Optimal surveillance strategies for patients with stage 1 cutaneous melanoma post primary tumour excision: three systematic reviews and an economic model

Luke Vale,[1] Patience Kunonga,[1] Diarmuid Coughlan,[1] Vasileios Kontogiannis,[1] Margaret Astin,[1] Fiona Beyer,[1] Catherine Richmond,[1] Dor Wilson,[1] Dalvir Bajwa,[2] Mehdi Javanbakht,[1] Andrew Bryant,[1] Wanwuri Akor,[3] Dawn Craig,[1] Penny Lovat,[4] Marie Labus,[5] Batoul Nasr,[6] Timothy Cunliffe,[7] Helena Hinde,[7] Mohamed Shawgi,[8] Daniel Saleh,[9,10] Pam Royle,[11] Paul Steward,[12] Rachel Lucas[12] and Robert Ellis[4,13*]

[1]Institute of Health & Society, Newcastle University, Newcastle upon Tyne, UK
[2]Institute of Cellular Medicine, Newcastle University, Newcastle upon Tyne, UK
[3]Northumbria Healthcare NHS Foundation Trust, North Shields, UK
[4]Institute of Translation and Clinical Studies, Newcastle University, Newcastle upon Tyne, UK
[5]Business Development and Enterprise, Newcastle University, Newcastle upon Tyne, UK
[6]Dermatological Sciences, Institute of Cellular Medicine, Newcastle University, Newcastle upon Tyne, UK
[7]Dermatology Department, James Cook University Hospital, Middlesbrough, UK
[8]Radiology Department, James Cook University Hospital, Middlesbrough, UK
[9]Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK
[10]Princess Alexandra Hospital Southside Clinical Unit, Faculty of Medicine, University of Queensland, Brisbane, QLD, Australia
[11]Patient representative, ITV Tyne Tees, Gateshead, UK
[12]Patient representative, Dermatology Department, James Cook University Hospital, Middlesbrough, UK
[13]South Tees Hospitals NHS Foundation Trust, Middlesbrough, UK

*Corresponding author robellis@nhs.net

Declared competing interests of authors: Luke Vale was a member of the National Institute for Health Research (NIHR) Clinical Evaluation and Trial Panel from 2015 to 2018. His partner is the Chief Executive Officer (employed by Newcastle University) of a Newcastle University spin-out company called AMLo Biosciences Ltd (Newcastle upon Tyne, UK). She is also a shareholder of the same, and holds patents for biomarkers in this area. AMLo Biosciences Ltd is developing a prognostic biomarker for melanoma. Dawn Craig is a member of the NIHR Health Services and Delivery Research programme's Prioritisation Committee (Researcher-led). Robert Ellis reports personal fees from AbbVie Inc. (Lake Bluff, IL, USA) and AMLo Biosciences Ltd outside the submitted work. Penny Lovat and Marie Labus report personal fees from AMLo Biosciences Ltd, outside the submitted work. In addition, Penny Lovat has a patent family arising from PCT/GB2015/053347 (biomarkers for disease progression in melanoma; assignee: AMLo Biosciences Ltd; inventors: Penny Lovat, Robert Ellis and Marie Labus; issued), patent 1818168.9 (monoclonal antibodies against AMBRA1; assignee: AMLo Biosciences Ltd; inventors: Penny Lovat, Robert Ellis, Ashleigh McConnell and Marie Labus; issued) and patent 118622.1 (monoclonal antibodies against loricrin; assignee: AMLo Biosciences Ltd; inventors: Penny Lovat, Robert Ellis, Ashleigh McConnell and Marie Labus; pending).

Published November 2021
DOI: 10.3310/hta25640

Scientific summary

Surveillance strategies for stage 1 cutaneous melanoma

Health Technology Assessment 2021; Vol. 25: No. 64
DOI: 10.3310/hta25640

NIHR Journals Library www.journalslibrary.nihr.ac.uk

Scientific summary

Background

Cutaneous melanoma is a cancer that develops from pigment-producing cells (melanocytes) in the skin, and is one of the deadliest skin cancers. It is aggressive, rapidly disseminates and, until recently, had a median overall survival of between 6 and 10 months once metastasis had occurred. The recent introduction of targeted immunotherapies has improved outcomes, with median overall survival now reaching at least 2 years. Cutaneous melanoma is the fifth most common cancer in the UK, with 17,000 patients diagnosed annually. It is also the UK's leading cause of cancer-related death among people aged 20–35 years.

After removal of the primary tumour, the majority of melanomas are cured. However, up to 30% of all primary melanomas progress to metastatic disease, with extremely poor 5-year survival rates of only 14%. Consequently, there are 2500 melanoma-associated deaths in the UK annually. The total annual cost due to skin cancer to the NHS was £106M–112M in 2008 and is expected to rise to > £180M by 2020.

Primary melanomas are staged according to the American Joint Committee on Cancer staging criteria. These include Breslow depth (the distance into the skin of the tumour invasion) and the presence of ulceration (loss of the epidermis overlying the tumour) to allow disease-risk stratification. In 2017, the eighth edition of American Joint Committee on Cancer staging, but when this study was conducted, most data used the seventh or earlier editions of staging. In the seventh edition of the American Joint Committee on Cancer, stage I tumours are identified as tumours up to 2 mm thick, with no ulceration, or < 1 mm thick, if ulceration is present. American Joint Committee on Cancer stage I disease represents the lowest mortality risk, compared with other stages of disease: up to 14% over 10 years.

Although surgical treatment of primary melanoma is effective, the pace of development is rapid, with the introduction of additional early investigatory techniques (e.g. sentinel lymph node biopsy and various radiological modalities) and advances with the treatment of metastatic disease. A structured, evidence-based model of patient follow-up after initial diagnosis is lacking. Current guidelines for surveillance vary across the world, with most based on anecdotal evidence and expert opinion. The recommendations make the assumption that earlier disease detection results in improved outcome, but often do not consider all elements used in the diagnosis and management of the condition, as well as the potential physical, psychological and financial costs of these surveillance regimens.

With low rates of metastasis, and early physiological stage of development, targeting American Joint Committee on Cancer stage I melanomas for appropriate follow-up strategies could improve, or at least maintain, outcomes at lower costs. Limited evidence suggests that low-risk patients may not need intensive clinician follow-up, as recommended. Conversely, a more appropriately structured follow-up regime for higher-risk patients may allow earlier detection of metastatic disease with associated benefits from earlier treatment.

With the rapid increase in melanoma rates, there is a need to develop a robust, evidence-based model of follow-up care for American Joint Committee on Cancer stage I patients: the majority of people affected by melanoma. The increase in diagnostic accuracy, development of potential prognostic biomarkers, new radiological modalities and the introduction of personalised systemic treatments could transform melanoma care. However, without a robust, evidence-based framework for implementation of such interventions, the potential health and economic benefits for the NHS will not be achieved.

Objectives

The aim of this research was to evaluate the effectiveness and cost-effectiveness of different surveillance strategies for patients with American Joint Committee on Cancer stage I melanoma after surgical excision of the primary cutaneous tumour. The objectives were to:

1. systematically review different strategies for surveillance and follow-up after surgical excision of the primary cutaneous tumour, including their effectiveness and cost-effectiveness
2. systematically review the prognostic performance of risk models used to determine the prognosis and risk stratification of patients with American Joint Committee on Cancer stage I melanoma after surgical excision of the primary cutaneous tumour
3. systematically review the diagnostic performance of tests used in surveillance and follow-up strategies in detecting new primaries, recurrence and metastatic diseases in patients with American Joint Committee on Cancer stage I melanoma after surgical excision of the primary cutaneous tumour
4. develop a decision-analytic model to estimate the effectiveness and cost-effectiveness of the surveillance and follow-up strategies after surgical excision of the primary cutaneous tumour
5. undertake value of information analysis to assess the need for further primary research.

The results of the systematic reviews conducted to meet the first three objectives were used to inform the design and conduct of an economic evaluation based on a decision model, which addressed the fourth and fifth objectives.

Surveillance review

Methods

A systematic review of comparative studies was conducted to identify various surveillance and follow-up strategies after surgical excision of American Joint Committee on Cancer stage I primary cutaneous melanomas in adults, and to assess their relative effectiveness. The review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines. Ten bibliographic databases, grey literature and guidelines from January 2011 to July 2019 were searched. To identify studies published prior to 2011, we assessed all references of an earlier review on melanoma surveillance by Cromwell *et al.* (Cromwell KD, Ross MI, Xing Y, Gershenwald JE, Royal RE, Lucci A, *et al.* Variability in melanoma post-treatment surveillance practices by country and physician specialty: a systematic review. *Melanoma Res* 2012;**22**:376–85), which searched up to that point. Furthermore, references from the clinical guidelines identified in an earlier component of this project were assessed. Included studies had to compare surveillance strategies (relevant strategies compared included a no active surveillance option). Outcomes included detection of new primary tumours, recurrence, metastases or overall survival.

Results

Searches identified 6205 records. One randomised controlled trial from the USA met the inclusion criteria. This trial evaluated the effect of surveillance using structured skin self-examination. New primaries, recurrences or metastases were detected in 49 out of 258 (19%) patients with stage IA or IB melanoma: 36 out of 203 (18%) in the intervention group and 13 out of 55 (24%) in the control group. The overall risk of bias for the trial was identified as being of some concern. Overall certainty of the evidence was low and future trials would be very likely to influence results.

Prediction model review

Methods

This review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses and the CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling

Studies (CHARMS) guidelines, and assessed the prognostic accuracy of risk prediction models to predict recurrence, new primary tumours and metastases. Searches of 10 databases were conducted, searching from 2000 (when sentinel lymph node biopsy was introduced in melanoma) to July 2019. Model performance measures included discrimination (ability to differentiate between high and low risk), calibration (agreement between observed and predicted risk) and overall performance (combination of the discrimination and calibration measures), assessed by the Brier score (a statistical measure of the accuracy of the measure; a higher score means higher inaccuracy) and R^2 statistic (a statistic describing the percentage of the variance to measure overall model performance).

Results

Searches identified 20,878 records and 11 different risk prediction models. The number of predictors per model ranged from 3 to 11. The most common were age, tumour site, tumour thickness, sex and ulceration. Discrimination was reported in six studies and the area under the operating curve (whereby 0.5 is fail and 1 is perfect) ranged from 0.59 to 0.88. Calibration measures were reported in three studies. One study reported a calibration slope of 0.88 ($p = 0.5$), and another reported concordance correlation coefficients of 0.9 and 0.93 for 5- and 10-year survival rates, both demonstrating high accuracy of the models. Two studies measured the overall performance of the model. One study assessed the Brier score of a new model and showed a slightly better (i.e. lower) Brier score than the American Joint Committee on Cancer scheme. The other study assessed the R^2 statistic and reported it as 0.47 (95% confidence interval 0.45 to 0.49), indicating that the model explains an estimated 47% of the variation. All studies were retrospective, and so were rated as having a high risk of bias; eight studies conducted internal validation using data from their development set.

Diagnostic performance review

Methods

This systematic review explored the diagnostic test accuracy of fine-needle biopsy and ultrasonography to detect new primaries, recurrence and locoregional metastases during follow-up of stage I melanoma. Searches of electronic databases were conducted from inception to July 2019. Data were extracted on study/participant characteristics and index test accuracy statistics. Risk of bias was independently assessed using Quality Assessment of Diagnostic Accuracy Studies-2 for each included study.

A bivariate random-effects meta-analysis model was planned. This approach would have enabled the calculation of summary estimates of sensitivity and specificity across different studies. Owing to paucity of data, a narrative approach was used and estimates of sensitivity and specificity for each study estimated.

Results

Database searches retrieved 2250 records. Two studies assessing different index tests relevant at different stages of diagnosis met the inclusion criteria. One Australian study (Doubrovsky A, Scolyer RA, Murali R, McKenzie PR, Watson GF, Lee CS, *et al.* Diagnostic accuracy of fine needle biopsy for metastatic melanoma and its implications for patient management. *Ann Surg Oncol* 2008;**15**:323–32.) assessed the accuracy of fine-needle biopsy (the study was rated as having a high risk of bias). Data were reported for stage I disease by the number of fine-needle biopsies performed ($n = 323$) in those with stage I melanoma. A German study (Krüger U, Kretschmer L, Thoms KM, Padeken M, Bertsch PH, Schön MP, Zutt M. Lymph node ultrasound during melanoma follow-up significantly improves metastasis detection compared with clinical examination alone: a study on 433 patients. *Melanoma Res* 2011;**21**:457–63.) from assessed ultrasound (the study was rated as having a high risk of bias) and included 669 investigations among individuals with stage I melanoma.

For the study assessing the diagnostic performance of fine-needle biopsy, the sensitivity was 0.93 (95% confidence interval 0.88 to 0.97) and specificity was 0.98 (95% confidence interval 0.95 to 1.00).

Results in the ultrasonography study were reported as the number of paired investigations of both clinical examination and ultrasonography conducted on the same day, with an average of three paired investigations per patient. These data were converted so that the unit of analysis was the participants. Sensitivity was reported as 1.00 (95% confidence interval 0.03 to 1.00) and specificity was 0.99 (95% confidence interval 0.96 to 1.00).

Sensitivity analyses were conducted for both study analyses; results did not change.

Economic evaluation

Methods

A review of cost-effectiveness studies identified 15 possibly relevant studies, but none directly addressed the study question. Therefore, an economic model was developed to assess the cost-effectiveness of alternative surveillance strategies and to estimate the value of information. The model took a lifetime horizon and an NHS and Personal Social Services perspective. A Markov microsimulation model, with monthly cycles, was developed in TreeAge 2019 R1.0 (TreeAge Software, Inc., Williamstown, MA, USA). Quality-adjusted life-years and costs were estimated and discounted at an annual rate of 3.5%. All costs are reported in 2018 Great British pounds.

Based on consultation with clinical team members and current National Institute for Health and Care Excellence surveillance guidelines, a total of 75 alternative NHS strategies for stage IA and 87 strategies for stage IB were identified for initial modelling. The main probabilities used in the model were the probabilities of recurrence, the probabilities of treatment success, patient self-diagnosis, 'false alarms' resulting in emergency visits, transition probabilities between the different stages of melanoma and mortality rates taken from various international data sources. EuroQol-5 Dimensions utility estimates were derived from a systematic review and meta-analysis of studies from Australia, Europe and North America. Both deterministic and probabilistic sensitivity analyses were used to explore uncertainty. The age and sex information from the individual patient data obtained from the University Hospital of North Durham (NREC 19/NE/004) were used to estimate the mortality rate of the simulated cohort.

Results

Initial modelling showed that strategies involving cancer nurse specialists providing clinical examinations were unlikely to be cost-effective, primarily because of the comparatively poorer diagnostic accuracy assumed.

From initial modelling, 20 surveillance strategies each for stages IA and IB were evaluated in more depth. For both stages, the evaluated strategies were similar in terms of quality-adjusted life-years, reflecting the relatively low rates of recurrence expected. The strategy of follow-up once for 1 year by a dermatologist was the least costly and most likely to be considered cost-effective if society were willing to pay £20,000 per quality-adjusted life-year (AJCC stage IA, 13%; AJCC stage IB, 13%). For stage IA, the strategy recommended by the National Institute for Health and Care Excellence performed similarly (12%). For stage IB, the strategy recommended by the National Institute for Health and Care Excellence performed poorly (4%). Although these probabilities are low, a large number of different surveillance strategies were compared. A sensitivity analysis showed that there may be value in improving the diagnostic performance of cancer nurse specialists and in the use of low-cost risk prediction tools for prognosis.

The highest value for research came from removing all uncertainty around probabilities of transitioning between the different stages of melanoma (stage IA, £380M; stage IB, £457M) and diagnostic accuracy (stage IA, £276M; stage IB, £193M). The value of removing uncertainty in utilities was lower, but still substantial.

Conclusions

Few data were available specific to surveillance of people after treatment for melanoma. Furthermore, few data were available for key components of a surveillance strategy that could be used to model alternative strategies. What data were available mainly related to studies using cancer staging classifications predating the publication of the American Joint Committee on Cancer eighth edition. Therefore, results are imprecise and considerable uncertainty exists. There is insufficient evidence to recommend any changes to the current surveillance guidelines produced by the National Institute for Health and Care Excellence. There are plausible surveillance strategies that may perform better than current recommendations for surveillance. However, for those treated for stage IA disease, the National Institute for Health and Care Excellence's strategy still performs comparatively well. For stage IB disease, the strategy recommended by the National Institute for Health and Care Excellence of follow-up every 3 months for 3 years then every 6 months for a further 2 years performs poorly, compared with other strategies considered, but there is insufficient evidence to support any changes.

Surveillance strategies whereby the clinical follow-up is conducted by a cancer nurse specialist may ease pressure on dermatologists and plastic surgeons. However, methods to enhance cancer nurse specialists' diagnostic performance may be needed, as the current limited evidence base suggests that their ability to correctly identify who does or does not have a recurrence is not as good as that of dermatologists. Likewise, encouraging and supporting patients in making accurate self-diagnosis of recurrence in stage I disease may reduce the need for any active surveillance strategy for those initially treated for stage I disease.

Recommendations for research

It is tempting to recommend that a randomised controlled trial should be conducted to compare surveillance strategies. However, a surveillance strategy is a complex intervention and research should first establish what sensible comparators there should be against current practice. What an appropriate comparator would be may vary between stage IA and stage IB disease, and establishing this requires improved evidence on how disease in patients with stage I melanoma develops over time. The economic modelling shows that both the incidence of recurrent and metastatic disease over time, and how it progresses are important. Further research would also be valuable on how well recurrent and metastatic disease is diagnosed, improving the diagnostic performance of practitioner groups like cancer nurse specialists, developing and evaluating low-cost tools that can better stratify patients into low or high risk of future recurrence and metastasis, and identifying the patient preferences for alternative methods of surveillance and on the impacts on health-related quality of life in patients with melanoma.

Study registration

This study is registered as PROSPERO CRD42018086784.

Funding

This project was funded by the National Institute for Health Research Health Technology Assessment programme and will be published in full in *Health Technology Assessment*; Vol 25, No. 64. See the NIHR Journals Library website for further project information.

Health Technology Assessment

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 4.014

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, the Cochrane Library and Clarivate Analytics Science Citation Index.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: journals.library@nihr.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

Health Technology Assessment (HTA) research is undertaken where some evidence already exists to show that a technology can be effective and this needs to be compared to the current standard intervention to see which works best. Research can evaluate any intervention used in the treatment, prevention or diagnosis of disease, provided the study outcomes lead to findings that have the potential to be of direct benefit to NHS patients. Technologies in this context mean any method used to promote health; prevent and treat disease; and improve rehabilitation or long-term care. They are not confined to new drugs and include any intervention used in the treatment, prevention or diagnosis of disease.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

This report

The research reported in this issue of the journal was funded by the HTA programme as project number 16/166/05. The contractual start date was in March 2018. The draft report began editorial review in October 2019 and was accepted for publication in July 2020. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health and Social Care.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).