
Article

Not peer-reviewed version

The level of agreement between self-assessments and expert assessments of melanocytic nevus counts: Findings from an evaluation of 4548 blinded double assessments

[Olaf Gefeller](#) * , [Isabelle Kaiser](#) , Emily M. Brockmann , [Wolfgang Uter](#) , [Annette B. Pfahlberg](#)

Posted Date: 28 February 2024

doi: [10.20944/preprints202402.1639.v1](https://doi.org/10.20944/preprints202402.1639.v1)

Keywords: cutaneous melanoma; screening; melanocytic nevus; self-counting; eye color



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

The Level of Agreement between Self-Assessments and Expert Assessments of Melanocytic Nevus Counts: Findings from an Evaluation of 4548 Blinded Double Assessments

Olaf Gefeller ^{*}, Isabelle Kaiser, Emily M. Brockmann, Wolfgang Uter and Annette B. Pfahlberg

Department of Medical Informatics, Biometry and Epidemiology, Friedrich-Alexander-Universität Erlangen-Nürnberg, 91054 Erlangen, Germany

^{*} Correspondence: olaf.gefeller@fau.de; Tel.: 49-9131-852-2750

Abstract: Cutaneous melanoma (CM) is a candidate for screening programs, since its prognosis is excellent when diagnosed in an early disease stage. Targeted screening of those at high risk for developing CM, a cost-effective alternative to population-wide screening, requires valid procedures to identify the high-risk group. Self-assessment of the number of nevi has been suggested as a component of such procedures, but its validity has not yet been established. We analyzed the level of agreement between self-assessments and expert assessments of the number of melanocytic nevi in the area between the wrist and the shoulder of both arms based on 4548 study subjects in whom mutually blinded double counting of nevi was performed. Nevus counting followed the IARC protocol. Study subjects received written instructions, photographs, a mirror, and a "nevometer" to support self-assessment of nevi larger than 2 mm. Nevus counts were categorized based on the quintiles of the distribution into five levels defining a nevus score. Cohen's weighted kappa-coefficient was estimated to measure the level of agreement. In the total sample, agreement between self-assessments and expert-assessments was moderate (raw agreement: 50.29%, weighted $\kappa=0.596$). Self-assessed nevus counts were higher than those determined by experts (mean difference: 3.33 nevi). The level of agreement was independent of sociodemographic and cutaneous factors; however, participants' eye color had a significant impact on the level of agreement. Our findings show that even with comprehensive guidance, only a moderate level of agreement between self-assessed and expert-assessed nevus counts can be achieved. Self-assessed nevus information does not appear to be reliable enough to be used in individual risk assessment to target screening activities.

Keywords: cutaneous melanoma; screening; melanocytic nevus; self-counting; eye color

1. Introduction

Incidence rates of cutaneous melanoma (CM) have shown a dynamic development in countries with predominantly fair-skinned populations in recent decades [1]. In Scandinavian countries and New Zealand, where population-based cancer registries have documented trends over a long period since the 1940s, incidence rates have increased 30- to 40-fold over the last 70 years, with annual increases of 3-7% [1-3]. CM, which was once a rare form of skin cancer until the middle of the last century, is currently an important public health issue in many countries worldwide.

CM is an obvious candidate for screening programs, since its prognosis is excellent when it is diagnosed in an early stage of the disease [4-6]. Current evidence on the benefits of CM screening, summarized for the US Preventive Services Task Force in a recent systematic review covering 20 studies [7], is, however, inconsistent. Due to the high financial costs of population-wide screening for CM and the potential for overdiagnosis, i.e. the identification of otherwise indolent melanoma or false-positive melanoma diagnoses that turn out to be benign lesions [8], most countries have not

issued a recommendation regarding CM screening thus far [9]. Only two countries have established national guidance on CM screening by physicians. Germany recommends visual skin examinations by dermatologists or general practitioners, who have been certified for performing skin cancer screening. Examinations can take place every two years (beginning at age 35) and are covered by statutory health insurance [10]. France recommends that general practitioners assess the CM risk of their patients via a checklist consisting of seven questions and refer high-risk patients for further diagnostic clarification to a dermatologist [11].

An alternative to population-wide screening is targeted screening of those at high risk for developing CM [12]. An efficient implementation of such a strategy requires that the CM risk can be assessed simply, reliably and validly. To this end, the number of melanocytic nevi has been consistently identified as a strong predictor of CM [13,14]. Nearly all prediction models developed to date to capture CM risk have incorporated the number of nevi, mostly in a categorized form, among the explanatory variables [15]. A prerequisite for using the number of nevi in order to identify individuals at high CM risk for targeted screening is, however, that this variable can be validly self-assessed. Otherwise physicians would need to be involved in identifying the high-risk group for actual screening, and potential savings compared to population-wide screening could not be achieved.

By addressing the aspect of validity when using self-assessment to identify the number of melanocytic nevi, the objective of our study was to analyze the level of agreement between self-assessment and expert assessment of nevus counts based on a large dataset of study subjects in whom mutually blinded double counting of nevi was performed. We report on the quantitative degree of agreement between the two assessments and on potential determinants of the agreement.

2. Materials and Methods

2.1. Design and Participants

The study started in April 2006 as a repeated cross-sectional survey and was conducted in an identical way twice a year (in April and November) until November 2019 at the University of Erlangen-Nuremberg. Every summer and winter semester, all students enrolled in the course 'Biometry and Epidemiology' were invited to participate in the study. Of the 5125 students enrolled during the 13.5-year study period, 4768 students attended the course on the day when the study took place. Of these, 4704 (98.7%) agreed to participate and provided written informed consent. The study was approved by the Ethics Committee of the University Erlangen-Nuremberg.

2.2. Nevus Counting

The criteria for the nevus counting procedure followed the IARC protocol for identifying and recording nevi [16]. Accordingly, all common acquired melanocytic nevi with a diameter of at least 2 mm were eligible for counting. For feasibility reasons, nevus counting was limited to the area between the wrist and the shoulder, as nevus counts on the arms are highly correlated with total body counts [17]. The number of melanocytic nevi on both arms was counted twice, once by the participant and once by one of six specifically trained examiners involved in the study (one dermatologist, two physicians of other disciplines, and three academic researchers who were instructed by the dermatologist), in a mutually blinded manner: participants were not aware of the counting result of the examiner when performing the nevus counting and vice versa (for logistical reasons, the sequence of nevus counting was not identical for all participants, in one subgroup examiners counted first, in the other subgroup examiners counted last).

Participants received detailed written instructions and photographs of various skin lesions to clarify the counting procedure and appearance of melanocytic nevi and to distinguish them from other skin lesions such as café au lait spots, Becker's nevus, viral warts, freckles, hemangiomas, dermatofibromas, and other potentially similar lesions. Additionally, they were given a mirror to facilitate counting on the dorsal upper areas of the arms and a "nevometer" [18], a small, 2-mm thick

transparent polymethylmethacrylate spatula with a hole of 2 mm in diameter, to be used as reference when deciding whether the size of a lesion was above the 2 mm threshold.

Standardization of the nevus counting procedure between the examiners who performed the expert assessments was essential for the study. To achieve this goal, regular training sessions were held before the study dates in each winter and summer semester, during which the examiners discussed all aspects of the nevus counting procedure and carried out practical exercises.

2.3. Questionnaire

In addition to nevus counting, information from the study participants was collected using a self-administered questionnaire that comprised sociodemographic variables (age, sex, degree course) and phenotype information (hair and eye color, freckling, skin type). Participants were provided with an eye and hair color reference chart, an illustration of different categories of freckling, and a detailed description of how to assess the Fitzpatrick skin type [19] in order to standardize the assessment of phenotype.

2.4. Statistical Analysis

Agreement between the two nevus assessments was analyzed based on (i) the raw results of the nevus counting and (ii) a categorized nevus score (categories defined by the quintiles of the nevus distribution). For these analyses all participants with Fitzpatrick skin types V and VI were excluded. For the raw nevus counts, the mean difference (\pm standard deviation (SD)) between the two assessments and Pearson's correlation coefficient, accompanied by its 95%-confidence interval (CI), were computed to quantify the relationship between the two nevus assessments. Additionally, the differences between the two nevus counting results were examined graphically using a Bland-Altman plot, which displays the difference in relation to the average count as a surrogate for the unknown true number of nevi. For the categorized nevus score, raw agreement (in percent) and chance-corrected agreement by Cohen's weighted kappa coefficient (κ) were estimated from the 5x5-contingency table showing the joint distribution of participants' and experts' nevus classification [20]. The precision of these estimates is depicted by 95%-CIs.

To evaluate potential determinants of agreement, subgroups of the total study group were defined by the variables sex (male/female), degree course (clinical medicine/other), semester (summer/winter), skin type (Fitzpatrick type I-IV), freckles (yes/no), hair color (red, blonde, brown, black), and eye color (dark blue, light blue, green, green-brown, light brown, dark brown). Weighted kappa coefficients were estimated in the subgroups and differences between subgroup-specific estimates were statistically evaluated for each of the seven variables by the test for equal kappa coefficients between independent groups developed by Fleiss et al. [21]. P-values less than 0.05 were considered to indicate statistical significance but should be interpreted in an explorative manner as no adjustment for multiple testing was made. All statistical analyses were carried out using R software version 4.3.2 (R Foundation for Statistical Computing, Vienna, Austria). In particular, the R package irrCAC was used for estimating weighted kappa coefficients.

3. Results

Of the recruited study group (n=4704), 90 participants were excluded from the analysis due to their skin type (i.e., Fitzpatrick type V-VI), 48 had incomplete information about their nevus status because one of the assessments was missing, and data from additional 18 participants with double counting results were excluded from the analysis because the expert assessment was carried out by a one-time substitute examiner. The remaining 4548 participants comprised 1689 (37%) men, 2845 (63%) women and 14 (0.3%) participants who did not reveal their gender. The participants' mean (\pm SD) age was 23.53 (\pm 3.36) years. Altogether, 4156 (91%) of the participants studied clinical medicine, and 392 (9%) came from other disciplines (molecular medicine, life science engineering, speech therapy). Nearly equal numbers of participants completed the study during the summer term (n=

2241, 49%) and the winter term (n=2307, 51%). A descriptive summary of the distribution of phenotype variables in the study sample is given in Table 1.

Table 1. Distribution of phenotype variables in the study sample (n=4548).

| Phenotype variable | Absolute number (n ¹) | Proportion (%) |
|-----------------------|-----------------------------------|----------------|
| Fitzpatrick skin type | | |
| Type I | 234 | 5.18 |
| Type II | 1524 | 33.70 |
| Type III | 2352 | 52.01 |
| Type IV | 412 | 9.11 |
| Freckling | | |
| none | 2632 | 58.02 |
| few | 1546 | 34.08 |
| many | 358 | 7.89 |
| Hair color | | |
| Red | 61 | 1.34 |
| Blonde | 1214 | 26.75 |
| Brown | 2971 | 65.47 |
| Black | 292 | 6.44 |
| Eye color | | |
| Dark Blue | 496 | 10.94 |
| Light Blue/Grey | 1003 | 22.12 |
| Green | 847 | 18.68 |
| Green/Brown | 1060 | 23.37 |
| Light Brown | 610 | 13.45 |
| Dark Brown | 519 | 11.44 |

¹ Absolute numbers of participants in the categories of the phenotypic variables do not add to the total sample size of 4548 due to missing information in some (8-26 depending on the variable) cases.

3.1. Distribution of Nevi Counts

Both distributions, that of the self-assessed number of nevi and that of the expert-assessed number of nevi, were strongly skewed. While the median number of self-assessed (expert-assessed) nevi was 14 nevi (12 nevi), the arithmetic mean was 18.9 nevi (15.6 nevi). When defining the nevus score, i.e., the categorized version of the counted number of nevi, we used the joint distribution of all nevus counts from participants and experts. The quintiles of this joint distribution were 5, 10, 16, and 26 nevi, leading to the following five categories for the nevus score: ≤5, 5-10, 10-16, 16-26, and >26.

3.2. Nevi Counts: Differences Between Assessments

The results of self-assessed and expert-assessed nevus counting showed a correlation of 0.774 (95%-CI: 0.762 – 0.786). The self-assessed nevus counts were somewhat greater than the expert-assessed counts, and the mean difference (\pm SD) in the total sample was 3.33 (\pm 11.29). Figure 1 displays the Bland-Altman plot for the data showing the individual mean differences in relation to the averages of the two counts.

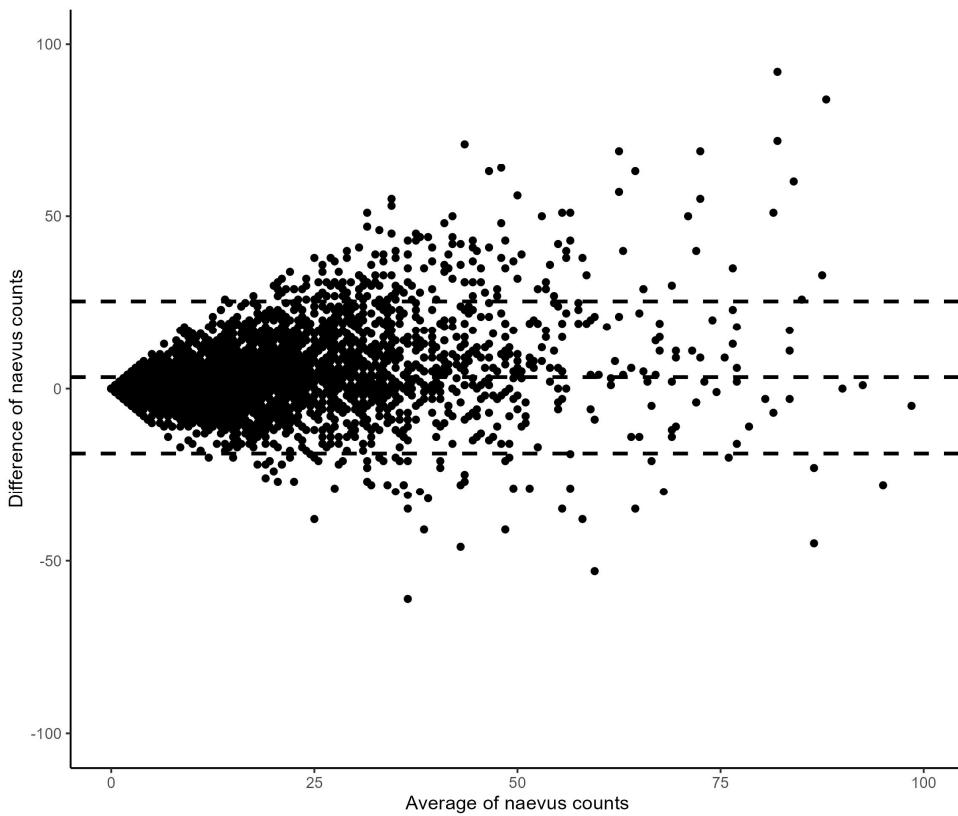


Figure 1. Bland-Altman plot showing the relationship between difference and average of the number of nevi counted by study participants and experts in a mutually blinded fashion (n=4539, nine data points with an average nevus count > 100 were outside the display range).

3.3. Nevus Score: Agreement Between Assessments

The joint distribution of the self-assessed and expert-assessed nevus scores is shown in Table 2. Based on the five-level score, the raw agreement in the total sample was 50.29% (95%-CI: 48.83% - 51.74%). The chance-corrected agreement was moderate ($\kappa=0.596$, 95%-CI: 0.581-0.611). Sensitivity analyses using a categorized nevus score with a different number of categories between 3 and 7 (using a data-adaptive definition of the specific categories, i.e. tertiles, quartiles, etc.) yielded stable results regarding estimated weighted kappa coefficients.

Table 2. Joint distribution of the five-level nevus score based on self-assessment (row variable) and expert assessments (column variable) for all 4548 double-counting results.

| | | Expert assessment | | | | | |
|-----------------|--|------------------------------|-----|-----|-----|-----|-----|
| | | [0,5](5,10](10,16](16,26]>26 | | | | | |
| | | [0,5] | 695 | 152 | 37 | 12 | 0 |
| | | (5,10] | 251 | 380 | 166 | 51 | 12 |
| Self-assessment | | (10,16] | 87 | 218 | 285 | 165 | 21 |
| | | (16,26] | 38 | 135 | 292 | 368 | 115 |
| | | > 26 | 11 | 55 | 132 | 311 | 559 |

Detailed results of the agreement between the two assessments in the subgroups defined by sociodemographic and phenotype variables are shown in Table 3. Overall, the level of agreement was quite homogenous across the subgroups. For all variables, except eye color, we found nonsignificant variation of limited magnitude in the weighted kappa estimates of the corresponding subgroups.

However, in the subgroups defined by participants' eye color, we observed a much lower weighted kappa of 0.512 (95%-CI: 0.462-0.561) for those with dark blue eyes and a higher weighted kappa of 0.642 (95%-CI: 0.598-0.685) for those with dark brown eyes, whereas the other four subgroups showed weighted kappas at the same level as in the total sample. Heterogeneity in the level of agreement between the six subgroups defined by participants' eye color reached significance ($p=0.01$).

Table 3. Agreement between self-assessed and expert-assessed nevi counts in subgroups defined by sociodemographic and phenotype variables. Raw agreement (in %) and weighted kappa, both accompanied by 95%-confidence intervals (CI), for the subgroups and result of the statistical evaluation of heterogeneity of subgroup estimates per variable.

| Subgroup | Raw agreement in % (95%-CI) | Weighted Kappa (95%-CI) | p-value |
|-----------------------|-----------------------------|-------------------------|---------|
| Sex | | | 0.08 |
| male | 47.90 (45.52 – 50.28) | 0.579 (0.554 – 0.604) | |
| female | 51.70 (49.87 – 53.54) | 0.607 (0.588 – 0.626) | |
| Degree course | | | 0.76 |
| Clinical medicine | 50.14 (48.62 – 51.66) | 0.596 (0.580 – 0.611) | |
| other | 51.79 (46.84 – 56.73) | 0.605 (0.554 – 0.655) | |
| Time | | | 0.54 |
| Summer term | 50.25 (48.18 – 52.32) | 0.601 (0.580 – 0.622) | |
| Winter term | 50.33 (48.28 – 52.37) | 0.592 (0.570 – 0.613) | |
| Fitzpatrick skin type | | | 0.72 |
| Type I | 53.42 (47.03 – 59.81) | 0.609 (0.542 – 0.677) | |
| Type II | 49.87 (47.36 – 52.38) | 0.585 (0.558 – 0.612) | |
| Type III | 49.53 (47.51 – 51.55) | 0.581 (0.560 – 0.603) | |
| Type IV | 55.10 (50.29 – 59.90) | 0.607 (0.556 – 0.658) | |
| Freckling | | | 0.89 |
| none | 50.27 (48.36 – 52.18) | 0.589 (0.569 – 0.609) | |
| few | 49.74 (47.25 – 52.23) | 0.588 (0.561 – 0.614) | |
| many | 52.79 (47.62 – 57.97) | 0.574 (0.514 – 0.633) | |
| Hair color | | | 0.15 |
| Red | 52.46 (39.93 – 64.99) | 0.561 (0.412 – 0.711) | |
| Blonde | 48.11 (45.29 – 50.92) | 0.565 (0.534 – 0.595) | |
| Brown | 50.32 (48.52 – 52.12) | 0.599 (0.581 – 0.618) | |
| Black | 58.90 (53.26 – 64.55) | 0.631 (0.571 – 0.691) | |
| Eye color | | | 0.01 |
| Dark Blue | 43.55 (39.18 – 47.91) | 0.512 (0.462 – 0.561) | |
| Light Blue/Grey | 48.75 (45.66 – 51.85) | 0.590 (0.558 – 0.622) | |
| Green | 51.71 (48.35 – 55.08) | 0.606 (0.570 – 0.641) | |
| Green/Brown | 49.91 (46.90 – 52.92) | 0.581 (0.548 – 0.613) | |
| Light Brown | 51.48 (47.51 – 55.44) | 0.595 (0.553 – 0.638) | |
| Dark Brown | 56.84 (52.58 – 61.10) | 0.642 (0.598 – 0.685) | |

4. Discussion

In a controlled study setting, we evaluated nevus self-counting by comparing self-assessed and expert-assessed nevus counts obtained in a mutually blinded fashion. Based on 4548 double counting results we observed only a moderate level of agreement between the two assessments. Self-assessed

nevus counts were higher than those determined by experts. The level of agreement was independent of sociodemographic and cutaneous factors. We found, however, a greater than average level of agreement in participants with dark brown eyes and a lower than average level of agreement in those with dark blue eyes. Participants' eye color was the only factor that had a significant impact on the level of agreement in our study.

Tackling aspects of validity and reliability of nevus self-counting is not a new topic in dermat-epidemiologic research. In a systematic literature search, we identified 21 studies published between 1991 and 2021 [18,22–41] that reported some data on the relationship between self-counting and expert counting of nevi on the same individuals; 13 of these studies focused primarily on this topic. All previous investigations were based on much smaller study samples comprising 46 to 1772 participants which limits the precision of quantifying the level of agreement. Most of them were embedded in studies with melanoma patients or used participants of screening programs for skin cancer as their study subjects. Unsurprisingly, the results of these studies were very heterogeneous as the investigations differed not only with respect to the origin and size of the study population but also with respect to the instructions for self-counting given to study participants. The lowest level of agreement was observed in a study of 1658 employees of bank and insurance companies who had only some minutes to fill out a self-administered questionnaire prior to a voluntary skin check by a dermatologist [39]. The questionnaire comprised one question asking participants to estimate the total number of moles on their body (0-10, 11-30, 31-50, 51-100, >100) without further explanation and instructions. These self-assessments of nevi showed virtually no agreement beyond chance with the results of nevus counting by the dermatologist (weighted $\kappa=0.03$). Other studies have reported higher levels of agreement, which is more consistent with our findings. For example, Jackson et al.'s study [26] in the United Kingdom, comprising eight general practitioner practices with 388 participants, found an identical level of agreement ($\kappa=0.60$) as our study for a three-level nevus score. A study by Mannino et al. [41] in three European dermatologic clinics that assisted participants in assessing their nevi with instructions and photographs similar to ours reported a moderate level of agreement (weighted $\kappa=0.45$), slightly below our result, in their sample of 744 patients.

The vast majority (91%) in our study sample were students attending a compulsory course in a medical degree program. Of course, such a sample is not representative of the general population. At the time of the study, however, the students had no previous dermatologic training and thus had no experience in distinguishing melanocytic nevi from other skin lesions. Nevertheless, we cannot exclude the possibility that their general medical training and attitudes toward this topic may have influenced their motivation and diligence when participating in the study, which may have biased the results toward a higher level of agreement in our study than would be expected in the general population.

When self-assessed information on melanoma risk factors is used to identify a subgroup in the population at high risk of developing melanoma as a target of screening activities, the self-assessed information needs to be accurate. Although prediction models developed for estimating melanoma risk usually include several variables, the frequency of nevi is a particularly important component of such models [15]. From the literature on melanoma prediction models there is some empirical evidence that the discriminatory properties of melanoma prediction models incorporating self-assessed nevus information are inferior to those using physician-assessed nevus information. Four publications [42–45] described melanoma prediction models derived from the same data, namely, the Australian Melanoma Family Study [46], and validated them using data from, again, the same population-based case-control study in the United Kingdom [47]. The prediction models differed with respect to the incorporated variables: two [44,46] used genotype information in addition to phenotype and UV exposure variables, while the remaining two [43,45] focused on nongenetic risk factors. The difference between the latter two models related to the incorporation of only self-assessed phenotype variables in [43] and the use of physician-assessed phenotype variables in [45]. The model performance differed significantly, the AUC parameter describing model discrimination ranged from 0.66 for the model including only self-assessed risk factors without genotype information to 0.79 for the model including physician-assessed phenotype variables and genotype information. The main

driver of the increase in the AUC was the use of physician-assessed nevus counts instead of self-assessed nevus information, while the use of genotype information had only a modest impact on the AUC.

Our large-scale investigation showed that even detailed instructions including photographs explaining the definition of melanocytic nevi, i.e., which skin lesions should and should not be counted, failed to increase the level of agreement between self-assessed and expert-assessed nevus counts to an acceptable level of substantial or near-perfect agreement. Together with the results of other studies it seems questionable to use self-assessed information on phenotype variables such as the number of nevi when defining a high-risk subgroup of the population for targeted melanoma screening.

5. Conclusions

Our results show that even with comprehensive guidance, only a moderate level of agreement between self-assessed and expert-assessed nevus counts can be achieved. Self-assessed nevus information does not appear to be reliable enough to be used in individual risk assessment for targeted screening programs.

Author Contributions: Conceptualization, O.G. and A.B.P.; methodology, O.G., I.K., and A.B.P.; software, E.B. and A.B.P.; validation, I.K. and W.U.; formal analysis, E.B.; investigation, O.G. and W.U.; resources, O.G.; data curation, I.K. and A.B.P.; writing—original draft preparation, O.G.; writing—review and editing, I.K., E.B., W.U., and A.B.P.; visualization, E.B. and A.B.P.; supervision, A.B.P.; project administration, A.B.P. All authors have read and agreed to the published version of the manuscript.”

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Ethics Committee at the Friedrich-Alexander-University of Erlangen-Nuremberg.

Informed Consent Statement: Written informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: We gratefully acknowledge the contribution of Martin Radespiel-Tröger, Martina Malzer, Jasmin Tarantino, and Jiang Li, who counted the participants' nevi during some part of the study period.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Garbe, C.; Keim, U.; Gandini, S.; Amaral, T.; Katalinic, A.; Holleczek, B.; Martus, P.; Flatz, L.; Leiter, U.; Whiteman, D. Epidemiology of cutaneous melanoma and keratinocyte cancer in white populations 1943-2036. *Eur J Cancer* **2021**, *152*, 18-25, doi:10.1016/j.ejca.2021.04.029.
2. Houghton, A.; Flannery, J.; Viola, M.V. Malignant melanoma in Connecticut and Denmark. *Int J Cancer* **1980**, *25*, 95-104, doi:10.1002/ijc.2910250113.
3. Bulliard, J.L.; Cox, B.; Elwood, J.M. Latitude gradients in melanoma incidence and mortality in the non-Maori population of New Zealand. *Cancer Causes Control* **1994**, *5*, 234-240, doi:10.1007/BF01830242.
4. Svedman, F.C.; Pillas, D.; Taylor, A.; Kaur, M.; Linder, R.; Hansson, J. Stage-specific survival and recurrence in patients with cutaneous malignant melanoma in Europe - a systematic review of the literature. *Clin Epidemiol* **2016**, *8*, 109-122, doi:10.2147/CLEP.S99021.
5. Enninga, E.A.L.; Moser, J.C.; Weaver, A.L.; Markovic, S.N.; Brewer, J.D.; Leontovich, A.A.; Hieken, T.J.; Shuster, L.; Kottschade, L.A.; Olariu, A.; et al. Survival of cutaneous melanoma based on sex, age, and stage in the United States, 1992-2011. *Cancer Med* **2017**, *6*, 2203-2212, doi:10.1002/cam4.1152.
6. Eriksson, H.; Utjes, D.; Olofsson Bagge, R.; Gillgren, P.; Isaksson, K.; Lapins, J.; Schultz, I.L.; Lyth, J.; Andersson, T.M. The Proportion Cured of Patients with Resected Stage II-III Cutaneous Melanoma in Sweden. *Cancers (Basel)* **2021**, *13*, doi:10.3390/cancers13102456.
7. Henrikson, N.B.; Ivlev, I.; Blasi, P.R.; Nguyen, M.B.; Senger, C.A.; Perdue, L.A.; Lin, J.S. Skin Cancer Screening: Updated Evidence Report and Systematic Review for the US Preventive Services Task Force. *JAMA* **2023**, *329*, 1296-1307, doi:10.1001/jama.2023.3262.
8. Bell, K.J.L.; Nijsten, T. Melanoma overdiagnosis: why it matters and what can be done about it. *Br J Dermatol* **2022**, *187*, 459-460, doi:10.1111/bjd.21750.

9. Ebell, M.H.; Thai, T.N.; Royalty, K.J. Cancer screening recommendations: an international comparison of high income countries. *Public Health Rev* **2018**, *39*, 7, doi:10.1186/s40985-018-0080-0.
10. Breitbart, E.W.; Waldmann, A.; Nolte, S.; Capellaro, M.; Greinert, R.; Volkmer, B.; Katalinic, A. Systematic skin cancer screening in Northern Germany. *J Am Acad Dermatol* **2012**, *66*, 201-211, doi:10.1016/j.jaad.2010.11.016.
11. Rat, C.; Blachier, L.; Hild, S.; Molinie, F.; Gaultier, A.; Dreno, B.; Nguyen, J.-M. Targeted screening for melanoma after a 5-year follow-up: Comparison of melanoma incidence and lesion thickness at diagnosis in screened (versus unscreened) patients. *La Presse Médicale Open* **2021**, *2*, 100013, doi:https://doi.org/10.1016/j.lpmope.2021.100013.
12. Bobrowska, A.; Murton, M.; Seedat, F.; Visintin, C.; Mackie, A.; Steele, R.; Marshall, J. Targeted screening in the UK: A narrow concept with broad application. *Lancet Reg Health Eur* **2022**, *16*, 100353, doi:10.1016/j.lanepe.2022.100353.
13. Gandini, S.; Sera, F.; Cattaruzza, M.S.; Pasquini, P.; Abeni, D.; Boyle, P.; Melchi, C.F. Meta-analysis of risk factors for cutaneous melanoma: I. Common and atypical naevi. *Eur J Cancer* **2005**, *41*, 28-44, doi:10.1016/j.ejca.2004.10.015.
14. Pampena, R.; Kyrgidis, A.; Lallas, A.; Moscarella, E.; Argenziano, G.; Longo, C. A meta-analysis of nevus-associated melanoma: Prevalence and practical implications. *J Am Acad Dermatol* **2017**, *77*, 938-945 e934, doi:10.1016/j.jaad.2017.06.149.
15. Kaiser, I.; Pfahlberg, A.B.; Uter, W.; Hepp, M.V.; Veierod, M.B.; Gefeller, O. Risk Prediction Models for Melanoma: A Systematic Review on the Heterogeneity in Model Development and Validation. *Int J Environ Res Public Health* **2020**, *17*, doi:10.3390/ijerph17217919.
16. International Agency for Research on Cancer. Epidemiological studies on melanocytic naevi: protocol for identifying and recording naevi. IARC Internal Report No. 90/002; International Agency for Research on Cancer: Lyon, 1990.
17. Farinas-Alvarez, C.; Rodenas, J.M.; Herranz, M.T.; Delgado-Rodriguez, M. The naevus count on the arms as a predictor of the number of melanocytic naevi on the whole body. *Br J Dermatol* **1999**, *140*, 457-462, doi:10.1046/j.1365-2133.1999.02709.x.
18. Carli, P.; Giorgi, V.d.; Nardini, P.; Mannone, F.; Palli, D.; Giannotti, B. Melanoma detection rate and concordance between self-skin examination and clinical evaluation in patients attending a pigmented lesion clinic in Italy. *British Journal of Dermatology* **2002**, *146*, 261-266, doi:10.1046/j.1365-2133.2002.04580.x.
19. Fitzpatrick, T.B. The validity and practicality of sun-reactive skin types I through VI. *Arch Dermatol* **1988**, *124*, 869-871, doi:10.1001/archderm.124.6.869.
20. Cohen, J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol Bull* **1968**, *70*, 213-220, doi:10.1037/h0026256.
21. Fleiss, J.L., Levin, B., Paik, M. C. The Measurement of Interrater Agreement. In *Statistical Methods for Rates and Proportions*; John Wiley & Sons, Inc. : New York, 2003.
22. WALTER, S.D.; MARRETT, L.D.; HERTZMAN, C. Reliability of interviewer and subject assessments of nevus counts in a study of melanoma. *Journal of Clinical Epidemiology* **1991**, *44*, 633-640, doi:10.1016/0895-4356(91)90024-4.
23. Gruber, S.B.; Roush, G.C.; Barnhill, R.L. Sensitivity and Specificity of Self-examination for Cutaneous Malignant Melanoma Risk Factors. *American Journal of Preventive Medicine* **1993**, *9*, 50-54.
24. Lawson, D.D.; Schneider, J.S.; Sagebiel, R.W. Nevus counting as a risk factor for melanoma: Comparison of self-count with count by physician. *Journal of the American Academy of Dermatology* **1994**, *31*, 438-444, doi:10.1016/s0190-9622(94)70207-1.
25. Little, P.; Keefe, M.; White, J. Self screening for risk of melanoma: Validity Of self mole counting by patients in a single general practice. *BMJ* **1995**, *310*, 912, doi:10.1136/bmj.310.6984.912.
26. Jackson, A.; Wilkinson, C.; Ranger, M.; Pill, R.; August, P. Can primary prevention or selective screening for melanoma be more precisely targeted through general practice? A prospective study to validate a self administered risk score. *BMJ* **1998**, *316*, 34-38, doi:10.1136/bmj.316.7124.34.
27. Buettnner, P.G.; Garbe, C. Agreement between self-assessment of melanocytic nevi by patients and dermatologic examination. *American Journal of Epidemiology* **2000**, *151*, 72-77.
28. Melia, J.; Harland, C.; Moss, S.; Eiser, J.R.; Pendry, L. Feasibility of targeted early detection for melanoma: a population-based screening study. *Br J Cancer* **2000**, *82*, 1605-1609, doi:10.1054/bjoc.2000.1183.
29. Harrison, S.L.; Buettnner, P.G.; MacLennan, R.; Kelly, J.W.; Rivers, J.K. How good are parents at assessing melanocytic nevi on their children? A study comparing parental counts, dermatologist counts, and counts obtained from photographs. *American Journal of Epidemiology* **2002**, *155*, 1128-1136, doi:10.1093/aje/155.12.1128.
30. Harbauer, A.; Binder, M.; Pehamberger, H.; Wolff, K.; Kittler, H. Validity of an unsupervised self-administered questionnaire for self-assessment of melanoma risk. *Melanoma Research* **2003**, *13*, 537-542, doi:10.1097/00008390-200310000-00013.

31. Richtig, E.; Santigli, E.; Fink-Puches, R.; Weger, W.; Hofmann-Wellenhof, R. Assessing melanoma risk factors: How closely do patients and doctors agree? *Public Health* **2008**, *122*, 1433–1439, doi:10.1016/j.puhe.2008.04.012.
32. Boone, S.L.; Stapleton, J.; Turrisi, R.; Ortiz, S.; Robinson, J.K.; Mallett, K.A. Thoroughness of skin examination by melanoma patients: Influence of age, sex and partner. *Australasian Journal of Dermatology* **2009**, *50*, 176–180, doi:10.1111/j.1440-0960.2009.00533.x.
33. Richtig, E.; Jung, E.; Asback, K.; Trapp, M.; Hofmann-Wellenhof, R. Knowledge and Perception of Melanocytic Nevi and Sunburn in Young Children. *PEDIATRIC DERMATOLOGY* **2009**, *26*, 519–523, doi:10.1111/j.1525-1470.2008.00854.x.
34. Quereux, G.; Nguyen, J.M.; Volteau, C.; Lequeux, Y.; Dreno, B. Creation and test of a questionnaire for self-assessment of melanoma risk factors. *EUROPEAN JOURNAL OF CANCER PREVENTION* **2010**, *19*, 48–54, doi:10.1097/CEJ.0b013e328333d113.
35. Morze, C.J.; Olsen, C.M.; Perry, S.L.; Jackman, L.M.; Ranieri, B.A.; O'Brien, S.M.; Cicero, R.A.; Whiteman, D.C. Good test-retest reproducibility for an instrument to capture self-reported melanoma risk factors. *Journal of Clinical Epidemiology* **2012**, *65*, 1329–1336, doi:10.1016/j.jclinepi.2012.06.014.
36. Fießler, C.; Pfahlberg, A.; Li, J.; Uter, W.; Gefeller, O. Accuracy and reliability of naevus self-counts. *Melanoma Research* **2014**, *24*, 611–616, doi:10.1097/cmr.0000000000000111.
37. Cust, A.E.; Pickles, K.M.; Goumas, C.; Vu, T.; Schmid, H.; Nagore, E.; Kelly, J.; Aitken, J.F.; Giles, G.G.; Hopper, J.L.; et al. Accuracy of self-reported nevus and pigmentation phenotype compared with clinical assessment in a population-based study of young Australian Adults. *Cancer Epidemiology Biomarkers and Prevention* **2015**, *24*, 736–743, doi:10.1158/1055-9965.Epi-14-1203.
38. Stapleton, J.L.; Turrisi, R.; Mallett, K.A.; Robinson, J.K. Correspondence between pigmented lesions identified by melanoma patients trained to perform partner-assisted skin self-examination and dermatological examination. *Cancer Epidemiology Biomarkers and Prevention* **2015**, *24*, 1247–1253, doi:10.1158/1055-9965.Epi-15-0218.
39. Winkler, A.; Plugfelder, A.; Weide, B.; Leiter, U.; Held, L.; Garbe, C.; Eigentler, T. Screening for skin cancer in bank and insurance employees: Risk profile and correlation of self and physician's assessment. *International Journal of Dermatology* **2015**, *54*, 419–423, doi:10.1111/ijd.12360.
40. Betz-Stablein, B.; Koh, U.; Plasmeijer, E.I.; Janda, M.; Aitken, J.F.; Soyer, H.P.; Green, A.C. Self-reported naevus density may lead to misclassification of melanoma risk. *British Journal of Dermatology* **2020**, *182*, 1488–1490, doi:10.1111/bjd.18802.
41. Mannino, M.; Sollena, P.; Esposito, M.; Farnolli, M.C.; Peris, K.; Nagore, E. Self-Assessment Questionnaire on Patient-Physician Concordance on Nevus Self-Count and Models Development to Predict High-Risk Phenotype > 50 Nevi. *DERMATOLOGY*, doi:10.1159/000523953.
42. Cust, A.E.; Goumas, C.; Vuong, K.; Davies, J.R.; Barrett, J.H.; Holland, E.A.; Schmid, H.; Agha-Hamilton, C.; Armstrong, B.K.; Kefford, R.F.; et al. MC1R genotype as a predictor of early-onset melanoma, compared with self-reported and physician-measured traditional risk factors: an Australian case-control-family study. *BMC Cancer* **2013**, *13*, 406, doi:10.1186/1471-2407-13-406.
43. Vuong, K.; Armstrong, B.K.; Weiderpass, E.; Lund, E.; Adam, H.O.; Veierod, M.B.; Barrett, J.H.; Davies, J.R.; Bishop, D.T.; Whiteman, D.C.; et al. Development and External Validation of a Melanoma Risk Prediction Model Based on Self-assessed Risk Factors. *JAMA Dermatol* **2016**, *152*, 889–896, doi:10.1001/jamadermatol.2016.0939.
44. Cust, A.E.; Drummond, M.; Kanetsky, P.A.; Australian Melanoma Family Study, I.; Leeds Case-Control Study, I.; Goldstein, A.M.; Barrett, J.H.; MacGregor, S.; Law, M.H.; Iles, M.M.; et al. Assessing the Incremental Contribution of Common Genomic Variants to Melanoma Risk Prediction in Two Population-Based Studies. *J Invest Dermatol* **2018**, *138*, 2617–2624, doi:10.1016/j.jid.2018.05.023.
45. Vuong, K.; Armstrong, B.K.; Drummond, M.; Hopper, J.L.; Barrett, J.H.; Davies, J.R.; Bishop, D.T.; Newton-Bishop, J.; Aitken, J.F.; Giles, G.G.; et al. Development and external validation study of a melanoma risk prediction model incorporating clinically assessed naevi and solar lentigines. *Br J Dermatol* **2020**, *182*, 1262–1268, doi:10.1111/bjd.18411.

46. Cust, A.E.; Schmid, H.; Maskiell, J.A.; Jetann, J.; Ferguson, M.; Holland, E.A.; Agha-Hamilton, C.; Jenkins, M.A.; Kelly, J.; Kefford, R.F.; et al. Population-based, case-control-family design to investigate genetic and environmental influences on melanoma risk: Australian Melanoma Family Study. *Am J Epidemiol* **2009**, *170*, 1541–1554, doi:10.1093/aje/kwp307.
47. Newton-Bishop, J.A.; Chang, Y.M.; Iles, M.M.; Taylor, J.C.; Bakker, B.; Chan, M.; Leake, S.; Karpavicius, B.; Haynes, S.; Fitzgibbon, E.; et al. Melanocytic nevi, nevus genes, and melanoma risk in a large case-control study in the United Kingdom. *Cancer Epidemiol Biomarkers Prev* **2010**, *19*, 2043–2054, doi:10.1158/1055-9965.EPI-10-0233.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.