weinstein M. What Matters Most for Predicting Survival? A Multinational Population-Based Cohort Study. *PLoS One* 2016;11:e0159273.
37. Tan A, Kuo YF, Goodwin JS. Predicting life expectancy for community-dwelling older adults from Medicare claims data. *Am J Epidemiol* 2013;178:974–83.
38. Ahmed I, Debray TP, Moons KG, *et al*. Developing and validating risk prediction models in an individual participant data meta-analysis. *BMC Med Res Methodol* 2014;14:3.
39. Mackenzie M, Collins C, McCartney G, *et al*. Troubled by Ubble. *Lancet* 2015;386:2055.
40. Debray TP, Moons KG, Ahmed I, *et al*. A framework for developing, implementing, and evaluating clinical prediction models in an individual participant data meta-analysis. *Stat Med* 2013;32:3158–80.