Published in final edited form as:
Methods Mol Biol. 2014 ; 1102: 227–258. doi:10.1007/978-1-62703-727-3_13.

Construction and analysis of multi-parameter prognostic models for melanoma outcome

Bonnie E. Gould Rothberg[1,2,4] and **David L. Rimm**[3,4]

[1]Department of Internal Medicine, Section of Medical Oncology, Yale School of Medicine

[2]Department of Epidemiology and Public Health, Yale School of Medicine

[3]Department of Pathology, Yale School of Medicine

[4]Yale Cancer Center

Summary

The outcome of Stage II melanoma is uncertain. Despite that 10-year melanoma-specific survival can approach 50% following curative-intent wide local excision and negative sentinel lymph node biopsy, the adverse risk-benefit ratio of interferon-based adjuvant regimens precludes their use in most patients. The discovery and translation of protein-based prognostic biomarkers into the clinic offers the promise for residual risk stratification of Stage II melanoma patients beyond conventional clinicopathologic criteria to identify an additional subset of patients who, based upon tumor molecular profiles, might also derive benefit from adjuvant regimens. Despite incorporation of Ki-67 assays into clinical practice, systematic review of REMARK-compliant, immunostain-based prognostic biomarker assays in melanoma suggests that residual risk of recurrence might be best explained by a composite score derived from a small panel of proteins representing independent features of melanoma biology. Reflecting this trend, to date, 5 such multi-parameter melanoma prognostic models have been published. Here, we review these 5 models and provide detailed protocols for discovering and validating multi-parameter models including: appropriate cohort recruitment strategies, comprehensive laboratory protocols supporting fully quantitative chromogenic or fluorescent immunostaining platforms, statistical approaches to create composite prognostic indices recommended steps for model validation in independent cohorts.

Keywords

Multi-parameter models; Prognosis; Systematic Review; Multivariate statistical modeling; Immunohistochemistry; Quantitative immunofluorescence; Tissue microarray

1. Introduction

The prognosis for Stage II melanoma is uncertain. Despite curative-intent wide local excision and negative sentinel lymph node dissection, the 10-year melanoma-specific survival ranges from 65% in Stage IIA to 40% for Stage IIC [1]. Yet, due to the morbidity associated with currently approved interferon-based therapeutic regimens, adjuvant therapy is only recommended for Stage IIB/C patients with '*high-risk*' features such as ulceration and/or in-transit metastases [2] with active surveillance as the standard of care for the remaining patients. To maximize the potential benefit of adjuvant treatment among Stage II melanoma patients, current melanoma research priorities include the identification of

Corresponding Author: David L. Rimm, MD, PhD, Brady Memorial Laboratories Room 116A, 310 Cedar Street, New Haven, CT 06520-8031, Phone: 203-737-4205, Fax:, david.rimm@yale.edu.

complementary biomarkers that can stratify additional Stage II patients as high-risk with the overall goal of reducing melanoma mortality among patients for whom cure is possible. In addition to the consideration of mitotic index, which has recently been shown to be independently prognostic among Stage II melanomas [3], a significant focus surrounds the adjudication of in situ expression of candidate proteins, as assessed by immunostaining on paraffin-embedded tissues, to identify signatures corresponding to melanomas with poor melanoma-specific survival.

To identify those protein candidates with potential for translation into the clinic as prognostic indicators from among the vast literature of evaluated candidates, we published a systematic review and meta-analysis critically evaluating the body of published melanoma immunostaining data through January 2008 [4]. Of the 455 identified manuscripts that reported immunostaining data on melanoma samples, only 37 manuscripts reported multivariable survival estimates across either prospective or retrospective cohorts while meeting REMARK criteria [5] for immunostaining assay rigor for 62 unique individual protein candidates. Schramm and Mann updated our review through December 2009, identifying an additional eligible 14 studies and bringing the total number of rigorously evaluated proteins to 86 [6]. For this report, we extended our systematic review search criteria through August 15th, 2012 and identified an additional 15 eligible manuscripts collectively evaluating 12 proteins including 9 candidates not previously considered in prior reports [7–21].

The incremental value of molecular markers added to clinicopathologic prognostic models is dependent on both the magnitude of the effect size and the inverse of the correlation with the clinicopathologic parameters already in the model [22]. Altogether, across the 66 included studies, statistically significant ($p<0.05$) multivariable adjusted Cox Proportional Hazards results were reported for 14/49 protein associations with overall survival, 25/47 for melanoma-specific survival and 20/31 for melanoma-free survival [10,4,7,9,6,8,11–21]. Interestingly, of the 58 significant associations, only 20 results representing 17 markers, including Ki-67, achieved a hazard ratio ≥3.0 with 13 of these based upon total sample sizes of <100 (Table 1). Furthermore, only metallothionein was independently validated with just the disease-free survival association repeated in an independent cohort [23,24]. None of the other extreme values were independently recapitulated. By comparison, individual proteins whose significant prognostic hazard ratios were independently recapitulated across multiple cohorts (e.g., p16/INK4A, Matrix metalloproteinase-2) displayed more modest, but consistent, p-values in each of these studies (Table 2). Taken together, these data suggest that the independent marginal gain in risk prediction over the base clinicopathologic prognostic models for most independent markers might not be of substantial clinical benefit to warrant further translation.

Consequently, one alternate strategy with the potential for augmenting the clinical relevance of melanoma-related molecular predictors is to assemble a portfolio of orthogonally prognostic proteins that, when combined as a single parameter, yield a high-magnitude independently prognostic measure. Creation of multi-marker prognostic models from differential transcriptome profiling is well-established with both the Oncotype DX [25,26] and MammaPrint [27,28] tests approved for breast cancer prognostic stratification and, similarly, the Oncotype Dx Colon Cancer [29] and ColoPrint [30] offer risk stratification for patients with Stage II colorectal cancer. Moreover, the immunohistochemistry-based 5-protein Mammostrat panel to identify estrogen receptor-positive breast cancer patients at high risk for relapse despite estrogen-modulating adjuvant therapy has recently been validated in a prospective, randomized clinical trial [31]. Similarly, in melanoma, multi-parameter-based approaches for prognostic biomarkers are beginning to emerge as our updated systematic review identified 5 manuscripts that each proposes a prognostic index

based upon the combination of 3 to 7 individual biomarkers with the potential for translation into the clinic (Table 3) [32,33,7,34,8]. Eighteen unique proteins were included across the 5 models with only 3 proteins – p16/INK4A, p53 and β-catenin – included in multiple models. Altogether, the 5 models highlight the diversity of both laboratory and statistical methodologies as well as the breadth of protein combinations. Although the most straightforward approaches leverage standard semi-quantitative chromogenic immunostaining, build a composite model from the subset of protein candidates that yield significant univariate survival relationships and create an index score reflecting the number of markers that exceed a specified threshold, more objective approaches that integrate fully quantitative measures of protein expression or more complex, parameterized statistical models are beginning to emerge.

In this review, we present a stepwise approach for discovering and validating novel multi-parameter biomarker-based prognostic models. We outline appropriate cohort recruitment strategies, comprehensively describe laboratory protocols supporting fully quantitative assessment of protein expression using either chromogenic or immunofluorescent platforms and present a suite of innovative statistical approaches for combining the assayed markers. Finally, we present recommended steps for validating the model in an independent cohort, necessary to support potential generalizability to the greater melanoma population.

2. MATERIALS

2.1. Tissue microarray (TMA) construction

1. Paraffin blocks from each of the index primary melanomas of all eligible individuals accessioned in the Discovery and Validation cohorts.
2. Paraffin blocks from tissue *reference controls* and *cell line reference controls*.
3. Stainless steel coring needle (Estigen OÜ, Tartu, Estonia).
4. Tape transfer system (Leica Microsystems, Buffalo Grove, IL).
5. Nitrogen chamber/dessicator (Terra Universal, Fullerton, CA).

2.2 Antigen retrieval

1. 60 °C hybridization oven (we use an HB-1000 hybridization oven (UVP, Upland, CA)).
2. Stainless steel, 30-slide staining dish (500 mL) and racks (Mopec, Oak Park, MI).
3. Xylene.
4. 100% ethanol and 75% ethanol.
5. Antigen Retrieval Buffer: (6.5 mM sodium citrate (3.84 g sodium citrate in 2L ddH_2O)), pH 6.0 or 1.3 mM EDTA (0.76 g EDTA in 2 L ddH_2O), pH=8.0.
6. Pre-Treatment (PT) Module automated dewaxing and epitope recovery device (Lab Vision, Fremont, CA).
7. 0.75% hydrogen peroxide in absolute methanol.
8. Tris-buffered saline (TBS, 1X, 8.76 g sodium chloride and 2.42 g trizma base/L, pH=8.0)
9. TBS with 0.5% Tween-20.
10. 0.3% bovine serum albumin (BSA) in TBS/0.5% Tween-20 (1.5 g BSA dissolved in 500 mL 1X TBS with 0.5% TWEEN-20).

2.3. Quantitative chromogenic immunohistochemistry (IHC) using the Aperio ScanScope™ CS brightfield platform

1. Anti-target primary mouse/rabbit monoclonal antibodies.
2. PAP pen (Kiyota International, Elk Grove Village, IL).
3. EnVision™-HRP anti-target species (mouse/rabbit) secondary antibody (DAKO, Carpenteria, CA).
4. 3,3′-diaminobenzidine (DAB) or3-amino-9-ethylcarbazole (AEC+) substrate chromogens and Substrate Buffer (DAKO, Carpenteria, CA).
5. TACHA hematoxylin (Biocare Medical, Concord, CA).
6. Cytoseal 60™ mounting medium (Richard-Allan Scientific, Kalamazoo, MI).
7. Aperio ScanScope® CS digital light microscope platform with *Console* and *Spectrum* software platforms (Aperio, Vista, CA).

2.4. Quantitative immunofluorescence (QIF) using the AQUA® technique

1. Anti-target primary mouse/rabbit monoclonal antibodies.
2. Mask antibodies: Rabbit anti-S100 polyclonal antibody (DAKO, Carpenteria, CA) and anti-gp100 polyclonal antibody (Abcam, Cambridge, MA) or Mouse anti-S100B monoclonal antibody (clone 15E2E2, Abcam, Cambridge, MA) and HMB45 monoclonal antibody (Abcam, Cambridge, MA).
3. PAP pen (Kiyota International, Elk Grove Village, IL).
4. AlexaFluor 546-conjugated goat anti-mask species secondary antibody (Life Technologies, Carlsbad, CA).
5. EnVision™-HRP anti-target species (mouse/rabbit) secondary antibody (DAKO, Carpenteria, CA).
6. 10x Cy5-tyramide and Amplification Buffer (Perkin-Elmer Life Sciences, Waltham, MA).
7. Prolong Gold mounting medium with 4′6-diamidino-2-phenylindole (DAPI) (Life Technologies, Carlsbad, CA).
8. Digital fluorescent image capture platform – Caliper Life Sciences Vectra® 2 (Caliper Life Sciences, Hopkinton, MA) or Aperio ScanScope® FL (Aperio, Vista, CA).
9. *AQUAsition*™ image capture and *AQUAnalysis*™ image analysis software packages (HistoRx, Branford, CT).

3. METHODS

3.1. Selection of cases for Discovery and Validation cohorts (see Note 1)

1. Obtain institutional review board approval for the use of human tissue specimens and to obtain their corresponding clinicopathological data and melanoma-related outcomes.
2. Define the sampling frame for each of the *Discovery* and *Validation* cohorts.
3. Define the study inclusion/exclusion criteria (e.g., Stage II melanoma cases only, exclusion of all new diagnoses in children <18 years of age).

4. Query the appropriate hospital records or population-based tumor registry to ascertain the complete list of all potentially eligible melanoma cases according to the defined inclusion/exclusion criteria.
5. Query the appropriate surgical/dermatopathological tissue repositories to identify the subset of eligible patients with available residual tissue.
6. Review diagnostic H&E sections and complete medical record abstraction for all eligible cases, whether residual tissue is available or not.
7. Within each of the Discovery and Validation cohorts, to rule out the potential for selection bias due to availability of residual tissue, compare the distribution of clinicopathologic criteria between those cases with and without available tissue.
8. To ensure the comparability of the *Discovery* and *Validation* cohorts, compare their respective distributions of the conventional clinicopathologic parameters planned for inclusion in multivariable modeling.

3.2. Construction of tissue microarrays (see Note 2)

TMAs are constructed using the standard method [35] and cut sections are adhered to the glass slides using a tape transfer technique.

1. Retrieve the paraffin blocks and their associated cut H&E slides corresponding to the primary tumors from all eligible *Discovery* and *Validation* cohort members from the tissue archive (*see* Note 3)
2. Review the H&E slide to confirm the presence of residual tumor and, using a fine-tip marker, circle on the H&E slide the region most representative of the primary tumor for inclusion onto the TMA.
3. To allow for adequate modeling of intra-tumoral heterogeneity and to guard against missing values, inclusion of 4 0.6 mm histospots per case in melanoma tumors is recommended (*see* Note 4).

---

[1]Correct study design is paramount for developing melanoma prognostic models with translational potential. REMARK criteria require that biomarker studies rigorously report the method of case selection, including details regarding methods of ascertainment (prospective or retrospective) and any inclusion/exclusion criteria that lead to stratification or matching as well as the time period from which cases were taken, then end of the follow-up period and the median follow-up time [5]. In our systematic reviews [4,69], we have further stipulated that study design must conform to the definition of a prospective or retrospective cohort [70]. The sampling frame must be defined with recruitment strategies striving for complete source population enrollment and documenting study non-participation. Convenience samples comprised of a subset of eligible cases (e.g., case series), no matter how large, risk introducing bias. We also advocate against conducting case-control studies since they do not support survival analysis. Simon *et al.* [71] advocate for additional study design rigor, requiring biomarker validation in prospective cohort studies or randomized clinical trials. At a minimum, two independent cohorts must be identified for the initial model development. A *Discovery Cohort* is used to explore the independent prognostic utility of the component markers, to conduct feature selection and to execute the initial multivariable Cox proportional hazard modeling. A *Validation Cohort* is required to verify the model's robustness through its generalizability across additional cohorts. In anticipation of their synergistic use, it is critical that the Discovery and Validation cohorts have parallel inclusion/exclusion criteria. For example, if the Discovery cohort is restricted to patients with Stage II melanoma, the correct Validation cohort would be similarly restricted to Stage II melanoma cases.

[2]Development of multi-parameter biomarker assays for melanoma prognosis requires a significant initial investment in candidate protein evaluation. In our experience [32], as well as that of others [7] between 6–10 candidates are triaged through the Training set for every marker selected for inclusion in the model. As a result, model discovery is most efficiently conducted using TMAs. Not only do TMAs allow for efficient rationing of small primary melanomas but they also ensure that all tumors are evaluated concurrently using the same batch of laboratory reagents. However, the risk of using TMAs is that they do not directly translate to patient care and require further validation cohorts that use conventional sections prior to introduction into laboratories for clinical usage.

[3]For melanoma patients who have multiple independent primary tumors, inclusion of material from multiple lesions is possible. However, these data must be summarized such that each patient enters only once into the survival analyses assessing prognosis.

[4]Although the within- and between-sample variance of expression, the principal determinants of optimized histospot number, is variable and marker-dependent [72], melanoma TMA validation experiments have shown over 95% concordance between 4 randomly selected cores and whole slides [73] and this approach has been successfully applied across a series of candidate proteins [74,75].

4. In addition to the cohort histospots, a series of controls must also be included in the TMA.

    a. *Tissue reference controls*. These controls allow for the normalization of the quantitative scores to a single reference point across multiple builds of a single TMA or between the training and validation TMAs, especially when these experiments were conducted at separate, distinct time points. Specifically, these provide the reference standards allowing correction for any laboratory drift that may occur. These are selected as follows:

        i. Three to four benign tissues with anticipated homogeneous expression are selected. Our group selects liver, kidney medulla and uterine smooth muscle but other tissue choices can be made.

        ii. Paraffin blocks of each targeted tissue from 8–10 individuals are retrieved and included in the TMA for a total of 24–30 0.6 mm histospots dedicated to these controls.

    b. *Cell line controls.* Formalin-fixed, paraffin-embedded pellets from 10–12 melanoma cell lines representing the diversity of known somatic mutation subtypes (e.g., BRAF mutants, NRAS mutants, RAC1 mutants and BRAF/NRAS/RAC1 wild type cell lines) are prepared (see Moeder *et al.* **Ref.** [36] for complete method) in parallel with a series of cell lysates. Cores from these pellets are included in duplicate on the TMA to both serve as known internal positive and negative controls (*see* **Section 3.3.1)**: Primary anti-biomarker antibody selection and validation) for the primary antibody and, among those positively expressing the antigen, to verify within-sample heterogeneity of expression.

5. To ensure an unbiased array of the tumor samples and random admixture of all the control samples, TMA row and column location are assigned using a random number generator. Each planned core is randomly assigned a number between 0 and 1, rank ordered according to that number assignment and then row/column addresses are sequentially given.

6. Donor blocks are aligned according to row/column assignment and 0.6 mm cores, created with a stainless steel coring needle (Estigen OÜ, Tartu, Estonia) are arrayed into the recipient paraffin block to create the microarray [35].

7. The completed TMA block is then faced and cut using the tape-transfer method (Leica Microsystems, Buffalo Grove, IL). Tape-transfer substantially reduces the number of lost histospots and distortion of the histospot array that can occur during sectioning [37] and, in our experience, does not interfere with immunostain quality or readability.

    a. Place an adhesive Tape Window to the block face. The tape supports and captures the section.

    b. The Tape Window is placed section-side-down on corresponding, complimentary adhesive-coated slides and laminated to the slide using the included hand roller.

    c. Expose the slide to 30–60 sec of 360 nm ultraviolet light to polymerize the slide adhesive layer into a hard, solvent-resistant plastic that anchors the section to the slide.

d. Remove the Tape Window by immersing the section in TPC solvent for 3 min.

e. Dip slide briefly in Xylene to dry and then coat the slide in paraffin and store in a dessicator (Terra Universal, Fullerton, CA) maintained at 10.5% relative humidity using nitrogen gas until use to preserve tissue antigenicity by reducing exposure of the cut face to ambient oxygen and relative humidity [38].

3.3. Methods for biomarker quantification

Diverse commercial quantitative immunostaining platforms with capabilities for brightfield, chromogenic immunostaining or quantitative immunofluorescence are available (Table 1). Brightfield platforms are based upon the visualization of chromogenic stains (e.g., DAB or AEC) against a hematoxylin counterstain. Machine-learning-based feature extraction algorithms or other more simple sized based methods can resolve relevant cellular (e.g., tumor, stroma, vasculature, inflammatory infiltrates) and subcellular (e.g., nuclear vs. cytoplasmic) compartments. The definition of these compartments may be dependent on the counterstain or other methods. Protein expression is then quantified by measuring the levels of white light absorbance by the chromogenic stain, typically operationalized as the average intensity of the stain's color (hue and saturation) across all pixels included in a selected compartment [39–41]. By comparison, QIF multiplexes the target antibody fluorophore (typically at or above 650 nm to avoid contamination with tissue autofluorescence) with fluorophores directed against relevant cellular and subcellular features. In this manner, cellular and subcellular compartments are resolved through co-localization with the selected fluorescent signal and protein expression is quantified by measuring the level of light emission at the appropriate wavelength corresponding to the target fluorophore within the compartment(s) of interest [42]. Multiplexing of targets can be executed by selecting fluorophores with non-overlapping emission spectra [43]. By example, we provide detailed methodologic protocols for quantitative chromogenic IHC optimized for the Aperio ScanScope CS brightfield platform and its associated Positive Pixel Count Algorithm as well as QIF using the AQUA® technique.

3.3.1. Primary anti-biomarker antibody selection and validation—Prior to initiating any immunostaining assays to assess levels of the candidate biomarker, advance consideration must be given towards choice of the primary anti-biomarker antibody. First, where possible, assay development with a monoclonal antibody is preferred. Once established, monoclonal antibody reagents are synthesized in vitro from an immortalized lymphocyte cell line which not only ensures a steady, consistent supply of quality reagent over the complete life cycle of assay development but also, if translation into the clinic is warranted, the antibody reagent quality and consistency can be validated according to Clinical Laboratory Improvement Amendments (CLIA) standards to support necessary clinical trial work. In contrast, while use of polyclonal antibodies can indeed uncover novel significant and biologically-relevant biomarkers for melanoma prognosis, translating their use into the clinic is less straightforward and uncertain at best. The most striking consideration is that, unlike an immortalized, clonal cell line capable of scalable monoclonal antibody production that is potentially unlimited, the availability of any polyclonal reagent is limited at any one time by the amount of blood obtained from host organism during the lifespan of the individual animal. The specific polyclonal mixture is exhausted upon the host's death and may never be regenerated in an identical manner due to variability in the immune systems in the host species, even when challenged with identical antigen. Although QC metrics for creating similarly effective polyclonal reagents for molecular techniques such as chromatin immunoprecipitation from across a population of immunized rabbits are routinely used [44], antibody-based projects requiring a high degree of standardization such

as the prognostic biomarker assays described here, are best served with monoclonal antibodies [45].

Next, all selected antibodies must undergo comprehensive validation procedures to verify that the antibody is both sensitive and specific for its presumed target. The method preferred by our lab has been recently published [46]. One simple method of antibody validation is to conduct a Western blot on a spectrum of cell lines including one with previously confirmed expression of the target protein and one where protein expression is known to be absent. In this assay, the goal is to verify specific band(s) at the confirmed expected molecular weight(s) for the target and to rule out visualization of any additional electrophoretic signals at spurious molecular weights. While a Western blot can eliminate an antibody due to cross-reactivity, absence of signal does not necessarily mean the antibody will not work in IHC or QIF applications. When the Western Blot is non-informative due to lack of antibody reactivity with the denatured antigen, it may still be valuable and specific in the less denatured conditions seen in tissue sections, so in situ validation methods are recommended. “Blocking peptide” methods where a soluble peptide expected to adsorb to the antibody are not recommended since they cannot confirm specificity as they do not rule out binding to off-target sites on the tissue sections. Rigorous in situ antibody validation is best achieved using small interfering-RNA (si-RNA) in vitro knockdown experiments on positively expressing cell lines or progressive induction of target antigen in non-expressing lines. Here, levels of protein expression on the selected cell line(s) are assayed by IHC or QIF in the presence/absence of siRNAs. If levels of the protein do not decrease with target knockdown, cross-reactivity with alternate antigens is suspected and the antibody should be abandoned. Alternatively, transfection and progressively increased induced expression can also show in situ specificity [47].

3.3.2. Antigen retrieval—The initial steps of the immunostaining protocol through application of the primary antibodies are identical for both brightfield and immunofluorescent techniques as described below:

1. Align slides in a metal slide rack and place the rack in a 60 °C oven for 20 min to overnight to melt off excess paraffin. For sections not adhered to the glass slides using an ultraviolet crosslinked tape-transfer method, overnight incubation at 60 °C is necessary to cure the slides.
2. Deparaffinize the tissues using two exchanges of 400 ml xylene for 20 min each.
3. Rehydrate the tissues through two exchanges of 400 ml 100% ethanol and 1 exchange of 75% ethanol for one min each and then place the slides under running water for 10 min.
4. Pour 2 L of Antigen Retrieval Buffer (6.5 mM sodium citrate, pH 6.0 or 1.3 mM EDTA pH=8.0) in the PT Module vessel and preheat the unit to 85 °C.
5. Once the buffer has reached the pre-heating temperature, place the slides in the Module using the accompanying slide rack, seal the device and set antigen retrieval at 97 °C for 20 min.
6. Following antigen retrieval, once the PT Module has cooled down to 75 °C, remove the slides from the Module, place them under cold running water for 10 min and then submerge the slides in 400 ml absolute methanol containing 0.75% hydrogen peroxide for 30 min to quench endogenous peroxidase activity.
7. Wash the slides in two exchanges of water for 30 sec and in 1X TBS for 5 min.
8. Blot each slide dry and apply 1 mL of TBS-Tween/0.3% BSA solution to the exposed tissue section for 30 min to adsorb non-specific binding sites.

9. Blot off the TBS-Tween/BSA and proceed to 3.3.3.1 (chromogenic immunostaining for brightfield platforms) or 3.3.3.2 (quantitative immunofluorescence) to begin with primary antibody application for the desired platforms.

3.3.3.1. Quantitative chromogenic IHC using the Aperio Scanscope™ CS brightfield platform (see Note 5)

1. Prepare the desired titer of the validated primary target antibody by diluting it into TBS-Tween/0.3% BSA to create a final volume of 1 mL.
2. Place the TMA slide in a level humid chamber and pipette the full 1 mL volume of the primary antibody cocktail onto the TMA section, ensuring that the full surface of the section is evenly covered with fluid. To create a hydrophobic barrier that ensures retention of the antibody cocktail on the section, outline the TMA boundary with a PAP pen.
3. Set the slide to incubate at 4 °C overnight.
4. Wash off the excess primary antibody by placing in 2 10-min exchanges of TBS-Tween followed by 1 10-min exchange of TBS.
5. Apply 1 mL of species-appropriate EnVision™ HRP-labeled polymer secondary antibody and incubate at room temperature for 1 hour.
6. Wash the EnVision™ by placing in 2 5-min exchanges of TBS followed by 1 5-min exchange of distilled water.
7. Apply 1 mL of the diluted AEC+ substrate chromogen (20 μL chromagen diluted into 980 μL Substrate Buffer), incubate for 5 min and then wash off by applying 2 5-min exchanges of TBS and 1 5-min exchange of distilled water.
8. Apply 1 mL TACHA hematoxylin to the TMA and incubate at room temperature for 1 min. In our experience, as the TACHA hematoxylin does not require a bluing step, it yields a more consistent color and reduces measurement error during automated assessment of staining intensity.
9. Rinse the slide in two exchanges of distilled water, dehydrate through an ethanol gradient of 1-min exchanges at 70%, 85%, 95% and 100% then place in a final 100% ethanol exchange for 5 min.
10. Transfer the slides into xylene for 5 min then remove each slide one at a time from the xylene, blot dry, apply 1–2 drops of CytoSeal 60™ mounting medium, coverslip each slide and allow the mounting medium to set.
11. Automated image capture is accomplished on the Aperio ScanScope CS using the *Console* application installed with the machine. Briefly, 1–5 stained slides are loaded into the tray and a photomicrograph snapshot of each slide covering all magnifications from 10x to 200x is captured.
12. The region of the slide containing the stained tissue is identified and each histospot is identified within the TMA by engaging the *Find focus points* algorithm. At this

[5]Automated quantitative brightfield platforms measure immunostain intensity as the level of white light absorbance in the range of the particular color of the selected chromogen, defined by its hue and saturation. To avoid possible mismeasurement of the selected antigen, the color of the chromogen must be distinct from any naturally occurring cellular melanin containing inclusion body. Recognizing that the range of browns associated with melanin can overlap with the color produced by the DAB chromogen, we suggest using a red chromogen such as AEC or Vector Red [76,77] to minimize possible target protein measurement error due to confounding by endogenous melanin pigment.

step, histospots not automatically identified can be marked by the user and slide artifacts erroneously tagged can be manually removed.

13. Automated quantification of levels of staining is then executed using the provided *Spectrum* software package. The '*area of interest*', or the region containing the tumor cells is identified using a two-step process where first a single, closed polygon is drawn around the entire tumor-containing region of the histospot using the freehand '*Pen Tool*' and then the '*Negative Pen Tool*' is used to exclude smaller regions of non-tumor areas within the larger encircled region. This step introduces a level of subjectivity into the analysis. Other software tools like the Definiens software (Definiens, Parsippany, NJ) or the Genie software (Flagship Biosciences, Flagstaff, AZ) may be used to make this step more objective.

14. With the tumor region now defined, application of available algorithms allows for quantitation of immunostain intensity and percent coverage.

    a. The *Nuclear Algorithm* defines included nuclei according to light absorption according to the counterstain colorization with the subsequent ability to define levels of chromogen intensity within this segregated region. Outputted values include: total area covered in the nuclear analysis, average nuclear size, percent of nuclei with any positive stain and the average intensity of stain across all included nuclei.

    b. The *Positive Pixel Count Algorithm* identifies all pixels within the Annotated region and assigns to them a score ranging from 0 to 255 reflecting the chromagen intensity at the pre-selected color. Three user-defined arbitrary cut-offs along this continuum are then selected to define the boundaries separating 'no stain (0)', 'weak positive (+1)', 'moderate positive (+2)' and 'strong positive (+3)'. Outputted metrics then include: Average intensity ([Σpixel intensity scores (0–255)]/total number of included pixels), number of weak positive, moderate positive and strong positive pixels and percent pixel positivity (number of pixels with any positivity/total number of pixels). Users may define methods to produce a continuous scale if desired.

15. Histospot-level data describing all measured parameters are then exported through the *Spectrum* software platform into Excel for further integration with the clinicopathologic and outcomes measures as described below (**Section 3.4**).

3.3.3.2. Quantitative Immunofluorescence using the AQUA® technique (see Note 6):
The methods for conducting QIF/AQUA® are provided, highlighting its similarities and differences with the brightfield chromogenic methods outlined above (**Section 3.3.3.1**).

1. For each TMA slide, prepare 1 mL primary antibody solution in TBS-Tween/0.3% BSA by adding the desired titer of the validated primary target antibody in addition to the melanoma 'mask' cocktail of S100/gp100 antibodies that will allow for the discrimination of melanoma cells from surrounding non-melanoma stromal elements in the absence of counterstain. When non-rabbit anti-target primary antibodies are selected (e.g., mouse, rat), the preferred rabbit-based mask includes 5 μL (1:200) anti-S100 polyclonal antibody and 40 μL (1:25) anti-gp100 polyclonal

[6]Quantitative immunofluorescence (QIF)/AQUA® measures levels of protein expression by measuring the intensity of light emitted by the target Cyanine (Cy) 5 or Cy7 label across the subset of pixels that co-localize to certain cellular and subcellular architectural compartments, the latter defined by additional, non-overlapping fluorescent labels (e.g., Cy3, Fluorescein isothiocyanate (FITC)/Cy2, 4′6-diamidino-2-phenylindole (DAPI)) that are required in the absence of hematoxylin counterstain [42,36].

antibody. Using rabbit anti-target primary antibodies necessitates a mouse-based mask of 10 μL (1:100) anti-S100B clone 15E2E2 and 10 μL (1:100) of HMB45).

2. Place the TMA slide in a level humid chamber and pipette the full 1 mL volume of the primary antibody cocktail onto the TMA section, ensuring that the full surface of the section is evenly covered with fluid. To create a hydrophobic barrier that ensures retention of the antibody cocktail on the section, outline the TMA boundary with a PAP pen.
3. Apply the primary antibody cocktail to each slide and incubate in a humid chamber at 4 °C overnight.
4. Wash off the excess primary antibody by placing in 2 10-min exchanges of TBS-TWEEN followed by 1 10-min exchange of TBS.
5. Prepare the secondary antibody cocktail by adding 10 μL of AlexaFluor 546-conjugated goat anti-mask species into each 1 mL target species-appropriate EnVision™ HRP-labeled polymer secondary antibody.
6. Apply 1 mL of the secondary antibody cocktail to each TMA and incubate at room temperature for 1 hour.
7. Wash the secondary antibody cocktail by placing in 2 10-min exchanges of TBS-TWEEN followed by 1 5-min exchange of TBS.
8. Dilute 20 μL of concentrated 10X Cy5-tyramide into 980 μL of the accompanying Amplification Buffer and add fluorescent label to the Envision™ by applying 1 mL of 1X Cy5-tyramide to the TMA for 10 min.
9. Wash the excess Cy5 by placing in 2 10-min exchanges of TBS-Tween followed by 1 5-min exchange of TBS.
10. Apply 120 μL of Prolong Gold mounting medium with DAPI, the latter to visualize the nuclei, and coverslip allowing the mounting medium to set overnight before commencing automated image capture.
11. Automated image capture is compatible with the Caliper Life Sciences Vectra® 2 or the Aperio ScanScope® FL microscopy instrumentation platforms using the respective native software platforms. Image capture on the PM-2000™ is executed with the *AQUAsition*™ platform (HistoRx, Branford, CT), described in detail in **Ref.** [48]. Image analysis is universally executed by the *AQUAnalysis*™ software package (HistoRx). The current version of AQUA® is robust to inter-operator variation [49] and minimizes coefficients of variability through standard instrument calibration procedures (e.g., calibration cube, light source and Cy5 optical path factors [50]. The AQUA® method is summarized below.
    a. A 4X low-resolution image of the entire slide in the DAPI channel is captured and used to localize the specific Cartesian coordinates defining each histospot.
    b. A series of 20X images, one image for each fluorescent channel used (Nuclear (DAPI), Mask (Cy3), Target (Cy5)), is captured for each histospot. To maximize the dynamic range of the target immunofluorescent signal, the exposure time is automated for each channel in each histospot to ensure in yielding only 0.02% saturated pixels for the selected channel.
    c. Global TMA quality assessment is executed using the stack of Mask-labeled Cy3 images as a reference. Histospots that are devoid of tumor,

where the histospot has fallen off or where the image is substantially out of focus should be eliminated outright. Among the included histospots, specific regions that contain major artifacts such as folded tissue, air bubbles should be cropped out leaving the remaining, adequately stained histospot regions available for analysis.

d. The tumor mask is determined as the set of pixels that display a level of Cy3 staining above a pre-specified minimal threshold (binary gating) followed by spatial image analysis procedures (e.g., filling holes) to create multiple continuous regions corresponding to the tumor locations.

e. Subcellular compartments (e.g., nuclear, cytoplasmic) are then defined within the tumor mask using an unsupervised k-means clustering algorithm [49]. Each pixel is plotted on a two-dimensional scatter plot according to the DAPI and Cy3 intensities. The clustering algorithm is then applied according to three fundamental assumptions: 1) background pixels will have low intensities for both channels, 2) nuclear pixels will have high DAPI but low Cy3 intensity and 3) cytoplasmic pixels will have high Cy3 but low DAPI intensity. Pixels are then assigned to each cluster according to Euclidean distance from each of the 3 cluster centroids.

f. Levels of target protein expression within the tumor mask and in each subcellular compartment are reported as a continuous value ranging from 0–33,000 that represents the mean Cy5 channel intensity across the subset of pixels included in the selected compartment. As high-resolution 8-bit image capture generates 256 discrete intensity values per pixel, raw pixel intensity reflects the mean intensity value across the 256 readings. Next, to account for the spot-to-spot variability in exposure time, pixel *power*, the raw pixel intensity/exposure time is calculated. Raw AQUA score is then calculated as the mean *pixel power* across all pixels included in the selected compartment. Finally, corrections reflecting the instrument calibration standards are applied to yield a *normalized AQUA score* which is the product of the raw AQUA score and each of the calibration cube, light source and optical path factors.

3.4. Computational methods for feature selection, model construction and internal validation

REMARK stipulates adjudication of individual markers using Cox Proportional Hazards. But creating a multi-marker model by entering multiple biomarkers into a single Cox model adjusted for clinicopathologic variables is not effective as the sample size required to allow for sufficient power to adequately consider not only the main effects of each protein but also all of their pairwise and higher-order interactions is excessive. An efficient alternative is to create a single multi-marker parameter that reflects the composite prognostic impact of all component markers for subsequent evaluation in adjusted survival analyses.

A diversity of statistical approaches exist for *dimension reduction* and *feature selection*, the process of condensing and culling from a larger set of individual protein candidates the subset that best explain the variance observed in the Discovery Cohort and their subsequent parameterization within the multi-parameter model including: 1) weighted summation of individual variables, 2) genetic algorithms, 3) classification and regression trees (CART), 4) random forest analysis, 5) k nearest neighbors and learning-based methods such as 6) support vector machines and 7) discriminant analysis. In this section, we provide detailed methods for the first three approaches as each have been used to define multi-parameter models with prognostic potential in melanoma. We also outline the steps required for

internal validation of the results on the Discovery Cohort and provide a strategy for evaluating their prognostic potential in the Validation Cohort.

3.4.1. Cox univariate regression coefficient-derived summary estimates—Of the five published multi-parameter models in melanoma, four [33,7,34,8] have their composite measures derived from the individual marker univariate Cox Proportional Hazards regression estimate (*see* Note 7). Here, candidate proteins are selected for the summary measure based upon their individual univariate prognostic value. The Cox univariate-based method featured in Meyer *et al*. is presented in detail (*see* Note 8) [7].

1. The crude prognostic value of the semi-quantitative score (0–4) captured for each eligible protein was determined using univariate Cox Proportional Hazards.
2. Markers achieving a p-value of 0.05 in the setting of a False Discovery Rate [51] of 0.15 were included in the multi-parameter model. Altogether, 7 markers, were thus included in their final model.
3. The risk score is then calculated by summing up the individual products of the immunostaining score (0–4) and the univariate Cox regression coefficient across all included markers with available data. To account for the situation of each individual, due to missing data, can contribute fewer than 7 markers to the summary score, the raw summary score is normalized through dividing by the number of included markers, according to the following equation (*see* Note 8):

$$Score\,(patient\ X) = \left(\sum_{i=1}^{D} \beta_i x_i\right) / \left(\sum_{i=1}^{D} \alpha_i\right)$$

   where $\alpha_i = 1$ if x_i immunostaining for marker x_i was successful or 0 if the experiment failed and the data is missing.
4. High vs. low-risk scores were determined by splitting the normalized scores at the median.

3.4.2. Genetic Algorithm-derived summary estimates—Genetic algorithms represent an iterative, learning-based approach to model building that exploit the working principles evolution and natural genetics to allow for an efficient, but comprehensive search through a large data space to identify the combination of markers from a larger set of available protein candidates that maximally optimizes prognostic stratification [52]. Beginning with a large panel of candidate proteins with (semi-)quantitative measures of expression, the algorithm arbitrarily selects a set of markers from which to construct a

[7]Univariate estimates can be obtained by modeling fully quantitative data as a continuous parameter or after categorization. Splitting the population, *a priori*, at the median, by tertiles or by quartiles yields unbiased estimates. If semi-quantitative scores are captured, dichotomization according to an *a priori* convention (e.g., <20% cells positive versus ≥20% cells positive) is also acceptable. A third approach of binarizing semi-quantitative or continuous data according to an 'optimal cut-point' that maximizes the survival differential requires special consideration. Although optimal cut-points may reflect non-intuitive but biologically driven boundaries, the volume of multiple comparisons required to search the entire space for an optimal cut-point risks selection of falsely positive values that do not validate in independent populations [78]. Strategies for p-value corrections that account for multiple comparisons such as applying Monte-Carlo simulations [79] should be applied in this setting.

[8]Meyer *et al.* applied several compelling statistical techniques during model discovery. First, rather than equally considering all component markers in creating the summary risk score, markers are weighted according to the magnitude of their Cox univariate regression coefficient. Next, component markers are not binarized; their original semi-quantitative immunostaining scores are used when calculating the summary risk score. Third, they account for the possibility of 'missing data' due to failed laboratory experiments by adjusting the composite scores according to the number of proteins with available data. Finally, when initially triaging among a 70-gene panel of eligible candidates, they corrected for a false discovery rate of 0.15 to avoid promoting falsely positive markers to the next steps of model construction [7].

prognostic model and uses the product-limit log-rank chi-square statistic returned from this initial assessment as a benchmark. Then, the algorithm attempts to improve on this initial model by iteratively altering a single parameter in the model ("generation") either by modifying the cut-point selected to binarize the (semi-) quantitative expression level variable for any individual component candidate ("mutation") or by swapping out one of the included markers for another choice from the larger dataset ("cross-over"). For each new model thusly created, the product-limit log-rank chi-square statistic is calculated and compared to the previous benchmark. If the resulting chi-square statistic exceeds the benchmark, the model is kept to become the new benchmark; otherwise, this model is discarded. The algorithm is iterated until 'convergence' is reached where no new benchmark can be established, typically occurring after over 20 million generations. We employed a genetic algorithm-based approach to generate our 5-marker melanoma prognostic model selected from 20 available candidates [32]. After marker and cut-point selection, binarized markers were scored, by convention, as '1' if they indicated "reduced risk" and '0' if they contrarily indicated "increased risk" with participants tallied according to the sum of "reduced risk" markers (0–5) which was subsequently used to calculate the benchmark log-rank score (*see* Note 9). We seeded our genetic algorithm according to the following parameters:

1. Cross-over rate – the rate at which a swap of markers will be introduced in successive generations, selected at 33%.
2. Mutation rate –the rate at which an alteration in marker cut-points is introduced, selected at 33%.
3. Minimum size of the eligible population – when missing data are present, this sets a lower bound for the number of individuals with complete data across the selected markers for model nomination. To ensure sufficient power for multivariable analyses, we required that a minimum of 100 of the 192 assayed individuals be included each model.
4. Threshold for the minimum number of individuals to be included in each arm of a binarized protein expression marker after cut-point selection. Our choice of 10% helped guard against excessive leverage of extreme observations in defining risk profiles.
5. Threshold for the minimum number of individuals to be included in category of the final ordinal score. Our choice of 15% further reduced the risk of extreme data point leverage and, to further maintain statistical robustness of our final model, we additionally required that no ordinal category enumerate fewer than 2 melanoma-related deaths.
6. The minimum and maximum number of parameters to be included in the final model. To balance maximizing the number of complementary markers with the desire for simplicity and ease of translation, we constrained our model to include at least 3 but no more than 8 eligible protein candidates.

[9]Genetic algorithms are not robust to missing data. Any individual with missing data for one or more selected markers will be excluded each time that protein is incorporated in a proposed model which can introduce bias if missingness is not completely at random [80]. In particular, as has been shown in melanoma [32] and in breast cancer [81], missing data on a TMA is more likely among smaller tumors that are either more difficult to sample or become exhausted among serial cuts such that patients with missing data tend to have better prognosis than those with complete data. One strategy to overcome this potential weakness is to impute values for the missing data to create a complete dataset for all available observations and a recent robust simulation study in breast cancer identified that, among available imputation methods including *mean substitution* and *multiple imputation* where survival time was either omitted (MI−) or included (MI+) in the imputation algorithm, the MI+ calculation yielded the most robust survival estimates [81] and multiple imputation has been used to develop a breast cancer prognostic model derived from QIF data [82].

3.4.3. Classification and Regression Tree (CART)-derived summary estimates—Recursive partitioning is a non-parametric, iterative algorithm that repeatedly splits subsets of the population into two descendant subsets, beginning with the full population itself according to the parameter that best optimizes the desired outcome among the subset X_i [53]. Thus, once the initial population has been split, the optimal variable to subdivide the left branch of the population might differ from the variable selected to further classify the right branch. Nested partitioning continues along each branch until endpoint assessment is no longer statistically feasible, with tree right-sizing accomplished through pruning of terminal branches to optimize experimental feasibility (e.g., no more than 5 protein assays required per terminal branch). Binary, ordinal and continuous independent variables can be included with the CART algorithm selecting the optimal cut-point for non-dichotomous variables that maximizes the outcome separation between the two sub-classes. Regression trees modify the baseline classification tree algorithm initially developed for binary outcomes (e.g., dead/alive) to accommodate time-to-failure data and right-censoring of observations [54].

While no multi-parameter molecular model has, to date, been constructed using this method, Gimotty *et al*. successfully applied the Segal algorithm [55] that employs the two-sample log-rank test statistic to determine the between-node heterogeneity measures to select each partitioning variable to refine prognostic strata among patients with thin melanomas (<1.0 mm) using clinicopathologic criteria [56]. Although overall SEER data from 1988–2001 indicate 97.4% and 90.2% 10-year survival rates among Stage IA and IB patients, respectively, the refined algorithm identified subsets from within each of these populations (e.g., Stage IA Clark level II/III men with lesions of 0.78–1.00 mm thick; 90.6% 10-year survival and Stage IB patients with Clark level IV/V, ulcerated melanomas; 69.8% 10-year survival) with survival estimates outside those generated for the whole population. This classification was subsequently validated on a second independent cohort to serve as a paradigm for a multi-parameter model that integrates clinicopathologic and molecular variables.

3.4.4. Confirmation of the independent prognostic value of the multi-parameter model in multivariable models—The REMARK criteria [5] require adjudication of all candidate molecular prognostic markers in a multivariable Cox Proportional Hazards model that includes all routinely assessed clinicopathologic variables to confirm the marker's independent prognostic value. This necessary step recognizes that these clinicopathologic evaluations reflect the current standard of care and that any molecular test must supplement this standard to justify the assay. For locoregional cutaneous melanoma, Breslow thickness (mm), presence of ulceration, number of mitotic figures/mm^2, presence of microsatellite lesions, sentinel lymph node status, age and gender [1,3] are validated parameters and should be included in the multivariable assessment. Clark level of invasion, while independently prognostic in the large, multi-center validation of both the 2001 and 2009 AJCC staging systems [57,1], is typically collinear with Breslow thickness in smaller populations and cannot be included. Additional clinicopathologic characteristics such as growth phase, histopathologic subtype, anatomic location, tumor vascularity and presence of regression, tumor-infiltrating lymphocytes, solar keratosis, vascular mimicry and lymphovascular invasion have all been associated with melanoma outcome [58–65] are not yet formally validated and their inclusion is optional.

3.4.5. Multi-parameter model internal validation using 10-fold and leave-one-out cross validation—Internal validation experiments assess the model's sensitivity to sample outliers and predict the overall likelihood for generalizability to the general locoregional melanoma population and guard against Type III or Procrustean Errors where the proposed model is suggested by the data due to the leverage of selected extreme data

points [66]. Two common methods for internal cross-validation, 10-fold and leave-one-out cross-validation are presented [7,67].

10-fold Cross-Validation

1. Divide the entire Discovery Cohort into 10 equally-sized groups.
2. Create 'training set #1' by combining groups 1–9, leaving group 10 as the 'test set #1'.
3. Using the raw data for 'training set #1', rebuild the multi-parameter model, following all of the steps exactly.
4. Apply the regression parameters obtained for the 'training set #1' to the 'test set #1' to obtain their predicted score for the prognostic model.
5. Calculate the univariate and multivariable (if sufficiently powered) melanoma-specific survival hazard ratio and 95% confidence interval for the individuals in 'test set #1'.
5. Repeat steps 2–4 nine more times to create training sets 2–10 and test sets 2–10, each time selecting a different group as the test set to obtain 9 more unique sets of hazard ratios and confidence intervals.
6. Calculate an aggregate hazard ratio and 95% confidence interval using the 10 'test set' values according to the general inverse variance method [68] to obtain a cross-validated estimate for the hazard ratio and 95% confidence interval for the proposed multi-parameter model.

Leave-one-out Cross-Validation

1. Starting with the entire Discovery Cohort (total sample size=n), *n* training sets are built that include all but one patient such that each patient is included in its own n=1 'test set' exactly once.
2. For each of the *n* 'training sets', apply the raw data for the included n−1 individuals to rebuild the multi-parameter model accordingly and then calculate the predicted value from the acquired regression parameters for the one individual included in the 'test set'.
3. Combine the predicted values from all *n* 'test set' iterations into a single dataset and use these values to conduct the univariate and Cox multivariable analyses to obtain a cross-validated estimate for the hazard ratio and 95% confidence interval.

3.5. External multi-parameter model validation

Construction concludes with confirmation of model generalizability through demonstrating prognostic relevance in at least one independent *Validation* cohort. This is done in two phases: demonstration of a significant hazard ratio upon multivariable Cox proportional hazards modeling and assessing the model prediction error. The latter step assesses whether the prognostic algorithm can prospectively assign an independent set of new cases (i.e., the *Validation* cohort) into high- vs. low-risk better than pure chance which would have a 50% success rate in randomly categorizing new individuals [67].

Delete

Multivariable Cox proportional hazards modeling

1. A TMA or whole-slide sections are prepared from the primary tissue blocks pertaining to the *Validation* cohort participants.

2. All biomarkers are assayed and scored using the exact methods and criteria applied for model discovery. This can be challenging and methods to standardize scoring between assays over time elapsed between discovery and validation set are critical, but beyond the scope of this chapter.
3. To overcome any potential laboratory drift, all quantitative measures must be normalized to a referent standard to ensure comparability between *Discovery* and *Validation* cohort values.
4. Categorize continuous scores using the same cut-points developed during the model discovery phase. For example, if the *Discovery* cohort was dichotomized at the median of the *Discovery* cohort scores, the *Validation* cohort should be divided at the score corresponding to the *Discovery* cohort median and not at the median *Validation* cohort score.
5. Calculate the prognostic index score for each individual according to the developed schema.
6. Complete a multivariable Cox proportional hazards model evaluating the *Validation* cohort prognostic index scores adjusting for the same conventional clinicopathologic covariates applied in the *Discovery* multivariable Cox model.

Prediction error estimation

1. Calculate the prognostic index score and its 95% confidence interval for each individual using the developed schema.
2. Divide the *Validation* cohort into “high risk” and “low risk” groups according to the index score cut-point selected during model discovery.
3. Binarize the *Validation* cohort based upon their true outcomes at a fixed timepoint (e.g., 5-year recurrence-free survival, 10-year melanoma-specific survival).
4. Create a 2 × 2 cross-tabular display of the data, calculating the sensitivity, specificity, positive predictive value and negative predictive values for the table.

References

1. Balch CM, Gershenwald JE, Soong SJ, Thompson JF, Atkins MB, Byrd DR, Buzaid AC, Cochran AJ, Coit DG, Ding S, Eggermont AM, Flaherty KT, Gimotty PA, Kirkwood JM, McMasters KM, Mihm MC Jr, Morton DL, Ross MI, Sober AJ, Sondak VK. Final version of 2009 AJCC melanoma staging and classification. J Clin Oncol. 2009; 27 (36):6199–6206. [PubMed: 19917835]
2. Petrella T, Verma S, Spithoff K, Quirt I, McCready D. Adjuvant interferon therapy for patients at high risk for recurrent melanoma: an updated systematic review and practice guideline. Clin Oncol (R Coll Radiol). 2012; 24 (6):413–423. [PubMed: 22245520]
3. Thompson JF, Soong SJ, Balch CM, Gershenwald JE, Ding S, Coit DG, Flaherty KT, Gimotty PA, Johnson T, Johnson MM, Leong SP, Ross MI, Byrd DR, Cascinelli N, Cochran AJ, Eggermont AM, McMasters KM, Mihm MC Jr, Morton DL, Sondak VK. Prognostic significance of mitotic rate in localized primary cutaneous melanoma: an analysis of patients in the multi-institutional American Joint Committee on Cancer melanoma staging database. J Clin Oncol. 2011; 29 (16):2199–2205. [PubMed: 21519009]
4. Gould Rothberg BE, Bracken MB, Rimm DL. Tissue biomarkers for prognosis in cutaneous melanoma: a systematic review and meta-analysis. J Natl Cancer Inst. 2009; 101 (7):452–474. [PubMed: 19318635]
5. Altman DG, McShane LM, Sauerbrei W, Taube SE. Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK): explanation and elaboration. PLoS Med. 2012; 9(5):e1001216. PMEDICINE-D-11-01220 [pii]. [PubMed: 22675273]

6. Schramm SJ, Mann GJ. Melanoma prognosis: a REMARK-based systematic review and bioinformatic analysis of immunohistochemical and gene microarray studies. Mol Cancer Ther. 2011; 10 (8):1520–1528. [PubMed: 21659462]
7. Meyer S, Fuchs TJ, Bosserhoff AK, Hofstadter F, Pauer A, Roth V, Buhmann JM, Moll I, Anagnostou N, Brandner JM, Ikenberg K, Moch H, Landthaler M, Vogt T, Wild PJ. A seven-marker signature and clinical outcome in malignant melanoma: a large-scale tissue-microarray study with two independent patient cohorts. PLoS ONE. 2012; 7(6):e38222.10.1371/journal.pone.0038222 [PubMed: 22685558]
8. Vaisanen A, Kuvaja P, Kallioinen M, Turpeenniemi-Hujanen T. A prognostic index in skin melanoma through the combination of matrix metalloproteinase-2, Ki67, and p53. Hum Pathol. 2011; 42 (8):1103–1111. [PubMed: 21334717]
9. Nodin B, Fridberg M, Jonsson L, Bergman J, Uhlen M, Jirstrom K. High MCM3 expression is an independent biomarker of poor prognosis and correlates with reduced RBM3 expression in a prospective cohort of malignant melanoma. Diagn Pathol. 2012; 7:82. 1746-1596-7-82 [pii]. [PubMed: 22805320]
10. Ladstein RG, Bachmann IM, Straume O, Akslen LA. Ki-67 expression is superior to mitotic count and novel proliferation markers PHH3, MCM4 and mitosin as a prognostic factor in thick cutaneous melanoma. BMC Cancer. 2010; 10:140. 1471-2407-10-140 [pii]. 10.1186/1471-2407-10-140 [PubMed: 20398247]
11. Ladstein RG, Bachmann IM, Straume O, Akslen LA. Prognostic importance of the mitotic marker phosphohistone H3 in cutaneous nodular melanoma. J Invest Dermatol. 2012; 132(4):1247–1252. jid2011464 [pii]. 10.1038/jid.2011.464 [PubMed: 22297638]
12. Jonsson L, Bergman J, Nodin B, Manjer J, Ponten F, Uhlen M, Jirstrom K. Low RBM3 protein expression correlates with tumour progression and poor prognosis in malignant melanoma: an analysis of 215 cases from the Malmo Diet and Cancer Study. J Transl Med. 2011; 9:114. 1479-5876-9-114 [pii]. 10.1186/1479-5876-9-114 [PubMed: 21777469]
13. Gremel G, Ryan D, Rafferty M, Lanigan F, Hegarty S, Lavelle M, Murphy I, Unwin L, Joyce C, Faller W, McDermott EW, Sheahan K, Ponten F, Gallagher WM. Functional and prognostic relevance of the homeobox protein MSX2 in malignant melanoma. Br J Cancer. 2011; 105(4): 565–574. bjc2011249 [pii]. 10.1038/bjc.2011.249 [PubMed: 21730974]
14. Zhang Z, Chen G, Cheng Y, Martinka M, Li G. Prognostic significance of RUNX3 expression in human melanoma. Cancer. 2011; 117(12):2719–2727.10.1002/cncr.25838 [PubMed: 21656750]
15. Chen G, Cheng Y, Zhang Z, Martinka M, Li G. Cytoplasmic Skp2 expression is increased in human melanoma and correlated with patient survival. PLoS ONE. 2011; 6(2):e17578.10.1371/journal.pone.0017578 [PubMed: 21386910]
16. Svobodova S, Browning J, MacGregor D, Pollara G, Scolyer RA, Murali R, Thompson JF, Deb S, Azad A, Davis ID, Cebon JS. Cancer-testis antigen expression in primary cutaneous melanoma has independent prognostic value comparable to that of Breslow thickness, ulceration and mitotic rate. Eur J Cancer. 2011; 47(3):460–469. S0959-8049(10)00963-9 [pii]. 10.1016/j.ejca.2010.09.042 [PubMed: 21115342]
17. Jafarnejad SM, Wani AA, Martinka M, Li G. Prognostic significance of Sox4 expression in human cutaneous melanoma and its role in cell migration and invasion. Am J Pathol. 2010; 177(6):2741–2752. S0002-9440(10)62903-3 [pii]. 10.2353/ajpath.2010.100377 [PubMed: 20952589]
18. Li J, Cheng Y, Tai D, Martinka M, Welch DR, Li G. Prognostic significance of BRMS1 expression in human melanoma and its role in tumor angiogenesis. Oncogene. 2011; 30(8):896–906. onc2010470 [pii]. 10.1038/onc.2010.470 [PubMed: 20935672]
19. Rotte A, Martinka M, Li G. MMP2 expression is a prognostic marker for primary melanoma patients. Cell Oncol (Dordr). 2012; 35(3):207–216.10.1007/s13402-012-0080-x [PubMed: 22669775]
20. Brown ER, Doig T, Anderson N, Brenn T, Doherty V, Xu Y, Bartlett JM, Smyth JF, Melton DW. Association of galectin-3 expression with melanoma progression and prognosis. Eur J Cancer. 2012; 48(6):865–874. S0959-8049(11)00714-3 [pii]. 10.1016/j.ejca.2011.09.003 [PubMed: 22071132]

21. Buljan M, Situm M, Tomas D, Milosevic M, Kruslin B. Prognostic value of galectin-3 in primary cutaneous melanoma. J Eur Acad Dermatol Venereol. 2011; 25(10):1174–1181.10.1111/j.1468-3083.2010.03943.x [PubMed: 21175876]
22. Pencina MJ, D'Agostino RB, Pencina KM, Janssens AC, Greenland P. Interpreting incremental value of markers added to risk prediction models. Am J Epidemiol. 2012; 176(6):473–481. kws207 [pii]. 10.1093/aje/kws207 [PubMed: 22875755]
23. Weinlich G, Eisendle K, Hassler E, Baltaci M, Fritsch PO, Zelger B. Metallothionein - overexpression as a highly significant prognostic factor in melanoma: a prospective study on 1270 patients. Br J Cancer. 2006; 94 (6):835–841. [PubMed: 16508630]
24. Weinlich G, Topar G, Eisendle K, Fritsch PO, Zelger B. Comparison of metallothionein-overexpression with sentinel lymph node biopsy as prognostic factors in melanoma. J Eur Acad Dermatol Venereol. 2007; 21 (5):669–677. [PubMed: 17447982]
25. Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, Baehner FL, Walker MG, Watson D, Park T, Hiller W, Fisher ER, Wickerham DL, Bryant J, Wolmark N. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med. 2004; 351 (27):2817–2826. Epub 2004 Dec 2810. [PubMed: 15591335]
26. Cronin M, Sangli C, Liu ML, Pho M, Dutta D, Nguyen A, Jeong J, Wu J, Langone KC, Watson D. Analytical validation of the Oncotype DX genomic diagnostic test for recurrence prognosis and therapeutic response prediction in node-negative, estrogen receptor-positive breast cancer. Clin Chem. 2007; 53(6):1084–1091. clinchem.2006.076497 [pii]. 10.1373/clinchem.2006.076497 [PubMed: 17463177]
27. Buyse M, Loi S, van't Veer L, Viale G, Delorenzi M, Glas AM, d'Assignies MS, Bergh J, Lidereau R, Ellis P, Harris A, Bogaerts J, Therasse P, Floore A, Amakrane M, Piette F, Rutgers E, Sotiriou C, Cardoso F, Piccart MJ. Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. J Natl Cancer Inst. 2006; 98(17):1183–1192. 98/17/1183 [pii]. 10.1093/jnci/djj329 [PubMed: 16954471]
28. Glas AM, Floore A, Delahaye LJ, Witteveen AT, Pover RC, Bakx N, Lahti-Domenici JS, Bruinsma TJ, Warmoes MO, Bernards R, Wessels LF, Van't Veer LJ. Converting a breast cancer microarray signature into a high-throughput diagnostic test. BMC Genomics. 2006; 7:278. 1471-2164-7-278 [pii]. 10.1186/1471-2164-7-278 [PubMed: 17074082]
29. Gray RG, Quirke P, Handley K, Lopatin M, Magill L, Baehner FL, Beaumont C, Clark-Langone KM, Yoshizawa CN, Lee M, Watson D, Shak S, Kerr DJ. Validation study of a quantitative multigene reverse transcriptase-polymerase chain reaction assay for assessment of recurrence risk in patients with stage II colon cancer. J Clin Oncol. 2011; 29(35):4611–4619. JCO.2010.32.8732 [pii]. 10.1200/JCO.2010.32.8732 [PubMed: 22067390]
30. Salazar R, Roepman P, Capella G, Moreno V, Simon I, Dreezen C, Lopez-Doriga A, Santos C, Marijnen C, Westerga J, Bruin S, Kerr D, Kuppen P, van de Velde C, Morreau H, Van Velthuysen L, Glas AM, Van't Veer LJ, Tollenaar R. Gene expression signature to improve prognosis prediction of stage II and III colorectal cancer. J Clin Oncol. 2011; 29(1):17–24. JCO.2010.30.1077 [pii]. 10.1200/JCO.2010.30.1077 [PubMed: 21098318]
31. Bartlett JM, Bloom KJ, Piper T, Lawton TJ, van de Velde CJ, Ross DT, Ring BZ, Seitz RS, Beck RA, Hasenburg A, Kieback D, Putter H, Markopoulos C, Dirix L, Seynaeve C, Rea D. Mammostrat as an Immunohistochemical Multigene Assay for Prediction of Early Relapse Risk in the Tamoxifen Versus Exemestane Adjuvant Multicenter Trial Pathology Study. J Clin Oncol. 2012 JCO.2012.42.8896 [pii]. 10.1200/JCO.2012.42.8896
32. Gould Rothberg BE, Berger AJ, Molinaro AM, Subtil A, Krauthammer MO, Camp RL, Bradley WR, Ariyan S, Kluger HM, Rimm DL. A melanoma prognostic model using tissue microarrays and genetic algorithms. J Clin Oncol. 2009; 27 (34):5772–5780. [PubMed: 19884546]
33. Kashani-Sabet M, Venna S, Nosrati M, Rangel J, Sucker A, Egberts F, Baehner FL, Simko J, Leong SP, Haqq C, Hauschild A, Schadendorf D, Miller JR 3rd, Sagebiel RW. A multimarker prognostic assay for primary cutaneous melanoma. Clin Cancer Res. 2009; 15(22):6987–6992. 1078-0432.CCR-09-1777 [pii]. 10.1158/1078-0432.CCR-09-1777 [PubMed: 19887476]
34. Piras F, Perra MT, Murtas D, Minerba L, Floris C, Maxia C, Demurtas P, Ugalde J, Ribatti D, Sirigu P. Combinations of apoptosis and cell-cycle control biomarkers predict the outcome of human melanoma. Oncol Rep. 2008; 20 (2):271–277. [PubMed: 18636186]

35. Kononen J, Bubendorf L, Kallioniemi A, Barlund M, Schraml P, Leighton S, Torhorst J, Mihatsch MJ, Sauter G, Kallioniemi OP. Tissue microarrays for high-throughput molecular profiling of tumor specimens. Nat Med. 1998; 4 (7):844–847. [PubMed: 9662379]
36. Moeder CB, Giltnane JM, Moulis SP, Rimm DL. Quantitative, fluorescence-based in-situ assessment of protein expression. Methods Mol Biol. 2009; 520:163–175. [PubMed: 19381954]
37. Rimm DL, Camp RL, Charette LA, Olsen DA, Provost E. Amplification of tissue by construction of tissue microarrays. Exp Mol Pathol. 2001; 70 (3):255–264. [PubMed: 11418004]
38. DiVito KA, Charette LA, Rimm DL, Camp RL. Long-term preservation of antigenicity on tissue microarrays. Lab Invest. 2004; 84 (8):1071–1078. [PubMed: 15195116]
39. Brennan DJ, Rexhepaj E, O'Brien SL, McSherry E, O'Connor DP, Fagan A, Culhane AC, Higgins DG, Jirstrom K, Millikan RC, Landberg G, Duffy MJ, Hewitt SM, Gallagher WM. Altered cytoplasmic-to-nuclear ratio of survivin is a prognostic indicator in breast cancer. Clin Cancer Res. 2008; 14(9):2681–2689. 14/9/2681 [pii]. 10.1158/1078-0432.CCR-07-1760 [PubMed: 18451232]
40. Rexhepaj E, Brennan DJ, Holloway P, Kay EW, McCann AH, Landberg G, Duffy MJ, Jirstrom K, Gallagher WM. Novel image analysis approach for quantifying expression of nuclear proteins assessed by immunohistochemistry: application to measurement of oestrogen and progesterone receptor levels in breast cancer. Breast Cancer Res. 2008; 10(5):R89. bcr2187 [pii]. 10.1186/bcr2187 [PubMed: 18947395]
41. Lloyd MC, Allam-Nandyala P, Purohit CN, Burke N, Coppola D, Bui MM. Using image analysis as a tool for assessment of prognostic and predictive biomarkers for breast cancer: How reliable is it? J Pathol Inform. 2010; 1:29.10.4103/2153-3539.74186 [PubMed: 21221174]
42. Camp RL, Chung GG, Rimm DL. Automated subcellular localization and quantification of protein expression in tissue microarrays. Nat Med. 2002; 8 (11):1323–1327. Epub 2002 Oct 1321. [PubMed: 12389040]
43. Neumeister V, Agarwal S, Bordeaux J, Camp RL, Rimm DL. In situ identification of putative cancer stem cells by multiplexing ALDH1, CD44, and cytokeratin identifies breast cancer patients with poor prognosis. Am J Pathol. 2010; 176(5):2131–2138. S0002-9440(10)60010-7 [pii]. 10.2353/ajpath.2010.090712 [PubMed: 20228222]
44. Goens G, Rusu D, Bultot L, Goval JJ, Magdalena J. Characterization and quality control of antibodies used in ChIP assays. Methods Mol Biol. 2009; 567:27–43.10.1007/978-1-60327-414-2_2 [PubMed: 19588083]
45. Larsson PH. Purification of antibodies. Methods Mol Med. 2008; 138:197–207.10.1007/978-1-59745-366-0_16 [PubMed: 18612609]
46. Bordeaux J, Welsh A, Agarwal S, Killiam E, Baquero M, Hanna J, Anagnostou V, Rimm D. Antibody validation. Biotechniques. 2010; 48(3):197–209. 000113382 [pii]. 10.2144/000113382 [PubMed: 20359301]
47. Welsh AW, Moeder CB, Kumar S, Gershkovich P, Alarid ET, Harigopal M, Haffty BG, Rimm DL. Standardization of estrogen receptor measurement in breast cancer suggests false-negative results are a function of threshold intensity rather than percentage of positive cells. J Clin Oncol. 2011; 29(22):2978–2984. JCO.2010.32.9706 [pii]. 10.1200/JCO.2010.32.9706 [PubMed: 21709197]
48. Gustavson, M.; Dolled-Filhart, M.; Christiansen, J.; Pinard, R.; Rimm, D. AQUA technology and molecular pathology. In: Platero, JS., editor. Molecular Pathology in Drug Discovery and Development. Wiley; Hoboken, NJ: 2009. p. 295-333.
49. Gustavson MD, Bourke-Martin B, Reilly DM, Cregger M, Williams C, Tedeschi G, Pinard R, Christiansen J. Development of an unsupervised pixel-based clustering algorithm for compartmentalization of immunohistochemical expression using Automated QUantitative Analysis. Appl Immunohistochem Mol Morphol. 2009; 17(4):329–337.10.1097/PAI.0b013e318195ecaa [PubMed: 19318915]
50. Gustavson MD, Molinaro AM, Tedeschi G, Camp RL, Rimm DL. AQUA analysis of thymidylate synthase reveals localization to be a key prognostic biomarker in 2 large cohorts of colorectal carcinoma. Arch Pathol Lab Med. 2008; 132(11):1746–1752. 2007-0718-OA [pii]. 10.1043/1543-2165-132.11.1746 [PubMed: 18976010]

51. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Statist Soc B. 1995; 57 (1):289–200.
52. Ooi CH, Tan P. Genetic algorithms applied to multi-class prediction for the analysis of gene expression data. Bioinformatics. 2003; 19 (1):37–44. [PubMed: 12499291]
53. Breiman, L.; Friedman, JH.; Olshen, RA.; Stone, CJ. Classification and Regression Trees. CRC Press; Boca Raton, FL: 1984.
54. Molinaro AM, Dudoit S, van der Laan MJ. Tree-based multivariate regression and density estimation with right-censored data. Journal of Multivariate Analysis. 2004; 90:154–177.
55. Segal MR, Bloch DA. A comparison of estimated proportional hazards models and regression trees. Stat Med. 1989; 8 (5):539–550. [PubMed: 2727474]
56. Gimotty PA, Elder DE, Fraker DL, Botbyl J, Sellers K, Elenitsas R, Ming ME, Schuchter L, Spitz FR, Czerniecki BJ, Guerry D. Identification of high-risk patients among those diagnosed with thin cutaneous melanomas. J Clin Oncol. 2007; 25 (9):1129–1134. [PubMed: 17369575]
57. Balch CM, Buzaid AC, Soong SJ, Atkins MB, Cascinelli N, Coit DG, Fleming ID, Gershenwald JE, Houghton A Jr, Kirkwood JM, McMasters KM, Mihm MF, Morton DL, Reintgen DS, Ross MI, Sober A, Thompson JA, Thompson JF. Final version of the American Joint Committee on Cancer staging system for cutaneous melanoma. J Clin Oncol. 2001; 19 (16):3635–3648. [PubMed: 11504745]
58. Massi D, Franchi A, Borgognoni L, Paglierani M, Reali UM, Santucci M. Tumor angiogenesis as a prognostic factor in thick cutaneous malignant melanoma. A quantitative morphologic analysis. Virchows Arch. 2002; 440 (1):22–28. [PubMed: 11942572]
59. de Giorgi V, Rossari S, Gori A, Grazzini M, Savarese I, Crocetti E, Cervadoro E, Massi D. The prognostic impact of the anatomical sites in the 'head and neck melanoma': scalp versus face and neck. Melanoma Res. 2012; 22(5):402–405.10.1097/CMR.0b013e3283577b96 [PubMed: 22922466]
60. Thies A, Mangold U, Moll I, Schumacher U. PAS-positive loops and networks as a prognostic indicator in cutaneous malignant melanoma. J Pathol. 2001; 195(5):537–542. [pii]. 10.1002/path.988 [PubMed: 11745688]
61. Gimotty PA, Guerry D. Prognostication in thin cutaneous melanomas. Arch Pathol Lab Med. 2010; 134(12):1758–1763. [pii]. 10.1043/2009-0653-RAR.1 [PubMed: 21128772]
62. Lindholm C, Andersson R, Dufmats M, Hansson J, Ingvar C, Moller T, Sjodin H, Stierner U, Wagenius G. Invasive cutaneous malignant melanoma in Sweden, 1990–1999. A prospective, population-based study of survival and prognostic factors. Cancer. 2004; 101(9):2067–2078.10.1002/cncr.20602 [PubMed: 15372475]
63. Storr SJ, Safuan S, Mitra A, Elliott F, Walker C, Vasko MJ, Ho B, Cook M, Mohammed RA, Patel PM, Ellis IO, Newton-Bishop JA, Martin SG. Objective assessment of blood and lymphatic vessel invasion and association with macrophage infiltration in cutaneous melanoma. Mod Pathol. 2012; 25(4):493–504. modpathol2011182 [pii]. 10.1038/modpathol.2011.182 [PubMed: 22080065]
64. Xu X, Chen L, Guerry D, Dawson PR, Hwang WT, VanBelle P, Elder DE, Zhang PJ, Ming ME, Schuchter L, Gimotty PA. Lymphatic invasion is independently prognostic of metastasis in primary cutaneous melanoma. Clin Cancer Res. 2012; 18(1):229–237. 1078-0432.CCR-11-0490 [pii]. 10.1158/1078-0432.CCR-11-0490 [PubMed: 22096024]
65. Yun SJ, Gimotty PA, Hwang WT, Dawson P, Van Belle P, Elder DE, Elenitsas R, Schuchter L, Zhang PJ, Guerry D, Xu X. High lymphatic vessel density and lymphatic invasion underlie the adverse prognostic effect of radial growth phase regression in melanoma. Am J Surg Pathol. 2011; 35(2):235–242. 00000478-201102000-00008 [pii]. 10.1097/PAS.0b013e3182036ccd [PubMed: 21263244]
66. Mills JL. Data torturing. N Engl J Med. 1993; 329 (16):1196–1199. [PubMed: 8166792]
67. Molinaro AM, Simon R, Pfeiffer RM. Prediction error estimation: a comparison of resampling methods. Bioinformatics. 2005; 21 (15):3301–3307. Epub 2005 May 3319. [PubMed: 15905277]
68. Deeks, JJ.; Altman, DG.; Bradburn, MJ. Chapter 15: Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis. In: Egger, M.; Smith, GD.; Altman, DG., editors. Systematic reviews in health care: Meta-analysis in context. 2. BMJ Press; Cornwall, UK: 2001. p. 285-312.

69. Gould Rothberg BE, Bracken MB. E-cadherin immunohistochemical expression as a prognostic factor in infiltrating ductal carcinoma of the breast: a systematic review and meta-analysis. Breast Cancer Res Treat. 2006; 100(2):139–148.10.1007/s10549-006-9248-2 [PubMed: 16791476]
70. Rothman, KJ.; Greenland, S. Chapter 7: Cohort studies. In: Rothman, KJ.; Greenland, S.; Lash, TL., editors. Modern Epidemiology. 3. Lippincott Williams and Wilkins; Philadelphia, PA: 2008.
71. Simon RM, Paik S, Hayes DF. Use of archived specimens in evaluation of prognostic and predictive biomarkers. J Natl Cancer Inst. 2009; 101(21):1446–1452. djp335 [pii]. 10.1093/jnci/djp335 [PubMed: 19815849]
72. Tolles J, Bai Y, Baquero M, Harris LN, Rimm DL, Molinaro AM. Optimal tumor sampling for immunostaining of biomarkers in breast carcinoma. Breast Cancer Res. 2011; 13(3):R51. bcr2882 [pii]. 10.1186/bcr2882 [PubMed: 21592345]
73. Pacifico MD, Grover R, Richman P, Daley F, Wilson GD. Validation of tissue microarray for the immunohistochemical profiling of melanoma. Melanoma Res. 2004; 14(1):39–42. 00008390-200402000-00006 [pii]. [PubMed: 15091192]
74. Pacifico MD, Grover R, Richman PI, Buffa F, Daley FM, Wilson GD. Identification of P-cadherin in primary melanoma using a tissue microarrayer: prognostic implications in a patient cohort with long-term follow up. Ann Plast Surg. 2005; 55 (3):316–320. [PubMed: 16106174]
75. Pearl RA, Pacifico MD, Richman PI, Wilson GD, Grover R. Stratification of patients by melanoma cell adhesion molecule (MCAM) expression on the basis of risk: implications for sentinel lymph node biopsy. J Plast Reconstr Aesthet Surg. 2008; 61:265–271. [PubMed: 17569608]
76. De Jong AS, Van Kessel-van Vark M, Raap AK. Sensitivity of various visualization methods for peroxidase and alkaline phosphatase activity in immunoenzyme histochemistry. Histochem J. 1985; 17 (10):1119–1130. [PubMed: 2416713]
77. Harlow E, Lane D. Detection of horseradish peroxidase-labeled reagents with aminoethylcarbazole. CSH Protoc. 2006; 2006(1) 2006/1/pdb.prot4334 [pii]. 10.1101/pdb.prot4334
78. Altman DG, Lausen B, Sauerbrei W, Schumacher M. Dangers of using “optimal” cutpoints in the evaluation of prognostic factors. J Natl Cancer Inst. 1994; 86 (11):829–835. [PubMed: 8182763]
79. Camp RL, Dolled-Filhart M, Rimm DL. X-tile: a new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. Clin Cancer Res. 2004; 10 (21):7252–7259. [PubMed: 15534099]
80. Greenland S, Finkle WD. A critical look at methods for handling missing covariates in epidemiologic regression analyses. Am J Epidemiol. 1995; 142 (12):1255–1264. [PubMed: 7503045]
81. Ali AM, Dawson SJ, Blows FM, Provenzano E, Ellis IO, Baglietto L, Huntsman D, Caldas C, Pharoah PD. Comparison of methods for handling missing data on immunohistochemical markers in survival analysis of breast cancer. Br J Cancer. 2011; 104(4):693–699. 6606078 [pii]. 10.1038/sj.bjc.6606078 [PubMed: 21266980]
82. Emerson JW, Dolled-Filhart M, Harris L, Rimm DL, Tuck DP. Quantitative assessment of tissue biomarkers and construction of a model to predict outcome in breast cancer using multiple imputation. Cancer Inform. 2009; 7:29–40. [PubMed: 19352457]
83. Alonso SR, Ortiz P, Pollan M, Perez-Gomez B, Sanchez L, Acuna MJ, Pajares R, Martinez-Tello FJ, Hortelano CM, Piris MA, Rodriguez-Peralto JL. Progression in cutaneous malignant melanoma is associated with distinct expression profiles: a tissue microarray-based study. Am J Pathol. 2004; 164 (1):193–203. [PubMed: 14695333]
84. Vaisanen AH, Kallioinen M, Turpeenniemi-Hujanen T. Comparison of the prognostic value of matrix metalloproteinases 2 and 9 in cutaneous melanoma. Hum Pathol. 2008; 39:377–385. [PubMed: 18187184]
85. Pacifico MD, Grover R, Richman PI, Daley FM, Buffa F, Wilson GD. Development of a tissue array for primary melanoma with long-term follow-up: discovering melanoma cell adhesion molecule as an important prognostic marker. Plast Reconstr Surg. 2005; 115 (2):367–375. [PubMed: 15692338]

86. Lin H, Wong RP, Martinka M, Li G. Loss of SNF5 expression correlates with poor patient survival in melanoma. Clin Cancer Res. 2009; 15(20):6404–6411. 1078-0432.CCR-09-1135 [pii]. 10.1158/1078-0432.CCR-09-1135 [PubMed: 19808872]

87. Ekmekcioglu S, Ellerhorst JA, Prieto VG, Johnson MM, Broemeling LD, Grimm EA. Tumor iNOS predicts poor survival for stage III melanoma patients. Int J Cancer. 2006; 119 (4):861–866. [PubMed: 16557582]

88. Straume O, Sviland L, Akslen LA. Loss of nuclear p16 protein expression correlates with increased tumor cell proliferation (Ki-67) and poor prognosis in patients with vertical growth phase melanoma. Clin Cancer Res. 2000; 6 (5):1845–1853. [PubMed: 10815907]

89. Thies A, Moll I, Berger J, Wagener C, Brummer J, Schulze HJ, Brunner G, Schumacher U. CEACAM1 expression in cutaneous malignant melanoma predicts the development of metastatic disease. J Clin Oncol. 2002; 20 (10):2530–2536. [PubMed: 12011132]

90. Florenes VA, Maelandsmo GM, Faye R, Nesland JM, Holm R. Cyclin A expression in superficial spreading malignant melanomas correlates with clinical outcome. J Pathol. 2001; 195 (5):530–536. [PubMed: 11745687]

91. Thies A, Schachner M, Moll I, Berger J, Schulze HJ, Brunner G, Schumacher U. Overexpression of the cell adhesion molecule L1 is associated with metastasis in cutaneous malignant melanoma. Eur J Cancer. 2002; 38 (13):1708–1716. [PubMed: 12175686]

92. Soltani MH, Pichardo R, Song Z, Sangha N, Camacho F, Satyamoorthy K, Sangueza OP, Setaluri V. Microtubule-associated protein 2, a marker of neuronal differentiation, induces mitotic defects, inhibits growth of melanoma cells, and predicts metastatic potential of cutaneous melanoma. Am J Pathol. 2005; 166 (6):1841–1850. [PubMed: 15920168]

93. Niezabitowski A, Czajecki K, Rys J, Kruczak A, Gruchala A, Wasilewska A, Lackowska B, Sokolowski A, Szklarski W. Prognostic evaluation of cutaneous malignant melanoma: a clinicopathologic and immunohistochemical study. J Surg Oncol. 1999; 70 (3):150–160. [PubMed: 10102344]

94. Piras F, Murtas D, Minerba L, Ugalde J, Floris C, Maxia C, Colombari R, Perra MT, Sirigu P. Nuclear survivin is associated with disease recurrence and poor survival in patients with cutaneous malignant melanoma. Histopathology. 2007; 50 (7):835–842. [PubMed: 17543072]

95. Tran TA, Ross JS, Carlson JA, Mihm MC Jr. Mitotic cyclins and cyclin-dependent kinases in melanocytic lesions. Hum Pathol. 1998; 29 (10):1085–1090. [PubMed: 9781646]

96. Florenes VA, Faye RS, Maelandsmo GM, Nesland JM, Holm R. Levels of cyclin D1 and D3 in malignant melanoma: deregulated cyclin D3 expression is associated with poor clinical outcome in superficial melanoma. Clin Cancer Res. 2000; 6 (9):3614–3620. [PubMed: 10999753]

97. McDermott NC, Milburn C, Curran B, Kay EW, Barry Walsh C, Leader MB. Immunohistochemical expression of nm23 in primary invasive malignant melanoma is predictive of survival outcome. J Pathol. 2000; 190 (2):157–162. [PubMed: 10657013]

98. Pacifico MD, Grover R, Richman PI, Buffa F, Daley FM, Wilson GD. nm23 as a prognostic marker in primary cutaneous melanoma: evaluation using tissue microarray in a patient group with long-term follow-up. Melanoma Res. 2005; 15 (5):435–440. [PubMed: 16179871]

99. Li Q, Murphy M, Ross J, Sheehan C, Carlson JA. Skp2 and p27kip1 expression in melanocytic nevi and melanoma: an inverse relationship. J Cutan Pathol. 2004; 31 (10):633–642. [PubMed: 15491322]

Table 1

Significant, multivariable-adjusted individual protein melanoma-related prognostic hazard ratios ≥ 3.0 (≤ 0.33)

Protein	Total n	Reference group	Hazard Ratio (95% Confidence interval)	p-value	References
Overall survival					
Bcl-xL	60	<10% cells positive	8.07 (1.77 – 36.89)	p=0.007	[83]
Bcl-6	60	No immunostain	3.98 (1.37 – 11.60)	p=0.01	[83]
p16/INK4A	60	<50% cells positive	0.29 (0.10 – 0.83)	p=0.02	[83]
p27/KIP1	60	<10% cells positive	3.08 (1.20 – 7.91)	p=0.02	[83]
Matrix metalloproteinase-2	50	<34% cells positive	4.5 (1.5 – 13.0)	p=0.006	[84]
Melanoma cell adhesion molecule/MUC18	76	No immunostain	16.34 (3.80 – 70.28)	p<0.001	[85]
SNF5	88	No immunostain	5.15 (1.48 – 17.89)	p=0.01	[86]
Melanoma-specific survival					
iNOS	132	<5% cells positive	4.63 (2.60 – 8.25)	p<0.001	[87]
Ki-67	187	<16% cells positive	3.7 (1.6 – 8.9)	p=0.003	[88]
MCM3	255	Nuclear H score ≤3 (of 9)	4.96 (1.77 – 13.87)	p=0.002	[9]
Metallothionein	1428	<10% cells positive	3.08 (2.02 – 4.68)	p<0.001	[23,24]
p53	187	No immunostain	8.9 (2.7 – 29.0)	p<0.001	[88]
SNF5	88	No immunostain	4.64 (1.15 – 18.63)	p=0.03	[86]
Disease-free survival					
CEACAM-1	100	<20% cells positive	7.17 (3.22 – 15.95)	p<0.001	[89]
Cyclin A	172	<5% cells positive	3.7 (3.4 – 4.1)	p=0.001	[90]
L1-CAM	100	<20% cells positive	4.38 (2.08 – 9.23)	p<0.001	[91]
Melanoma cell adhesion molecule/MUC18	76	No immunostain	14.83 (5.20 – 42.24)	p=0.01	[75]
Metallothionein	1428	<10% cells positive	3.77 (2.73 – 5.22)	p<0.001	[23,24]
Microtubule-associated protein-2	37	<70% cells positive	0.18 (0.06 – 0.56)	p=0.003	[92]
Proliferating cell nuclear antigen	93	≤15% cells positive	4.00 (2.05 – 7.81)	p=0.039	[93]
Survivin	50	No nuclear immunostain	7.32 (1.43 – 37.38)	p=0.017	[94]

Methods Mol Biol. Author manuscript; available in PMC 2014 February 04.

Table 2

Individual proteins with multivariable hazard ratios independently assayed across multiple cohorts

Protein	Study	Total n	Reference group	Hazard Ratio (95% Confidence interval)	p-value	References
Overall survival						
Cyclin A						
	Tran 1998	66	<5% cells positive	5.00 (0.56 – 44.69)	p=0.15	[95]
	Florenes 2001	172	<5% cells positive	0.80 (0.40 – 1.60)	p=0.43	[90]
	Alonso 2004	60	<5% cells positive	0.76 (0.27 – 2.14)	p=0.60	[83]
Ki-67						
	Niezabitowski 1999	93	<20% cells positive	5.17 (2.17 – 12.29)	p=0.0002	[93]
	Alonso 2004	60	<20% cells positive	1.24 (0.49 – 3.14)	p=0.65	[83]
Cyclin D3						
	Florenes 2000	172	<5% cells positive	1.00 (0.60 – 1.67)	p>0.99	[96]
	Alonso 2004	59	<5% cells positive	3.63 (0.81 – 16.27)	p=0.09	[83]
Cyclin D1						
	Florenes 2000	172	<5% cells positive	1.00 (0.40 – 2.50)	p>0.99	[96]
	Alonso 2004	60	<5% cells positive	1.22 (0.35 – 4.25)	p=0.75	[83]
nm23						
	McDermott 2000	145	Weak/moderate immunostain	0.79 (0.41 – 1.51)	p=0.48	[97]
	Pacifico 2005a	76	Weak/moderate immunostain	0.44 (0.10 – 1.99)	p=0.29	[98]
skp2						
	Alonso 2004	59	<5% cells positive	1.04 (0.38 – 2.85)	p=0.94	[83]
	Li 2004	104	<5% cells positive	1.07 (0.55 – 2.08)	p=0.84	[99]
Melanoma-specific survival						
Metallothionein						
	Weinlich 2006	1270	<10% cells positive	3.49 (2.25 – 5.41)	p<0.0001	[23]
	Weinlich 2007	158	<10% cells positive	0.85 (0.21 – 3.45)	p=0.82	[24]
Matrix metalloproteinase-2						
	Vaisanen 2008	157	≤20% cells positive	2.60 (1.32 – 5.07)	p=0.006	[84]
	Rotte 2012	330	≤Moderate immunostaining	2.38 (1.39 – 4.09)	p=0.002	[19]
p16/INK4A						
	Straume 2000	187	Weak stain	0.4 (0.24 – 0.67)	p=0.007	[88]

Protein	Study	Total n	Reference group	Hazard Ratio (95% Confidence interval)	p-value	References
	Gould Rothberg 2009	187	Weak nuclear immunostain (1st quartile AQUA score)	2nd quartile: 0.46 (0.27 – 0.88) 3rd quartile: 0.42 (0.22 – 0.81) 4th quartile: 0.60 (0.32 – 1.15)	p=0.04	[32]
Galectin-3						
	Bulijan 2011	104	Weak/moderate immunostain	3.54 (0.57 – 22.00)	p=0.18	[21]
	Brown 2012	314	Weak/moderate nuclear immunostain	0.74 (0.55–1.00)	p=0.06	[20]
Disease-free survival						
Metallothionein						
	Weinlich 2006	1270	<10% cells positive	3.94 (2.77 – 5.60)	p<0.0001	[23]
	Weinlich 2007	158	<10% cells positive	2.98 (1.31 – 6.78)	p=0.009	[24]

Table 3

Published melanoma multi-parameter prognostic models

Study	Assay method	Parameter triage method	Selected markers and cut-points	Prognostic index categories	Discovery cohort Hazard ratios (95% CI)	p-value	Validation cohort Hazard ratio (95% CI)	p-value
Piras 2008 [34]	Qualitative chromogenic immunostaining	Significance in univariate log-rank analysis	Total p53: any + Nuclear p16: any + Nuclear survivin: any +	0–1 altered 2–3 altered	1.00 4.77 (2.14 – 10.59)	p<0.0001	NA	-
Kashani-Sabet 2009 [33]	Semi-quantitative chromogenic immunostaining	Independent significance in multivariable models	Osteopontin[1] NCOA3/AIB-1 RGS1	Score of −3 to 0 Score of 1 to 3	1.00 1.29 (1.19 – 1.40)	p=0.002	1.00 1.34 (1.22 – 1.47)	p=0.001
Gould Rothberg 2009 [32]	Quantitative immunofluorescence	Genetic Algorithms	ln(ATF2 ratio)> −0.052 β-catenin>38.68 Fibronectin≤57.93 ln(p16 ratio)≤ −0.083 Nuclear p21>12.98	≤3 markers + 4–5 markers +	1.00 2.84 (1.46 – 5.49)	p=0.002	1.00 2.72 (1.12 – 6.58)	p=0.027
Vaisanen 2011 [8]	Semi-quantitative chromogenic immunostaining	Significance in univariate log-rank analysis	Ki-67: >40% cells + p53: any + MMP-2: >20% cells +	0 altered 1 altered 2 altered 3 altered	NA[2]	-	NA	-
Meyer 2012 [7]	Semi-quantitative chromogenic immunostaining	Significance in univariate log-rank analysis	Bax[3] Bcl-X β-catenin CD20 Cox-2 MTAP PTEN	0^{th}–50^{th} %ile 51^{st}–100^{th} %ile	1.00 5.1 (1.4 – 18.2)	p=0.012	1.00 14.45 (1.68 – 124.49)	p=0.015

[1] Selected cut-points were not defined in the published manuscript

[2] The multi-parameter assay was developed but only assessed as a univariate parameter. The reported multivariable Cox proportional hazards model simultaneously included MMP-2 and Ki-67 as independent parameters, which were both independently significant.

[3] Cut-points were not used. Each marker was scored as a 5-level semi-quantitative variable (0–4) and this value was used to calculate the summary score.