Enhancing survival prognostication in patients with choroidal melanoma by integrating pathologic, clinical and genetic predictors of metastasis

Antonio Eleuteri
Department of Medical Physics and Clinical Engineering,
Royal Liverpool University Hospital,
Duncan Building,
Daulby St., Liverpool L7 8XP, UK
E-mail: antonio.eleuteri@gmail.com

Bertil Damato
Ocular Oncology Service,
Royal Liverpool University Hospital,
Prescot St., Liverpool, L7 8XP, UK
E-mail: bertil.damato@btinternet.com

Sarah E. Coupland
Department of Molecular and Clinical Cancer Medicine,
University of Liverpool,
Duncan Building, Daulby Street,
Liverpool, L69 3GA, UK
E-mail: s.e.coupland@liverpool.ac.uk

Azzam F.G. Taktak*
Department of Medical Physics and Clinical Engineering,
Royal Liverpool University Hospital,
1st Floor, Duncan Building,
Daulby St., Liverpool L7 8XP, UK
Fax: 0151 7065803
E-mail: afgt@liverpool.ac.uk
*Corresponding author

Abstract: Survival in choroidal melanoma was modelled using accelerated failure time models. We combined pathological, clinical and genetic data, using imputation techniques. Performance was assessed using n-fold cross-validation. Using data from 3653 patients, we generated two models; the first using clinical data only and the second using clinical and laboratory data. The c-index of discrimination was 0.75 for the clinical model and 0.79 for the laboratory model. Calibration showed good correlation between predicted and observed mortality (p-value: 0.699 for clinical model and 0.801 for laboratory model). We conclude that our model provides reasonably reliable prognosis relevant to individual patients.
Enhancing survival prognostication in patients with choroidal melanoma

Keywords: uveal melanoma; mortality; chromosome aberrations; prognostication; histology; mathematical-modelling.


Biographical notes: Antonio Eleuteri received his Laurea Degree from the University of Salerno, Italy, in 1998 and a PhD, from the University of Napoli, Italy, in 2004. Since 2006, he has been with the Royal Liverpool University Hospital, UK, where he is a Research Fellow. His research interests focus on the theory and application of machine learning and mathematical modelling. So far he has around 20 publications on peer-reviewed journals and conference proceedings.

Bertil Damato is consultant ophthalmologist at the Royal Liverpool University Hospital and Clinical Lead of the Ocular Oncology Service, which receives around 700 new patients each year. He is Honorary Professor in the Department of Molecular and Clinical Cancer Medicine at Liverpool University. His research interests include: development of treatments for ocular tumours; uveal melanoma cytogenetics; prognostication; and quality of life. He has published over 160 papers in peer-reviewed journals in addition to several chapters and a textbook on ocular oncology. He is Vice-President of the International Society of Ocular Oncology and committee member of the European Ophthalmic Oncology Group.

Sarah E. Coupland is Professor/Honorary Consultant in Pathology at RLBUHT. She obtained her MBBS at the University of NSW, Sydney (1988) and a PhD (Med) at the University of Sydney (1994). She undertook her Pathology training at the University of Benjamin Franklin, Berlin (completed 2002) and defended her Associate Professor thesis in 2005. She commenced her work in Liverpool in December 2005, and has since then established a successful ocular oncology research team. She has published 127 manuscripts and 14 textbook Chapters. She is Secretary of the European Ocular Oncology Group (OOG) and Chair of the NCRI Lymphoma Biological Studies Subgroup.

Azzam F.G. Taktak is a Consultant Clinical Scientist at the Royal Liverpool University Hospital and an Honorary Lecturer at the University of Liverpool, UK. He has a BEng Degree with Honours in Electronic and Electrical Engineering, a PhD in Biomedical Engineering and a Diploma in Statistics. He is a Fellow of the Institute of Physics and Engineering in Medicine (FIPEM). His main research interests are mathematical and statistical modelling in medicine. To date he has published more than 60 peer-reviewed papers and edited a book on outcome prediction in cancer.

1 Introduction

About 50% of patients with uveal melanoma develop metastatic disease, which almost always involves the liver (Kujala et al., 2003). It is not known whether treatment
of the ocular tumour prevents metastasis and if so in whom (Damato, 2010). Most patients with metastatic disease die within a year of the onset of symptoms, and treatment only rarely seems to prolong life significantly (Augsburger et al., 2009). Studies on systemic adjuvant therapy have not shown significant survival benefit (Voelter et al., 2008).

As with other cancers, accurate prognostication is required, so that patients with a good survival probability can be reassured and prevented from unnecessary screening for metastasis. Such intensive care would be reserved for high-risk patients, thereby conserving resources. More reliable prognostication would also enhance prospects for randomised studies evaluating whether systemic adjuvant therapy prevents or delays metastatic disease. Accurate survival predictions would also make it possible to use tissue samples to identify novel prognostic markers without waiting several years for each patient’s survival to be determined.

Many factors are known to be predictive of metastatic disease from choroidal melanoma, the most important being: largest basal tumour diameter; ciliary body involvement; extracocular spread; epithelioid melanoma cytomorphology; extravascular matrix patterns such as closed loops (henceforth referred to as closed loops); high mitotic count; and a variety of genetic abnormalities, particularly chromosome 3 loss, chromosome 8q gain and class 2 gene expression profile (Coupland et al., 2008; Angi et al., 2011; Rummelt et al., 1995; Onken et al., 2010; Prescher et al., 1996; McLean et al., 1983; Damato et al., 2010). In previous studies, when choroidal melanomas showed lethal genetic abnormalities, the patient’s survival time correlated with the absence of adverse clinical and histologic risk factors (Damato et al., 2008, 2010).

The 7th edition of the Tumour, Node Metastasis (TNM) staging system for cancer categorises uveal melanomas according to: involvement of choroid, ciliary body and iris; basal tumour diameter; tumour height and extracocular extension (Finger, 2009). A list of pathological risk factors recommended for collection is included, together with a system for defining histologic grade according to melanoma cell type. There are no instructions or guidelines for multivariate analysis of clinical, histologic and genetic predictive factors. As a result, prognostication is imprecise and therefore only applicable to large groups of patients, not individuals.

In the past, models have been developed using neural networks for integrating pathological, clinical and genetic data to enhance prognostic accuracy so that estimates of survival probability are relevant to individual patients (Damato et al., 2008; Taktak et al., 2004; Eleuteri et al., 2007). Validation studies showed these models to perform adequately in patients treated by local resection or enucleation when a full dataset of clinical and laboratory information was available (Taktak et al., 2007). However, prediction was unreliable when only biopsy specimens were analysed, as in patients treated with radiotherapy or phototherapy. This was because biopsy precluded mitotic counts and assessment of extravascular matrix patterns and because the neural networks were unable to compensate adequately for the missing data. Such missing information has become more common in recent years as prognostic biopsy of irradiated melanomas has become routine, at least in our hospital.

The aim of this study was to create a prognostic model that combined pathological, clinical and genetic data, using imputation techniques to compensate for missing information, also taking competing risks into account.
2 Methodology

2.1 Subjects

Patients were selected from the database of the Liverpool Ocular Oncology Centre for the time period 1984–2009 if:

1. diagnosed with uveal melanoma, clinically or histopathologically
2. primarily treated by the first author (BD) or an associate at the Tennent Institute of Ophthalmology, Glasgow, before January 1993 or at the Royal Liverpool University Hospital between January 1993 and July 2006
3. resident in mainland Britain.

Patients were excluded because of:

1. bilateral melanoma
2. missing data regarding basal tumour dimension or anterior tumour extension
3. iris or ciliary body tumour not involving choroid; or
4. residence overseas, including Northern Ireland.

The reason for including patients only from mainland Britain is that these patients were enrolled National Health Service (NHS) Cancer Registry, which automatically informed us of date and cause of death resulting in complete follow-up on these patients.

2.2 Data

Baseline and follow-up assessments were performed as described previously (Damato et al., 2008). A summary of the input variables is provided in Table 1. Clinical data included: ciliary body involvement, categorised according to whether or not the tumour extended anterior to the ora serrata (‘ora’); extraocular spread (‘EO’) and largest basal tumour dimension measured with B-scan ultrasonography (‘LUD’).

Presence of epithelioid melanoma cells (‘Epi’) and mitotic count (‘Mitosis’) were determined by light microscopy using hematoxylin and eosin staining. Closed loops (‘Loops’) were identified using sections stained with periodic-acid Schiff stain without counterstaining and viewed under a green filter, as previously described (Damato et al., 2007). Extraocular tumour spread was detected by slit-lamp examination, ultrasonography, or pathological examination (Coupland et al., 2008).

Between 1999 and 2007, chromosome 3 loss (i.e., monosomy 3, ‘Mon3’) was identified by Fluorescent In-Situ Hybridization (FISH), which was performed as a clinical service to all patients undergoing local resection or enucleation (Damato et al., 2007; Pardue and Gall, 1975). Since 2007, genetic typing was based on Multiplex Ligation-dependent Probe Amplification (MLPA), this being offered also to patients treated with radiotherapy, whose tumour was biopsied trans-sclerally or trans-retinally (Damato et al., 2010; Van Dijk et al., 2005).
Survival probabilities of the general population were estimated using the UK Government Actuary’s Department (GAD) interim life table figures for 2004 (http://www.gad.gov.uk).

Table 1  
Input variables for the model including their description and coding scheme

<table>
<thead>
<tr>
<th>Variable name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Age treatment in years</td>
</tr>
</tbody>
</table>
| Sex           | 0: Female  
               1: Male |
| LUD           | Largest tumour diameter from ultrasound in mm |
| Ora           | Anterior margin:  
               0: Post-ora  
               1: Pre-ora |
| EO            | Extra-ocular extension  
               0: No  
               1: Yes |
| Epi           | Tumour cell type  
               0: Spindle  
               1: Epithelioid/mixed |
| Loops         | Presence of extra-vascular closed-loop matrices  
               0: No  
               1: Yes |
| Mitosis       | Mitotic count per 40 high power fields  
               0: 0–1  
               1: 2–3  
               2: 4–7  
               3: > 7 |
| Mon3          | Chromosome 3 loss  
               0: No  
               1: Yes |

This study was conducted in accordance with the Declaration of Helsinki and Good Clinical Practice Guidelines. Consent for the use of tissues and data for research was obtained from all patients. Institutional Review Board/Ethics committee approval was obtained.

2.3  Data analysis

We first attempted to model the data using the Cox Proportional Hazards (PH) model. However, evaluation of the Grambsch-Therneau residuals showed that the null hypothesis of the proportionality of hazards was rejected for both the clinical \((p < 0.001)\) and laboratory \((p < 0.001)\) models (Cox and Snell, 1968). The variable that violated the PH assumptions most strongly was the basal tumour diameter.
We therefore explored an alternative approach, by formulating the problem in terms of an Accelerated Failure Time model. This predictive model was created by specifying an analytical expression that described the relationships between the variables and survival time. The formula developed for this model, linking the survival time \( T \) to the vector of prognostic factors \( X \) was the following:

\[
\log(T) = f(X) + \sigma \epsilon,
\]

where \( \epsilon \) is a logistic random variable with scale \( \sigma \). The function \( f \) was a linear combination of prognostic factors. Because age and basal tumour diameter were continuous variables, complex nonlinear interactions with time were anticipated; therefore, these continuous variables were modelled by 5 knots-restricted cubic splines (Harrell, 2001):

\[
g(x) = b_0 + b_1 x + \sum_{k=2}^{5} b_{k,1}(x-t_k)^3,
\]

where the operator \((\cdot)\) was defined as:

\[
(u) = \begin{cases} 
  u, & u > 0 \\ 
  0, & u \leq 0
\end{cases}
\]

The knots \( t_k \) were chosen based on the quantiles of the variable (Harrell, 2001): 0.05, 0.275, 0.5, 0.725, 0.95.

Mitotic count was treated differently, being expanded nonlinearly by a combination of linear plus dummy factors (Harrell, 2001) (i.e., creating non-uniform categories). This was necessary because of the sparseness of mitotic counts. Survival times were modelled assuming a log-logistic distribution of the events (Figure 1).

Two versions of the model were created: one with clinical factors only and the other including pathological and genetic data (i.e., ‘laboratory’ data). Values for missing data were estimated using the Alternating Conditional Expectations algorithm, which essentially estimated each of the missing variables as a function of the other variables (Harrell, 2001). For example, if mitotic count and extravascular matrix patterns were not known, these were estimated by modelling their relationships with all the other baseline variables. This process approximated the joint distribution of the baseline variables. If only the clinical features and melanoma cytomorphology were known, with missing data on mitotic count, closed loops, and chromosome 3 status, then survival was predicted with and without chromosome 3 loss respectively, estimating the likely mitotic count and closed loops in each case. A weighted-average scheme was then applied based on the probability of chromosome 3 loss, which was estimated according to all available information. The confidence intervals were adjusted to take account of any error introduced. If melanoma cytomorphology and chromosome 3 status were both unknown, then the prognostication defaulted to the clinical model.

Bootstrap re-sampling was then used to estimate the model parameters. In other words, the entire dataset was repeatedly and randomly split into training and test datasets. The model was then fitted to the training data and its performance tested using the test data. This was repeated 200 times so that the statistical variability of the performance of the model could be assessed, using the c-index and residuals. To reduce the chance of the model over-fitting the data, we applied Bayesian regularisation in the form of penalised
maximum likelihood (Harrell, 2001) which involved adding a quadratic penalty term to the likelihood function, corresponding to a Gaussian prior distribution on the model parameters. This simplified the model and tailored it to the complexity of the data so that predictions could be extrapolated to all patients and not only those used to train the model. Such simplification was achieved by forcing the parameters to have as small a value as possible.

**Figure 1** Probability density function of the event data with the log-logistic fit model

All modelling and data analysis were carried out in the R software package, which can be found at the URL: http://cran.r-project.org

The model was implemented for use in the MATLAB language (The MathWorks Inc., Natick).

### 2.4 Validation

Two validation measures were calculated, discrimination and calibration. Discrimination described the ability of the model to rank the outcomes as a function of the prognostic factors. Discrimination was expressed in terms of the $c$-index, which determines the diagnostic power of a test applied to censored data. Calibration described the precision of the predictions compared with actuarial outcome for different risk groups. Calibration was assessed by calculation of Cox-Snell residuals, which take into account the censoring mechanism (Cox and Snell, 1968; Taktak et al., 2008).

### 3 Results

#### 3.1 Cohort

The cohort comprised 3653 patients. Histologic data were available for 1778 tumours, of which, all contained the variable Epi (i.e., melanoma cytomorphology). Of these, 1502 (84.5%) also contained either Loops, Mitosis or both. Mon3 data were available in 738 subjects of which 712 (96.5%) also contained the variable Epi.
A total of 1235 (34.1%) patients had died from all causes (33.8%). The median follow-up time was 5.6 years (range, 0.01 – 40.54). A total of 2013 (55.1%) patients had survival time of more than five years and 1013 (27.7%) more than 10 years. A summary of univariate analysis is shown in Figure 2.

Figure 2  Univariate analysis of the input parameters. For continuous variables: (a) age and (c) LUD; these are represented as a box plot showing the range of the variables against outcome at five years (0 – died, 1 – survived). For ordinal variables: (b) sex; (d) ora; (e) EO; (f) epi; (g) loops; (h) mitosis and (i) mon3, these are shown as a bar chart representing percentage surviving >5 years for at each level. The numbers in brackets represent the number of subjects in each category.
Figure 2  Univariate analysis of the input parameters. For continuous variables: (a) age and (c) LUD; these are represented as a box plot showing the range of the variables against outcome at five years (0 – died, 1 – survived). For ordinal variables: (b) sex; (d) ora; (e) EO; (f) epi; (g) loops; (h) mitosis and (i) mon3, these are shown as a bar chart representing percentage surviving >5 years for each level. The numbers in brackets represent the number of subjects in each category (continued)
Figure 2  Univariate analysis of the input parameters. For continuous variables: (a) age and (c) LUD; these are represented as a box plot showing the range of the variables against outcome at five years (0 – died, 1 – survived). For ordinal variables: (b) sex; (d) ora; (e) EO; (f) epi; (g) loops; (h) mitosis and (i) mon3, these are shown as a bar chart representing percentage surviving >5 years for at each level. The numbers in brackets represent the number of subjects in each category (continued)
Figure 2  Univariate analysis of the input parameters. For continuous variables: (a) age and (c) LUD; these are represented as a box plot showing the range of the variables against outcome at five years (0 – died, 1 – survived). For ordinal variables: (b) sex; (d) ora; (e) EO; (f) epi; (g) loops; (h) mitosis and (i) mon3, these are shown as a bar chart representing percentage surviving >5 years for at each level. The numbers in brackets represent the number of subjects in each category (continued)
Enhancing survival prognostication in patients with choroidal melanoma

Figure 2 Univariate analysis of the input parameters. For continuous variables: (a) age and (c) LUD; these are represented as a box plot showing the range of the variables against outcome at five years (0 – died, 1 – survived). For ordinal variables: (b) sex; (d) ora; (e) EO; (f) epi; (g) loops; (h) mitosis and (i) mon3, these are shown as a bar chart representing percentage surviving >5 years for at each level. The numbers in brackets represent the number of subjects in each category (continued)

![Bar chart showing survival rates](image)

(i)

3.2 Model validation

The c-index of discrimination was 0.75 (95% CI, 0.74–0.76) for the clinical model and 0.79 (95% CI, 0.76–0.82) for the laboratory model. Calibration tests showed good agreement between predicted and observed mortality for both the clinical and laboratory models as shown in Figure 3. The estimated survival function $\hat{S}(t)$ was evaluated at the observed and censored event times. According to theory, the cumulative distribution $1-\hat{S}(t)$ should be uniformly distributed in (0, 1). This equality can be formally checked by a Kolmogorov-Smirnov test statistic, modified to handle random censoring, the null hypothesis being the equality of the distributions. For the clinical model the KS statistic was 0.8774204 ($p = 0.699$). For the laboratory model the KS statistic was 0.7980748 ($p = 0.8005$). In both cases the null hypothesis could not be rejected.

As can be seen from Figure 4(a) and (b), the clinical model had narrower confidence intervals because a larger dataset was used in fitting the model. Further validation was carried out between the predicted and observed survival according to the estimated risk of metastatic death at five years (i.e., low, medium or high), both for the clinical model (Figure 4(a)) and the laboratory model (Figure 4(b)). All pairs of survival curves correlated well with each other.
Figure 3  Cox-Snell censored residuals matching observed vs. predicted mortality to assess goodness of fit in the (a) clinical model and (b) the laboratory model. A perfect fit lies on the 45° line.
Figure 4 Kaplan-Meier and model predictions stratified according to estimated risk of mortality at five years (i.e., low, medium and high) in the (a) clinical model and (b) laboratory model. The model predictions were obtained by averaging the estimated survival curves of all the patients in each group.

Continuous line: Model.
Dashed lines: Kaplan-Meier.
4 Discussion

Our models accurately predicted survival after treatment of choroidal melanoma by integrating histopathologic results with clinical and genetic findings. Good prognostication was achieved even when data were incomplete; this enhanced the value of prognostic tumour biopsy, increasing the scope of this procedure in clinical practice despite missing mitotic count and closed loops.

Unlike Kaplan-Meier analysis our models allow multivariate analysis and cope with continuous variables, such as basal tumour diameter. In contrast to Cox analysis, our models do not require hazards to be proportional over time. Our present models are superior to our previous neural networks in that they are better at interpolating results to compensate for missing data such as mitotic count and closed loops.

The main limitations arise the clinical, histological and genetic data. Clinical assessment of basal tumour diameter is imprecise, especially if tumour margins are diffuse. Histological assessment of melanoma cytomorphology is subjective even between experienced pathologists, and mitotic counts can be difficult (e.g., in necrotic tumours or in small biopsies). Genetic typing is technically demanding and lethal abnormalities may be missed if the techniques applied lack sufficient resolution or if intratumoral genetic heterogeneity results in sampling error.

Before the model can be applied to data from other centres, external validation needs to be carried out especially if examination methods and terminology are different from ours (e.g., if basal tumour diameter is measured ophthalmoscopically instead of by ultrasonography or if pars plana is not considered to form part of the ciliary body).

Until now, treatment of metastases has rarely prolonged life so that any impact on the training of our models has been minimal. In future, however, patients with a high risk of developing metastases will be closely followed with the aim of treating them early (e.g., surgical excision of isolated metastases). Patients responding to such systemic or hepatic therapy may need to be excluded to avoid bias.

The model used all-cause mortality and not disease-specific mortality as an output. However, because of the rarity of uveal melanomas it was possible to estimate the likelihood of metastatic death by comparing the patient’s survival curve with that of the matched general population (Hakulinen and Dyba, 2006). The main advantage of this strategy is that our model does not rely on certified cause of death, which is known to be unreliable. Another advantage is that there is no bias caused by censoring patients who die of competing risks, such as old age.

There is much scope for further work. At present, our models do not include gains in chromosome 6p and 8q, which respectively are associated with good prognosis and poor prognosis (Damato et al., 2010). Once a sufficient numbers of deaths have occurred, we hope to re-train our models so as to use as much genetic data as possible. Ultimately, we hope that data on all 31 genetic loci tested will be incorporated into the analysis thereby avoiding the need for subjectively deciding whether to categorise chromosome 3 as being normal or abnormal.

Others have described neural networks for predicting survival after treatment of uveal melanoma, but these studies are limited by small patient numbers and short follow-up (Kaiserman et al., 2005).

Our work has several clinical and research implications. Our improved prognostication is already enabling us to reassure patients with a good prognosis and to advise them that systemic screening for metastases is unlikely to be helpful to them.
This enhances their sense of well-being and spares them from unnecessary and repeated investigations over many years, which are stressful, time-consuming and wasteful of healthcare resources. Conversely, patients with a poor prognosis are alerted to their condition, after confirming that they do indeed want to know their survival prospects. Such patients undergo six-monthly screening with biochemical liver function tests and abdominal imaging. As a result, several have undergone partial hepatectomy or been entered into trials of chemotherapy. Clinical trials investigating methods of screening for metastasis and pharmacological treatment of asymptomatic disease have been undertaken. These benefits have occurred thanks to the improved prognostication resulting from our multivariate analysis of clinical, histological and genetic data.

In future, we hope to be able to enter our patients into randomised trials of systemic adjuvant therapy, using our models to exclude patients without sufficient risk of metastasis. A study on adjuvant intra-arterial chemotherapy showed a non-significant trend towards clinical benefit (Voelter et al., 2008); however, patients were selected only on the basis of their tumour dimensions and it is likely that a significant proportion of these patients had a non-lethal melanoma. If such patients had been excluded, statistical significance may have been achieved, resulting in adoption of this adjuvant therapy.

Our prognostic methods also enhance opportunities for basic research on tissue specimens, so that immunohistochemical, genetic and other features can be correlated with lethality without waiting several years for survival time to be measured.

5 Conclusions

Prognostication is enhanced by integrating histopathological findings with clinical and genetic data, also taking normal life expectancy into account. This has improved our patient care. It also improves prospects for clinical trials evaluating the impact of ocular or systemic therapy on survival. By classifying tumours according to risk of metastasis, our models also enhance opportunities for identifying histological and genetic predictors of survival. Although choroidal melanomas are rare, the methods we have described should be applicable to other cancers.

Acknowledgements

We are grateful to the National Commissioning Group for funding our ocular oncology service and the ophthalmic pathology service in Liverpool.

References


