



ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

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

A Modeling Study of the Cost-Effectiveness of a Risk-Stratified Surveillance Program for Melanoma in the United Kingdom

Edward C.F. Wilson, PhD^{1,*}, Juliet A. Usher-Smith, PhD², Jon Emery, DPhil³, Pippa Corrie, FRCP⁴, Fiona M. Walter, FRCGP²

¹Cambridge Centre for Health Services Research, University of Cambridge, Cambridge, UK; ²Primary Care Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK; ³Department of General Practice, Centre for Cancer Research, Faculty of Medicine, Dentistry and Health Science, Victorian Comprehensive Cancer Centre, University of Melbourne, Melbourne, Victoria, Australia; ⁴Cambridge Cancer Centre, Cancer Research UK Cambridge Institute, Li Ka Shing Centre, University of Cambridge, Cambridge, UK

ABSTRACT

Background: Population-wide screening for melanoma is unlikely to be cost-effective. Nevertheless, targeted surveillance of high-risk individuals may be. **Objectives:** To estimate the cost-effectiveness of various surveillance strategies in the UK population, stratified by risk using a simple self-assessment tool scoring between 0 and 67. **Methods:** A decision model comparing alternative surveillance policies from the perspective of the UK National Health Service over 30 years was developed. The strategy with the highest expected net benefit for each risk score was identified, resulting in a compound risk-stratified policy describing the most cost-effective population-wide strategy. The overall expected cost and quality-adjusted life-years (QALYs), the incremental cost-effectiveness ratio, and associated uncertainty were reported. **Results:** The most cost-effective strategy is for those with a Williams score of 15 to 21 (relative risk [RR] of 0.79–1.60 vs. a mean score of 17 in the United Kingdom) to be offered a one-off full-body skin examination, and for those with a

score of 22 or more (RR 1.79+) to be enrolled into a quinquennial monitoring program, rising to annual recall for those with a risk score greater than 43 (RR 20.95+). Expected incremental cost would be £164 million per annum (~0.1% of the National Health Service budget), gaining 15,947 additional QALYs and yielding an incremental cost-effectiveness ratio of £10,199/QALY gained (51.3% probability <£30,000). **Conclusions:** The risk-stratified policy would be expensive to implement but cost-effective compared with typical UK thresholds (£20,000–£30,000/QALY gained), although decision uncertainty is high. Phased implementation enrolling only higher risk individuals would be substantially less expensive, but with consequent foregone health gain. **Keywords:** cost-effectiveness, decision modeling, economic evaluation, melanoma, screening.

Copyright © 2018, International Society for Pharmacoeconomics and Outcomes Research (ISPOR). Published by Elsevier Inc.

Introduction

Approximately 14,500 new cases of malignant melanoma are diagnosed and approximately 2,600 deaths occur in the United Kingdom every year [1]. Early detection is critical: 90% of patients survive for 5 or more years, but this falls to 25% of women and less than 10% of men with metastatic disease at diagnosis [1]. The cost of treating metastatic melanoma far outweighs the cost of treating primary melanoma, and the relative increase has risen sharply with the recent introduction of several high-cost drugs that palliate for the most part. For example, nivolumab costs approximately £70,000 per patient per year for an additional gain of 1.3 quality-adjusted life-years (QALYs) compared with dacarbazine [2]. Screening programs are therefore of increasing relevance. The UK National Screening Committee has not formally reviewed whether a program for melanoma would be an efficient

use of public funds [3]. Nevertheless, existing evidence suggests such a program would have difficulty identifying the target population [4], raises concerns about whether a comprehensive program could be cost-effective [5], and cites lack of evidence on the cost-effectiveness of full-body skin examination (FBSE), except in those with a history of melanoma [6].

Two recent systematic reviews [7,8] concluded that although skin cancer prevention initiatives are highly cost-effective [7], there is a lack of evidence on the cost-effectiveness of early detection programs [7], and future research should focus on targeted screening/surveillance in high-risk populations [8]. On the basis of this, the US Preventive Services Task Force (2016) reiterated its previous recommendation [9] that the “current evidence is insufficient to assess the balance of benefits and harms of visual skin examination ... to screen for skin cancer in adults” [10].

*Address correspondence to: Edward C.F. Wilson, Cambridge Centre for Health Services Research, University of Cambridge, Cambridge CB2 0SR, UK.

E-mail: ed.wilson@medschl.cam.ac.uk

1098-3015/\$36.00 – see front matter Copyright © 2018, International Society for Pharmacoeconomics and Outcomes Research (ISPOR).

Published by Elsevier Inc.

<https://doi.org/10.1016/j.jval.2017.11.009>

Several tools have been developed to enable identification of higher risk individuals [11]. One of the better performing was developed from a case-control study in the United States by Williams et al. [12]. It is a self-assessed clinical risk estimation model not requiring expert FBSE that, in a split-sample validation population, had an area under the receiver operator characteristic curve of 0.70 (95% confidence interval 0.64–0.77) and was able to identify 15% of the population in whom 50% of melanomas would be expected to develop [12]. We have recently shown that it is both feasible and acceptable to collect data on the risk of melanoma in the waiting rooms of UK family practices and that using the Williams model produces a distribution of risk in the attending population, which allows identification of subgroups at different levels of risk [13].

The purpose of this study was to establish whether using the Williams model and resulting score to risk-stratify the population and guide future management is a cost-effective approach to reducing mortality and morbidity from melanoma in a UK setting. Key to this is determining the risk score at which it is most cost-effective to enroll patients into a surveillance program. If the score is set too low, primary care capacity will be absorbed examining patients with an extremely low risk of melanoma at the expense of other patients with a greater capacity to benefit. If set too high, then patients will be falsely reassured and any benefits in terms of reduced melanoma morbidity and mortality will be foregone. Specifically, therefore, this study aimed to identify the optimal cutoff scores from the Williams self-assessment tool [12] at which users are recommended to either 1) visit their primary care practitioner for a one-off FBSE or 2) be entered into a routine primary care-based monitoring program, and if so, 3) the optimal frequency of visits, ranging from 5-yearly to annually.

Methods

We substantially adapted and modified a decision model we previously developed for a novel diagnostic aid for melanoma [14]. The adapted model was a patient-level simulation following a simulated cohort of participants (UK general public) one by one. Uncertainty was propagated through the model via Monte-Carlo simulation (distributions of parameters are specified in Table 1). The code was written in R (R Foundation for Statistical Computing, Vienna, Austria) [15–17] and run on the University of Cambridge High Power Cluster computing facility. The code is available on request from the corresponding author. Ethical approval was not required for this study.

The Williams Self-Assessment Tool

The scenarios we model focus on the Williams self-assessment tool (Appendix 1) [12]. This is a rapid questionnaire comprising eight questions on sex, age, hair color, density of freckles, history of severe sunburn in childhood and adolescence, number of raised moles on the arms, and history of nonmelanoma skin cancer yielding a summary score between 0 (lowest risk) and 67.

Model Definition

The model comprises two modules: natural history and clinical (Fig. 1). The link between the two is determined by the comparator policies, described later. Cohorts of a given age, sex, and Williams score [12] are simulated. In year 0, the distribution of prevalent melanomas and their disease stages in each cohort is estimated on the basis of UK prevalence data and stage at diagnosis [18,19] adjusted for risk score. The natural history module is a Markov-like model and simulates

patients' trajectories over a period of 30 years: each year patients are at risk of new melanomas developing according to UK incidence by age and sex [19] adjusted for risk score [12], and undiagnosed (and hence untreated) melanomas progress according to estimated rates of progression [20]. When the model determines that contact is made with the health service, the simulated patient "breaks out" of the natural history module into the clinical module, which has a decision-tree structure. Once reaching a terminal node of the decision tree, the patient is returned to the natural history module.

Natural history module

Cutaneous melanoma is categorized into four main types (superficial spreading, lentigo maligna, acral lentiginous, and nodular) [21], each with nine stages of invasion (stages 1a–4) plus an in situ stage for all except nodular melanoma (which is by definition invasive) [22]. We assumed that invasive disease would progress at the same rate irrespective of primary melanoma subtype, but allowed the rate of progression from in situ disease to vary by subtype, yielding a total of 12 discrete stages describing the disease. The model also included "no melanoma" and "dead" health states. The overall prevalence of undiagnosed melanoma in the community in year 0 was estimated at 0.162%, assumed the same as that observed in a population screening study in Northern Germany [18] (review details are given in Appendix 2). This was distributed according to risk score by combining with UK-relevant epidemiological data [12,19,23,24]. The parameters of the resulting risk function are presented in Table 1. The annual incidence was estimated using an analogous approach. Full details are provided in Appendix 3.

Data on the rate of progression of untreated melanoma do not exist and it would be most unethical to conduct a prospective cohort study to establish this empirically. Therefore, data were elicited from a representative group of experts in melanoma [20] (Table 1; Appendix 4). Age- and sex-specific background and melanoma-specific mortality data were extracted from UK life tables [25] (Appendix 5) adjusted for the odds ratio [22] (Appendix 6).

Clinical module

The clinical module describes the patient pathway after health service contact (Fig. 1). The model allows two ways for patients to present in primary care: of their own initiative with a mole that they are concerned about or because they have been advised to do so after a risk assessment. Any suspicious moles are inspected during an FBSE by a primary care practitioner, and the patient is either referred to secondary care or discharged. Figure 1 (right-hand side) illustrates the pathway; the natural history component of the model will have determined whether a patient is healthy (D–) or has melanoma (D+). For a patient with a melanoma, the probability of the primary care practitioner identifying it and referring a patient to secondary care is the sensitivity of the practitioner, denoted $P(T+|D+)$, and is based on data from the control arm of a recent study of a diagnostic aid in primary care [26]. Likewise, the probability of correctly discharging a patient without melanoma is the specificity (denoted $P(T-|D-)$ in Fig. 1) extracted from the same source. Data are summarized in Table 1.

Patients with melanoma correctly referred (true positives, with probability $P(T+|D+)$) receive appropriate treatment according to disease stage ($D&T_{stage}$ in Fig. 1; see the "Costs" section later for details). They are then flagged as having a history of melanoma and are at risk of mortality as described in the natural history module (data based on stage-specific prognosis postdiagnosis [22]). Patients with melanoma who are

Table 1 – Input parameters.

Parameter	Distribution	Hyperparameters	Mean	Median	SE	95% CrI	Source
<i>Probability of prevalent melanoma in year 0 by risk score*</i>							
α	N	(-8.454, 0.119)	-8.454		0.119		
β	N	(0.100, 0.008)	0.100		0.008		
<i>Annual probability of incident melanoma by risk score*</i>							
α	N	(-10.270, 0.186)	-10.270		0.186		
β	N	(0.117, 0.008)	0.117		0.008		
<i>Transition probabilities (from state > to state)[†]</i>							
ISLM > ISLM	mCM	(0.167 0.102 0.023 0.999)		0.92		0.02–1.00	[1]
ISLM > 1A		(9.624 2.133 0.073 0.956)		0.06		0.00–0.88	
ISLM > 1B		(9.885 9.988 0.039 0.604)		0.00		0.00–0.11	
ISLM > 2A		(2.412 9.960 0.101 0.994)		0.00		0.00–0.07	
ISLM > 2B		(9.964 3.223 0.019 0.665)		0.00		0.00–0.09	
ISLM > 2C		(9.810 9.970 0.649 0.964)		0.00		0.00–0.07	
ISLM > 3A				0.00		0.00–0.02	
ISSS > ISSS	mCM	(0.739 0.319 0.003 0.996)		0.83		0.03–1.00	
ISSS > 1A		(2.069 2.033 0.610 1.000)		0.14		0.00–0.82	
ISSS > 1B		(9.055 6.031 0.012 0.973)		0.02		0.00–0.15	
ISSS > 2A		(9.687 2.788 0.034 0.903)		0.01		0.00–0.08	
ISSS > 2B		(7.133 9.546 0.264 0.990)		0.00		0.00–0.02	
ISSS > 2C		(0.450 0.137 0.038 0.621)		0.00		0.00–0.01	
ISSS > 3A				0.00		0.00–0.01	
ISAL > ISAL	mCM	(0.839 0.386 0.000 0.999)		0.79		0.03–1.00	
ISAL > 1A		(1.492 0.248 0.000 0.759)		0.13		0.00–0.70	
ISAL > 1B		(10.000 10.000 0.000 0.969)		0.03		0.00–0.23	
ISAL > 2A		(6.603 8.078 0.000 0.777)		0.01		0.00–0.09	
ISAL > 2B		(8.718 9.295 0.000 0.671)		0.01		0.00–0.05	
ISAL > 2C		(9.634 10.000 0.247 0.810)		0.01		0.00–0.06	
ISAL > 3A		(8.705 9.196 0.000 0.973)		0.00		0.00–0.02	
ISAL > 3B				0.00		0.00–0.00	
1A > 1A	mCM	(1.034 0.568 0.000 0.999)		0.72		0.05–1.00	
1A > 1B		(9.734 9.973 0.004 1.000)		0.14		0.00–0.53	
1A > 2A		(2.932 9.951 0.002 1.000)		0.03		0.00–0.16	
1A > 2B		(0.052 0.233 0.009 0.985)		0.00		0.00–0.26	
1A > 3A		(9.865 9.992 0.000 0.365)		0.01		0.00–0.08	
1A > 2C		(0.011 0.126 0.031 0.989)		0.00		0.00–0.10	
1A > 3B		(9.867 9.903 0.021 0.993)		0.02		0.00–0.17	
1A > 3C		(0.000 9.995 0.014 0.962)		0.00		0.00–0.00	
1A > 4		(0.000 9.971 0.009 0.999)		0.00		0.00–0.00	
1A > ISSS				0.02		0.00–0.16	
1B > 1B	mCM	(0.448 0.294 0.098 0.924)		0.69		0.10–0.92	
1B > 2A		(10.000 9.915 0.000 0.957)		0.14		0.03–0.52	
1B > 2B		(0.893 1.659 0.000 0.963)		0.03		0.00–0.33	
1B > 2C		(0.000 9.696 0.393 0.910)		0.03		0.00–0.18	
1B > 3A		(0.232 0.398 0.000 0.939)		0.01		0.00–0.18	
1B > 3B		(9.841 9.398 0.000 0.934)		0.01		0.00–0.12	
1B > 3C		(9.526 9.769 0.065 0.968)		0.01		0.00–0.07	
1B > 4		(9.809 0.845 0.081 0.931)		0.01		0.00–0.06	
1B > 1A				0.00		0.00–0.00	
2A > 2A	mCM	(1.688 1.004 0.000 0.918)		0.61		0.11–0.91	
2A > 2B		(2.725 2.357 0.023 0.962)		0.18		0.03–0.60	
2A > 3A		(8.995 9.106 0.000 0.629)		0.05		0.01–0.19	
2A > 2C		(0.524 0.623 0.220 0.884)		0.05		0.01–0.25	
2A > 3B		(5.262 8.588 0.266 0.691)		0.02		0.00–0.10	
2A > 4		(0.078 1.900 0.470 0.818)		0.01		0.00–0.07	
2A > 3C				0.01		0.00–0.05	

continued on next page

Table 1 – continued

Parameter	Distribution	Hyperparameters	Mean	Median	SE	95% CrI	Source
2B > 2B	mCM	(0.260 0.257 0.009 0.994)	0.52	0.14	0.01–0.99	0.00–0.49	
2B > 2C							
2B > 3A							
2B > 3B							
2B > 3C							
2B > 4							
2B > 2A							
2C > 2C	mCM	(1.977 2.396 0.014 0.954)	0.44	0.14	0.09–0.83	0.01–0.82	
2C > 3A							
2C > 3B							
2C > 3C							
2C > 4							
2C > 2B							
3A > 3A	mCM	(1.393 0.993 0.013 0.977)	0.61	0.15	0.08–0.96	0.01–0.63	
3A > 3B							
3A > 3C							
3A > 4							
3A > 1A							
3A > 1B							
3A > 2A							
3A > 2B							
3A > 2C							
3B > 3B							
3B > 3C							
3B > 4							
3C > 3C	mCM	(3.785 2.926 0.008 0.988)	0.57	0.31	0.21–0.87	0.08–0.66	
3C > 4							
<i>Odds ratio of death by disease stage (vs. 1A)[‡]</i>							
1A	–	–	1		0		[2]
1B	LN	(1.449, 0.007)	4.261		0.007		
2A	LN	(2.506, 0.007)	12.250		0.007		
2B	LN	(3.045, 0.007)	21.000		0.007		
2C	LN	(3.731, 0.008)	41.741		0.008		
3A	LN	(2.626, 0.010)	13.821		0.010		
3B	LN	(3.486, 0.008)	32.667		0.008		
3C	LN	(4.215, 0.011)	67.667		0.011		
4	LN	(5.743, 0.006)	312.104		0.006		
<i>Sensitivity and specificity of primary care practitioner at detecting melanoma</i>							
Sensitivity, P(T+ D+)	β	(29, 7)	80.1%			66.3%–91.6%	[3]
Specificity, P(T– D–)	B	(864, 536)	61.7%			59.2%–64.2%	[3]
<i>Costs</i>							
GP consultation (per min)		Constant	£3.80				[4]
GP consultation time (min)	N	(22.1, 3.2)	22.1				[3]
Initial referral	Γ	(20.408, 5.439)	£111				[5]
Biopsy excision	Γ	(10.443, 13.694)	£143				[5]
Definitive surgery	Γ	(3.762, 41.468)	£156				[5]
CXR	Γ	(12.430, 2.414)	£30				[5]
CT scan	Γ	(13.616, 9.695)	£132				[5]
Liver function test	Γ	(4.041, 0.742)	£3				[5]
FBC	Γ	(4.041, 0.742)	£3				[5]
Sentinel node biopsy	Γ	(1.165, 24.887)	£29				[5]
Radical lymph node dissection	Γ	(1.808, 547.925)	£991				[5]
Surgical removal of localized metastases	Γ	(1.256, 577.101)	£725				[5]
Radiotherapy (planning)	Γ	(8.890, 82.673)	£735				[5]
Radiotherapy (per fraction)	Γ	(17.014, 7.758)	£132				[5]

continued on next page

Table 1 – continued

Parameter	Distribution	Hyperparameters	Mean	Median	SE	95% CrI	Source
Chemotherapy (dacarbazine, procurement)	Γ	(1.330, 209.827)	£279				[5]
Chemotherapy (dacarbazine, delivery first attendance)	Γ	(6.934, 26.823)	£186				[5]
Chemotherapy (dacarbazine, delivery subsequent)	Γ	(3.239, 62.988)	£204				[5]
Dermatology follow-up	Γ	(12.183, 7.962)	£97				[5]
<i>Summary costs</i>							
GP consultation			£83.98				
Chemotherapy, cycle			£1485				
Radiotherapy, 10-fraction cycle			£2055				
<i>By disease stage</i>							
D&T in situ (SS, LM, AL)			£396				
D&T stage 1a, 1b			£1463				
D&T stage 2a			£1880				
D&T stage 2b, 2c			£2048				
D&T stage 3a, 3b, 3c			£3171				
D&T stage 4			£4761				
Final year of life in situ, 1a (i.e., states “dead IS” and “dead 1a”)			£0				
Final year of life stage 1b–4			£4265				
<i>Health state utilities</i>							
No melanoma	C	–	1.00		–	–	
Undetected disease	C	–	1.00		–	–	
<i>Detected and treated, and post-D&T</i>							
Stage 0	N	(0.93, 0.013)	0.93				[6]
Stage 1	N	(0.93, 0.013)	0.93				[6]
Stage 2	N	(0.87, 0.057)	0.87				[6]
Stage 3	N	(0.89, 0.046)	0.89				[6]
Stage 4	N	(0.52, 0.117)	0.52				^s
Last year of life with in situ or 1a disease (states “dead 0” and “dead 1a”)	N	(0.93, 0.013)	0.93				[6]
Last year of life with stage 1b–4 disease	N	(0.52, 0.117)	0.52				[6]
Dead	C	–	0				

1A, 1B, etc., invasive melanoma of stage 1A, 1B, etc.; AL, acral lentiginous; C, constant; CrI, credibility interval; CXR, chest x-ray; D&T, diagnosed and treated; FBC, full blood cell count; GP, general practitioner; ISAL, in situ acral lentiginous melanoma; ISLM, in situ lentigo maligna melanoma; ISSS, in situ superficial spreading melanoma; LM, lentigo maligna; LN, lognormal; mCM, modified Connor-Mosimann distribution; N, normal; SE, standard error; SS, superficial spreading.

* $P(\text{prevalent or incident melanoma in year } 0) = e^{(\alpha + \beta W)}$, where W is risk score.

† Parameters are presented as a $(k - 1) \times 4$ matrix. Parameters relate to 6-mo transition probabilities (elicited probabilities were over a 6-mo and not a 12-mo time horizon), which form the inputs to the mCM. Columns representing mean, median, 95% CrI represent the respective moments of the 12-mo transition marginal probability distribution. Please see Appendix 4 for further details. Note medians are aggregated from individual summaries and so do not total 100%.

‡ Parameters of lognormal distribution are the natural log of the mean and the SE of the natural log of the mean.

^s Assumption based on Ref. [6].

not referred (false negatives, with probability $1 - P(T+ID+)$) are discharged and they return to the natural history module in which they are at risk of disease progression and mortality. Patients without melanoma who are referred incur the cost of referral followed by discharge to the community. Finally, patients without melanoma who are not referred are reassured and discharged back to the community directly from primary care.

Contact with Health Service: Comparator Policies

There are seven alternative policies. The first is the status quo. This assumes an ad hoc presentation by a member of the public concerned about a skin lesion, the probability of which is estimated at 0.73% per annum [19,24–28] (Appendix 7). The second policy is to invite all at-risk persons to primary care for a one-off FBSE by a primary care practitioner in year 0;

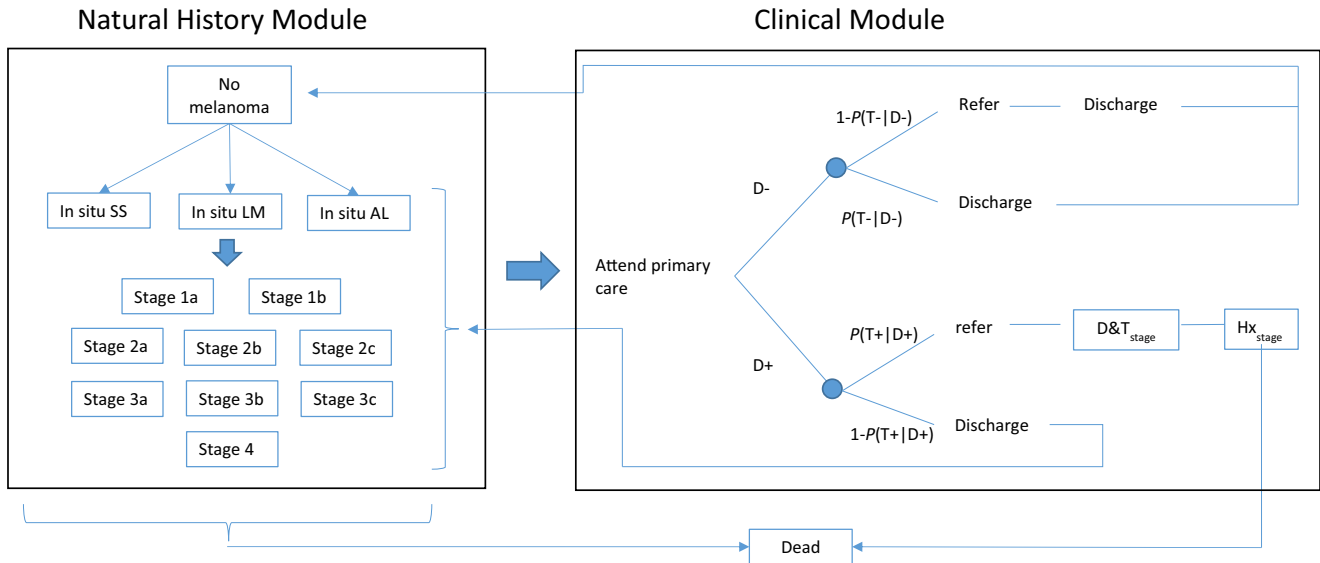


Fig. 1 – Model schematic. AL, acral lentiginous; D+/D–, patient with or without melanoma; $P(T+|D+)$, probability of a positive diagnosis given the patient has melanoma (sensitivity); $P(T-|D-)$, probability of a negative diagnosis given the patient does not have melanoma (specificity); $D\&T_{stage}$, diagnosis and treatment in secondary care according to disease stage; Hx_{stage} , patient with history of treated disease of a given stage; LM, lentigo maligna; SS, superficial spreading.

patients then present opportunistically (with lesions of concern) in the remaining years. Policies 3 to 7 represent enrollment into a primary care–based monitoring program with increasing frequency of recall from 5-yearly to annually. Thus, under policy 3, patients attend for a body examination in years 0, 5, 10, 15, and so on (with ad hoc presentation in the intervening years). Under policy 4, patients present in years 0, 4, 8, 12, and so forth.

In policies 2 to 7, we assume that the Williams tool is used by members of the public to assess their own risk before contact with the health service. This could be administered, for example, via a leaflet in pharmacies, in primary care waiting rooms or other public places, or electronically via a smartphone app. The objective of this analysis is to determine the optimal cutoff scores at which each of the seven policies is recommended.

Costs

The perspective and price year of the analysis is the UK National Health Service (NHS) and 2015, with future costs discounted at a rate of 3.5%. Unit costs were extracted from standard NHS sources [29,30], and care pathways and primary care consultation time were from current guidelines [31] and a recent clinical trial [26] (Table 1; Appendix 8).

Health State Utilities

A systematic review of health-related quality of life in patients with melanoma identified three distinct periods of impact of the disease: at diagnosis, treatment, and follow-up [32]. We assumed that patients who are unaware they have melanoma suffer no impairment in quality of life (assigned a utility of 1), whereas from the point of diagnosis, a health utility impairment was assigned as per the authors' previous model in a related area [14], adapted from a study of health-related quality-of-life measurement in patients with melanoma [33] (Table 1).

Model Calculation and Analysis

To determine the appropriate cohort size and number of iterations, one of the seven scenarios under one age/sex/risk score (baseline scenario: 35-year-old male, risk score 17) was run under a range of cohort sizes and iterations for a total of 50 times. The coefficient of variation (CV) of the expected values and standard errors (SEs) of cost and QALYs were calculated from these, with a “target” CV of 2% or less considered “stable.” A cohort size of 1000 and 1000 iterations yielded CVs of 0.39%, 2.21%, 0.01%, and 1.93% for mean cost, SE mean cost, mean QALYs, and SE mean QALYs, respectively (Appendix 9).

The model therefore generates 1000 patients of a given age, sex, and risk score group and simulates their development, progression, and treatment of melanoma over 30 years under each of the seven policies 1000 times. We estimated the expected cost, QALYs, and net benefit (defined as the QALY gain multiplied by the thresholds of £20,000 and £30,000 less the cost) for each policy. The model was calculated for each of the seven policies at seven selected values for the Williams risk score (10, 17, 20, 25, 30, 50, and 60), males and females, and four starting ages (35, 45, 55, and 65 years). Results were weighted for the age/sex of the UK population to yield costs, outcomes, and net benefits by risk score alone. The risk scores at which the model was evaluated were chosen to get a spread of scores, but included 17 because this is the mean risk score for the UK population [13]. Net benefits for intervening scores were estimated by linear interpolation.

The policy yielding the highest net benefit was noted for each risk score, and the cutoff scores at which the optimal policy changed were identified. This risk-stratified policy describes the most cost-effective strategy given the epidemiology and demographic characteristics of the UK population. The overall expected cost and QALY gains of this risk-stratified policy are applied to the UK population and compared with the expected cost and QALYs of the status quo, thus estimating the overall incremental cost per QALY gained of the stratified policy versus status quo.

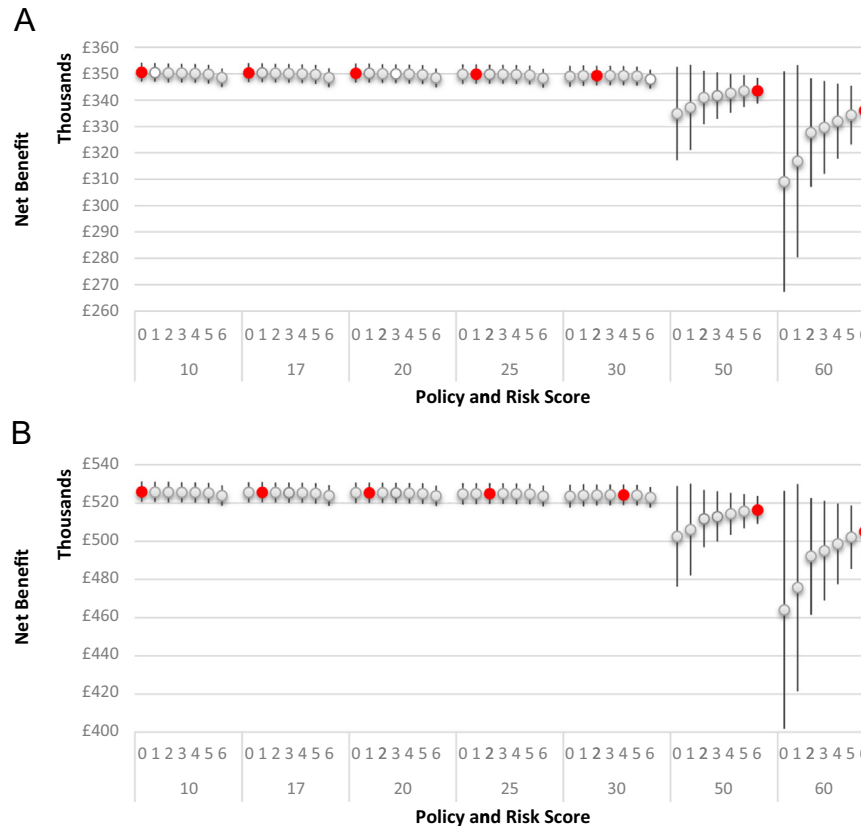


Fig. 2 – (A) Expected net benefit \pm 95% credibility interval, threshold = £20,000/QALY. (B) Expected net benefit \pm 95% credibility interval, threshold = £30,000/QALY. Policies are ranked in order of intensity: 0 = status quo; 1 = one-off examination; 2–6 = enrollment in monitoring program with examination every 5 y to annually, respectively. Expected net benefit and 95% credibility intervals are shown for each policy for each of the seven example risk scores. The option with the highest expected net benefit at each risk score is identified in red. For example, given a willingness-to-pay threshold of £30,000/QALY, the most cost-effective strategy for those with a risk score of 20 is a one-off FBSE, whereas those with a risk score of 30 should be enrolled in a monitoring program with 3-yearly recall. FBSE, full-body skin examination; QALY, quality-adjusted life-year.

Results

Figure 2A and 2B shows the expected net benefit and 95% credibility intervals from each policy as a function of selected risk scores at a willingness-to-pay threshold of £20,000 and £30,000/QALY, respectively (data in Appendix 10; additional figures in Appendix 11). At lower risk scores, all options have a very similar expected net benefit. As the risk score increases, the expected net benefit of status quo (no screening) and the less intensive policies drop below that of the more intensive policies.

Given a willingness-to-pay threshold of £30,000/QALY, the optimal policy is for those with a risk score between 15 and 21 to be offered a one-off FBSE to check for melanoma. Those with a risk score of 22 and more should be enrolled into a monitoring program with quinquennial recall, rising to annual recall for those with a score greater than 43.

If this “compound” policy were to be enacted across the United Kingdom, the expected additional cost per person older than 30 years would be £164.89, yielding an extra 0.016 QALYs per person. The incremental cost per QALY gained is thus £10,199 (Table 2; Appendix 10). The 95% credibility ellipse (Fig. 3) and cost-effectiveness acceptability curve (Fig. 4) illustrate the high decision uncertainty; at the National Institute for Health and Care Excellence’s threshold of about £20,000 to £30,000, there is only a 51.0%

probability that the policy is cost-effective. Thus, although the compound risk-adjusted policy yielding the highest expected net benefit can be identified (Table 2), there is a great deal of decision uncertainty. This is a function of both the small absolute difference in net benefit between policies at lower risk scores (Fig. 2) and the substantial parameter uncertainty (Table 1). The probability of cost-effectiveness does not exceed 51% to 52% because of the proportion of the probability mass in the northwest quadrant of the cost-effectiveness plane, representing scenarios when the policy is more expensive but less effective (i.e., yields fewer QALYs) than the status quo (Fig. 3).

Discussion

Interpretation of Results

The results in Table 2 and Figure 2 show how the recommended intensity of surveillance increases with risk score, from status quo (no screening program) rising to enrollment in a monitoring program of increasing frequency of recall for the highest risk individuals. This suggests that the model has face validity. We are able to identify the most efficient cutoff scores for these recommendations: all those with a Williams risk score greater

Table 2 – Recommendations by risk score.

Risk score	Equivalent RR	Optimal policy	% of population	Status quo*			Monitoring program†		
				£ (SE)	QALYs (SE)	Cov (£, Q)	£ (SE)	QALYs (SE)	Cov (£, Q)
0–14	0.14–0.70	Do nothing	38.6%	£34.99 (£6.79)	28.388 (0.346)	–0.449	£34.99 (£6.79)	28.388 (0.346)	–0.449
15–21	0.79–1.60	One-off examination	33.9%	£24.99 (£7.47)	16.347 (0.107)	–0.113	£141.3 (£18.01)	16.352 (0.107)	–0.072
22–28	1.79–3.62	5-yearly monitoring	19.4%	£40.87 (£15.48)	16.705 (0.094)	–0.476	£402.09 (£49.34)	16.727 (0.09)	–0.100
29–32	4.07–5.78	3-yearly monitoring	4.6%	£57.61 (£27.27)	14.638 (0.103)	–3.135	£573.01 (£75.42)	14.684 (0.086)	–0.567
33–42	6.50–18.63	2-yearly monitoring	3.2%	£219.62 (£109.45)	19.354 (0.256)	–40.098	£1128.51 (£129.7)	19.592 (0.121)	–4.259
43+	20.95+	Annual monitoring	0.2%	£170.45 (£123.05)	8.909 (0.263)	–35.990	£1105.36 (£176.67)	9.124 (0.085)	1.184
		Weighted average ICER		£40.05 (£23.22)	21.066 (0.234)	–1.824	£204.93 (£38.47)	21.082 (0.229)	–0.377
							£10,198.57		

Cov, covariance; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life-year; RR, relative risk of incident melanoma vs. mean risk score of 17 in the United Kingdom [13] (calculations are based on ratio of expected incidences at respective risk scores as per equation described in Appendix 2); SE, standard error.

* Status quo costs and QALYs represent the current expected costs and QALYs accrued by members of the population with various risk scores over a period of 30 y with ad hoc detection rates (figures discounted at 3.5% per annum).

† Monitoring program shows the expected cost and QALYs accrued by those same patients under the “optimal policy” option. Thus, there is no change in cost or outcomes for those with a risk score of less than 14. The extra cost for those with a score of 15–21 represents the expected cost of the one-off examination and subsequent referral and treatment when incurred (= £141.30 – £24.99 = £116.31). The added benefit in these patients is 16.352 – 16.347 = 0.005 QALYs.

than 15 should have a one-off FBSE with a primary care practitioner. Those with a risk score greater than 22 should be enrolled into a primary care-based monitoring program with a 5-year recall, rising to annual monitoring for those with a score greater than 43.

The mean risk score in the UK population is 17 [13]. Implementing this policy in the United Kingdom would involve inviting

an estimated 61% of the adult population to at least one examination (approximately 29.9 million people), at an extra cost of £4.9 billion over 30 years, or approximately £164 million per year (0.1% of the 2016 NHS budget). This cost is the present value discounted at 3.5% per annum and includes the cost of monitoring as well as subsequent referrals and surgery. This would, however, yield approximately 15,947 additional QALYs per year:

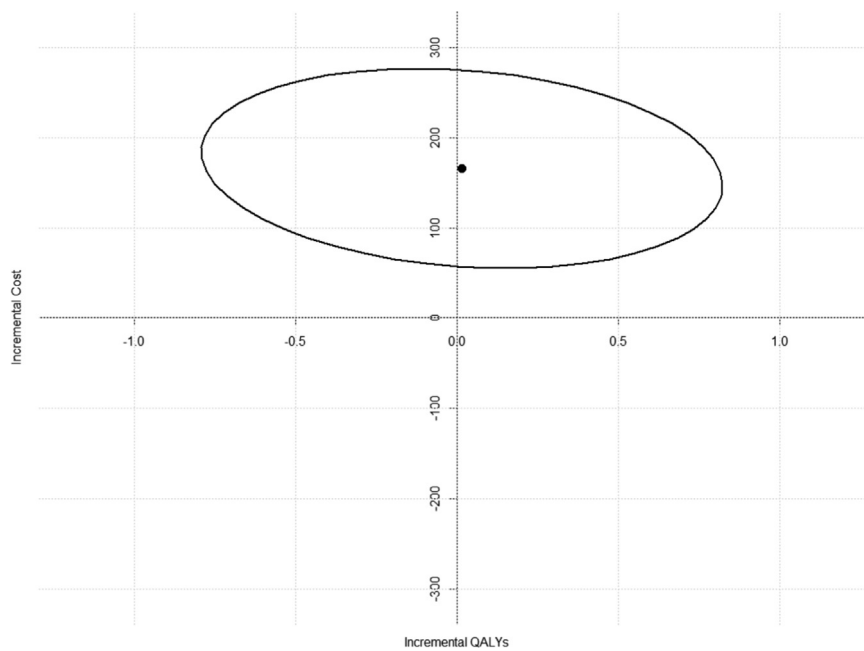


Fig. 3 – 95% credibility ellipse. The central locus is the expected incremental cost and QALYs gained per person enrolled into the “compound” strategy in the United Kingdom. Uncertainty in the point estimates is illustrated with the 95% credibility ellipse. It is almost certain that the strategy will be cost-increasing, but there is a great deal of uncertainty as to whether it will yield a health benefit. QALY, quality-adjusted life-year.

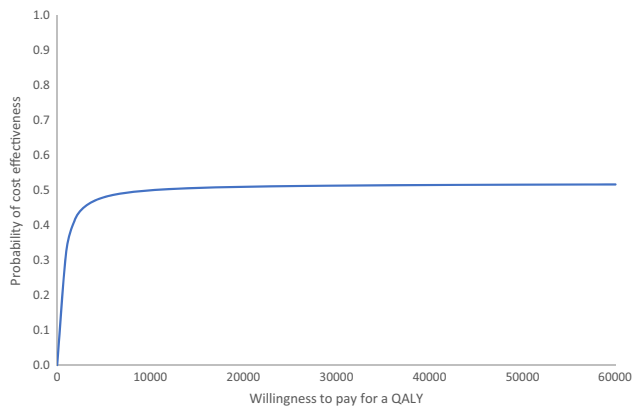


Fig. 4 – Cost-effectiveness acceptability curve of compound risk-stratified policy vs. status quo. QALY, quality-adjusted life-year.

an incremental cost per QALY of £10,199, well within what is usually considered cost-effective in the United Kingdom (£20,000–£30,000/QALY) [34]. By way of comparison, the existing breast screening program in the United Kingdom adds approximately £42.5 million to NHS expenditure, but generates about 2,040 extra QALYs (£20,800/QALY gained) [35]. A phased implementation involving only higher risk individuals would be substantially less expensive, but with consequent foregone health gain.

These recommendations are based on expected values rather than on the results of hypothesis tests. This approach is consistent with statistical decision theory [36], a key assumption of which is that decision makers are risk-neutral [37] and thus interested in maximizing expected outcomes subject to budgetary constraints. Uncertainty in decisions therefore should not be a factor in whether to adopt one particular strategy or another, but is critical to guide future research, ideally via value of information analysis [38,39]. Our analysis suggests that there is a great deal of decision uncertainty, with only a 51% to 52% probability that such a surveillance program is cost-effective (Fig. 3) [40]. A phased implementation as suggested earlier must provide the opportunity to reduce decision uncertainty, for example, through a cluster or stepped-wedge randomized controlled trial. This would inform the decision to either expand the scheme to lower risk individuals, limit to higher risk, or disinvest entirely. Additional preparatory research to establish the feasibility, sensitivity, and specificity of nurses in conducting an FBSE as part of a screening program is also critical.

Strengths and Weaknesses

As with any decision model, the robustness of the policy recommendations is contingent on the quality of the modeling and availability of source data. Our model is a patient-level simulation of a complex decision framework with a total of 476 possible compound policies (seven strategies at 68 risk scores). The model development process was methodical and rigorous, gathering the most appropriate evidence on all input parameters.

The major limitations in the model were due to lack of relevant data. Specifically, the risk of progression in undiagnosed melanoma was based on expert opinion [20]. This limitation is common to many decision analyses, particularly of screening studies; to quantify the added health benefit of screening, it is necessary to know the disease course of those who would otherwise not be identified and treated. Prospectively withholding treatment from patients with melanoma to observe this would clearly be deeply unethical; therefore, the only alternative

is to seek expert opinion. A number of techniques exist (e.g., Ref. [41]), a key feature of which is that they focus on eliciting experts' uncertainty (in terms of a range of plausible values, weighted according to strength of belief) rather than a single "best guess" for a particular parameter. We conducted an elicitation exercise in a transparent and replicable manner to address this issue [20]. Because of a lack of evidence [10], we were also unable to include potential screening-related harms in the model. These include risk of overdiagnosis [42], side effects of treatment, or psychological harms [43] and are important when considering any future screening program.

Treatment of late-stage disease in the model is based on 2010 guidelines, which do not include newer, expensive treatments of varying cost-effectiveness [44–46]. If these add substantial cost with limited health gain, it becomes even more cost-effective to detect (and thus treat) earlier in the disease process.

Further limitations include the costs and practicalities of introducing such surveillance. We assumed that all patients initially receive an FBSE by a general practitioner (GP). Not offering an FBSE may reduce the cost of the consultation, but at the risk of lower sensitivity. A community nurse conducting the examinations would be less expensive than a GP, but because the current criterion standard is examination in secondary care, costs of providing and training for both GP and nurse would need consideration [47]. Surveillance could be offered in other locations such as community clinics and via telemedicine [48].

We assumed perfect adherence and did not explicitly account for recurrent or multiple lesions. Lower adherence will reduce both costs and health gain from the program. Recurrent skin cancers were indirectly accounted for in the post-treatment survival functions [22]. Patients with multiple lesions will have increased surgical costs, but the marginal cost is likely to be small compared with the cost of nonsurgical treatment at later stages of the disease and so is unlikely to alter our conclusions substantially.

The baseline utility for patients was assumed equal to perfect health (i.e., 1). Population norms suggest a declining utility with age [49]. The model, however, applies an absolute reduction in utility (and hence QALYs) with various health states, and thus the incremental QALY gain is insensitive to this.

A final limitation was that we considered a maximum recall interval of 5 years. It may be more efficient for recall to be less frequent for medium-risk individuals, for example, decennially. This was, however, out of the scope of this analysis.

Despite these caveats, the model is based on the best evidence available to the authors at the time of writing: no decision model is perfect and can always be improved. Decisions as to what concepts to include in a decision model must be balanced against the resources available to conduct it and the need for a timely policy recommendation. Acknowledged limitations of a decision model, in terms of both structure and data inputs, provide an agenda for future research in the area.

Conclusions

Current evidence is highly uncertain but suggests that, on balance, a UK-wide program to identify patients at risk of melanoma using the Williams self-assessment tool is potentially cost-effective. Nevertheless, a program would be expensive to implement because of its scale. Additional research into the feasibility, sensitivity, and specificity of nurses in conducting FBSE as part of a program is required. Ultimately, a phased implementation targeting only the highest risk groups may be practical to implement, but must be embedded within a rigorous randomized trial to reduce decision uncertainty and hence inform further rollout.

Source of financial support: This study was supported in part by F. M. Walter's Clinician Scientist Award from the National Institute for Health Research (NIHR) (grant no. RG 68235). E. C. F. Wilson is funded by the NIHR Cambridge Biomedical Research Centre. J. Usher-Smith was funded by an NIHR Clinical Lectureship. J. Emery is funded by a National Health and Medical Research Council (NHMRC) Practitioner Fellowship. The analysis was performed using the Darwin Supercomputer of the University of Cambridge High Performance Computing Service (<http://www.hpc.cam.ac.uk/>), provided by Dell, Inc., using Strategic Research Infrastructure Funding from the Higher Education Funding Council for England and funding from the Science and Technology Facilities Council. Views expressed in this publication are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at <https://doi.org/10.1016/j.jval.2017.11.009>.

REFERENCES

- Cancer Research UK. Skin cancer statistics. Available from: <http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/skin-cancer>. [Accessed August 2, 2016].
- National Institute for Health and Care Excellence. Nivolumab for treating advanced (unresectable or metastatic) melanoma: evidence—cost-effectiveness. Available from: <https://www.nice.org.uk/guidance/ta384/chapter/3-Evidence#cost-effectiveness>. [Accessed October 2, 2017].
- UK National Screening Committee. Current UK NSC recommendations. Available from: http://legacy.screening.nhs.uk/screening-recommendations.php?sortBy=R_LastReview_date&direction=ASC. [Accessed August 2, 2016].
- Melia J, Harland C, Moss S, et al. Feasibility of targeted early detection for melanoma: a population-based screening study. *Br J Cancer* 2000;82:1605–9.
- Cancer Research UK. Skin cancer screening. Available from: <http://www.cancerresearchuk.org/about-cancer/type/skin-cancer/about/skin-cancer-screening>. [Accessed August 2, 2016].
- Wolff T, Tai E, Miller T. Screening for skin cancer: an update of the evidence for the U.S. Preventive Services Task Force. *Ann Intern Med* 2009;150:194–8.
- Gordon LG, Rowell D. Health system costs of skin cancer and cost-effectiveness of skin cancer prevention and screening: a systematic review. *Eur J Cancer Prev* 2015;24:141–9.
- Wernli KJ, Henrikson NB, Morrison CC, et al. Screening for skin cancer in adults: updated evidence report and systematic review for the US Preventive Services Task Force. *JAMA* 2016;316:436–47.
- US Preventive Services Task Force. Calonge N, Petitti DB, DeWitt TG, et al. Screening for skin cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med* 2009;150:188–93.
- US Preventive Services Task Force. Bibbins-Domingo K, Grossman DC, Curry SJ, et al. Screening for skin cancer: US Preventive Services Task Force recommendation statement. *JAMA* 2016;316:429–35.
- Usher-Smith JA, Emery J, Kassianos AP, et al. Risk prediction models for melanoma: a systematic review. *Cancer Epidemiol Biomarkers Prev* 2014;23:1450–63.
- Williams LH, Shors AR, Barlow WE, et al. Identifying persons at highest risk of melanoma using self-assessed risk factors. *J Clin Exp Dermatol Res* 2011;2:pii 1000129.
- Usher-Smith JA, Kassianos AP, Emery JD, et al. Identifying people at higher risk of melanoma across the U.K.: a primary-care-based electronic survey. *Br J Dermatol* 2017;176:939–48.
- Wilson EC, Emery JD, Kinmonth AL, et al. The cost-effectiveness of a novel SLAsopic diagnostic aid for the management of pigmented skin lesions in primary care: a decision-analytic model. *Value Health* 2013;16:356–66.
- R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing, 2016.
- Johnson SG. The NLOpt nonlinear-optimization package. Available from: <http://ab-initio.mit.edu/nlopt>. [Accessed March 3, 2018].
- Wickham H, Francois R. dplyr: a grammar of data manipulation. R package version 0.5.0. 2016. Available from <https://CRAN.R-project.org/package=dplyr>. [Accessed March 3, 2018].
- Breitbart EW, Waldmann A, Nolte S, et al. Systematic skin cancer screening in Northern Germany. *J Am Acad Dermatol* 2012;66:201–11.
- Cancer Research UK. Skin cancer incidence by age. Available from: <http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/skin-cancer/incidence#heading-One>. [Accessed August 3, 2016].
- Wilson E, Usher-Smith J, Emery J, et al. Expert elicitation of multinomial probabilities for decision analytic modelling: an application to rates of disease progression in undiagnosed and untreated melanoma, in press.
- Liu V, Mihm MC. Pathology of malignant melanoma. *Surg Clin North Am* 2003;83:31–60.
- Balch CM, Gershenwald JE, Soong SJ, et al. Final version of 2009 AJCC melanoma staging and classification. *J Clin Oncol* 2009;27:6199–206.
- Office for National Statistics. Population estimates for UK, England and Wales, Scotland and Northern Ireland, Mid 2013. Available from: <http://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/populationestimatesforukenglandandwalesscotlandandnorthernireland>. [Accessed August 9, 2016].
- Usher-Smith JA, Kassianos AP, Emery JD, et al. Identifying people at higher risk of melanoma across the UK: a primary care-based electronic survey. *Br J Dermatol* 2017;176(4):939–48.
- Office for National Statistics. National life tables, United Kingdom: 2012–2014. Available from: <http://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesunitedkingdomreferencetables>. [Accessed August 9, 2016].
- Walter FM, Morris HC, Humphrys E, et al. Effect of adding a diagnostic aid to best practice to manage suspicious pigmented lesions in primary care: randomised controlled trial. *BMJ* 2012;345:e4110.
- Cox NH. Evaluation of the U.K. 2-week referral rule for skin cancer. *Br J Dermatol* 2004;150:291–8.
- Office for National Statistics. Population estimates for UK, England and Wales, Scotland and Northern Ireland, Mid-2011 and Mid-2013. Available from: <http://www.ons.gov.uk/ons/rel/pop-estimate/population-estimates-for-uk-england-and-wales-scotland-and-northern-ireland/mid-2011-and-mid-2012/index.html>. [Accessed May 2, 2014].
- Curtis L, Burns A. Unit Costs of Health and Social Care 2015. Canterbury, Kent: University of Kent, 2015.
- Department of Health. National schedule of reference costs, 2014–2015. Available from: <https://www.gov.uk/government/publications/nhs-reference-costs-2014-to-2015>. [Accessed August 9, 2016].
- Marsden JR, Newton-Bishop JA, Burrows L, et al. Revised U.K. guidelines for the management of cutaneous melanoma 2010. *Br J Dermatol* 2010;163:238–56.
- Cornish D, Holterhues C, van de Poll-Franse LV, et al. A systematic review of health-related quality of life in cutaneous melanoma. *Ann Oncol* 2009;20(Suppl. 6):vi51–8.
- Bendeck S, Hadley J, Bonaccorsi P, et al. Quality of life impact by melanoma as measured by utilities. Presented at: Annual Meeting of the Society for Medical Decision Making, Atlanta, GA, 2004:E49.
- National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013. Available from: <https://www.nice.org.uk/process/pmg9/chapter/foreword>. [Accessed March 1, 2018].
- Pharoah PD, Sewell B, Fitzsimmons D, et al. Cost effectiveness of the NHS breast screening programme: life table model. *BMJ* 2013;346:f2618.
- Pratt J, Raiffa H, Schlaifer R. Introduction to Statistical Decision Theory. Cambridge, MA: Massachusetts Institute of Technology, 1995.
- Arrow KJ, Lind RC. Uncertainty and evaluation of public investment decisions. *Am Econ Rev* 1970;60:364–78.
- Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. *J Health Econ* 1999;18:341–64.
- Wilson EC. A practical guide to value of information analysis. *Pharmacoeconomics* 2015;33:105–21.
- Fenwick E, Claxton K, Sculpher M. Representing uncertainty: the role of cost-effectiveness acceptability curves. *Health Econ* 2001;10:779–87.
- Oakley JE, O'Hagan A. SHELF: The Sheffield Elicitation Framework (Version 2.0). Sheffield, UK: School of Mathematics and Statistics, University of Sheffield, 2010.
- Welch HG, Woloshin S, Schwartz LM. Skin biopsy rates and incidence of melanoma: population based ecological study. *BMJ* 2005;331:481.
- Gambichler T, Senger E, Rapp S, et al. Deep shave excision of macular melanocytic nevi with the razor blade biopsy technique. *Dermatol Surg* 2000;26:662–6.

- [44] National Institute for Health and Care Excellence. Summary of Appraisal Committee's key conclusions: TA319 ipilimumab for previously untreated advanced (unresectable or metastatic) melanoma. Available from: <https://www.nice.org.uk/guidance/TA319/chapter/4-Consideration-of-the-evidence>. [Accessed November 12, 2016].
- [45] National Institute for Health and Care Excellence. Summary of Appraisal Committee's key conclusions: TA385 nivolumab for treating advanced (unresectable or metastatic) melanoma. Available from: <https://www.nice.org.uk/guidance/ta384/chapter/4-Committee-discussion>. [Accessed November 12, 2016].
- [46] National Institute for Health and Care Excellence. Summary of Appraisal Committee's key conclusions: TA321 dabrafenib for treating unresectable or metastatic BRAF V600 mutation-positive melanoma. Available from: <https://www.nice.org.uk/guidance/ta321/chapter/4-Consideration-of-the-evidence>. [Accessed November 12, 2016].
- [47] Grange F, Woronoff AS, Bera R, et al. Efficacy of a general practitioner training campaign for early detection of melanoma in France. *Br J Dermatol* 2014;170:123–9.
- [48] Landow SM, Oh DH, Weinstock MA. Teledermatology within the Veterans Health Administration, 2002–2014. *Telemed J E Health* 2015;21:769–73.
- [49] Kind P, Hardman G, Macran SUK. Population Norms for EQ-5D. York, UK: Centre for Health Economics, University of York, 1999.