Dermatology Practical & Conceptual

Systematic Review of Nevus Counting and Reporting Methodologies in Contemporary Studies of the General Population

Dilki Jayasinghe[1], Nathasha Naranpanawa[2], Daniel Paul Ashley[3,4], Åsa Ingvar[5,6], Brigid Betz-Stablein[7], H Peter Soyer[7,8], Monika Janda[1]

1 Centre for Health Services Research, Faculty of Medicine, The University of Queensland, Brisbane, Queensland, Australia
2 School of Electrical Engineering and Computer Science, The University of Queensland, Brisbane, Queensland, Australia
3 Faculty of Medicine, The University of Queensland, Brisbane, Queensland, Australia
4 Ochsner Clinical School, Ochnser Health, New Orleans, Louisiana, United States of America
5 Department of Dermatology, Skåne University Hospital, Lund University, Lund, Sweden
6 School of Public Health and Preventive Medicine, Monash University, Melbourne, Victoria, Australia
7 Frazer Institute, The University of Queensland, Dermatology Research Centre, Brisbane, Queensland, Australia
8 Dermatology Department, Princess Alexandra Hospital, Brisbane, Queensland, Australia

Key words: melanoma risk, standardization, reliability, validity, inter- and intra-observer variation

Citation: Jayasinghe D, Naranpanawa N, Ashley DP, et al. Systematic Review of Nevus Counting and Reporting Methodologies in Contemporary Studies of the General Population. *Dermatol Pract Concept*. 2024;14(4):e2024223. DOI: https://doi.org/10.5826/dpc.1404a223

Accepted: May 8, 2024; **Published:** October 2024

Funding: None.

Competing Interests: None.

Authorship: All authors have contributed significantly to this publication.

Corresponding Author: Dilki Jayasinghe, Centre for Health Services Research, Faculty of Medicine, The University of Queensland, 34 Cornwall St, Woolloongabba, Queensland, 4102, Australia. Tel: +61 434878262; E-mail: dilki.jayasinghe@uq.edu.au

ABSTRACT

Introduction: Having many melanocytic nevi on the skin is a risk factor for melanoma. However, the reproducibility of nevus counts in previous studies is limited due to high inter- and intraobserver variation. Despite the introduction of a protocol for counting and reporting of nevi in 1990 by the International Agency for Research on Cancer (IARC), significant variations in nevus counting methods persist across studies.

Objectives: We sought to review the variations in nevus counting and reporting methods, adherence and deviations from the IARC protocol, and the reproducibility of nevus counting studies.

Methods: A systematic search of Embase, PubMed, and Web of Science was conducted. The review was limited to nevus (>2 mm) counting studies of general population adults conducted between 2000 and 2022 and studies using skilled examiners.

Results: Out of the eight studies which were eligible for inclusion, none followed the IARC protocol. Three studies used a predefined criterion to count nevi. Five studies provided training for their observers. Three studies assessed the inter- or intraobserver variation using the correlation coefficient (>0.75), and three studies attempted to verify the validity and the reproducibility of the counts. There was little to no agreement in nevus counting and reporting procedures in the reviewed studies, and most studies did not report their procedures adequately.

Conclusion: This review highlights the need for an easily accessible and feasible protocol for the identification, counting, and reporting of nevi, which also considers nevus counting from total-body imaging and automated nevus counts since these technologies are expected to become widely available for future studies.

Introduction

Melanoma incidence is predicted to increase by 57% worldwide by 2040, with an estimated 68% rise in mortality [1]. Having many melanocytic nevi is the strongest phenotypic risk factor for melanoma. Nevus numbers are thought to increase into middle age and decrease thereafter, but this is mostly based on cross-sectional studies. A recent longitudinal study reported dynamic changes of nevus numbers even into late adulthood [2].

A major constraint in reporting and comparing nevus prevalence across studies of different age groups, risk groups, and geographic locations is the lack of a standard procedure for counting and recording them. Addressing this issue, English et al. (1990) [3] formulated a set of guidelines for differentiating a nevus from other skin lesions and reporting nevus counts. This has been recognized as the standard international protocol for counting nevi by the International Agency for Research on Cancer (IARC). Despite this IARC protocol, comparability of studies may be low if it is not implemented consistently. Differences in the size of nevi included in the analyses, inadequate information about nevus counting methods, and reporting of nevus counts versus nevus density are commonly seen inconsistencies in previous studies. In fact, a previous meta-analysis reported strong overall heterogeneity in nevus counting methods [4]. A recent scoping review of the body site distribution of nevi [5] also highlighted the difficulty in comparing prevalence of nevi by anatomic site due to differences in aggregating body sites in previous studies.

Reproducibility of nevus counts in research studies is also hampered by high inter- and intraobserver variation [3, 6-8], which may account for up to 10% of variation in total nevus count [9]. Numerous observers with varying experience were involved in counting nevi in previous studies, including dermatologists, physicians, nurses, and trained interviewers or self-report. Intraobserver reliability increases with training and experience of the person counting nevi [6].

Given these previous reports, this systematic review evaluated variations in nevus counting and reporting methods, adherence to or deviations from the IARC protocol, and reproducibility of previous nevus counting studies. Criteria used for identification and counting of nevi >2 mm in diameter, body sites, reported outcomes (nevus count, nevus density, or categorized nevus count), personnel performing the nevus counting and related training, and assessment of inter- and intraobserver variation were also investigated.

Methods

This systematic review followed PRISMA guidelines [10]. A protocol for the review was registered with the International Prospective Register for Ongoing Systematic Reviews (PROSPERO)[11] (registration ID: CRD42021278664). Studies reporting total body melanocytic nevus counts of >2 mm in diameter in adults (most of the reported sample being >18 years) from the general population using skilled examiners for nevus counting were eligible. Studies recruiting from dermatology clinics, those using self-reported nevus counts as well as studies of atypical nevus syndrome, those specifically counting a subtype of nevi (e.g., Unna nevi, Spitz nevi), or studies focusing on children were excluded.

A systematic search of Embase, PubMed, and Web of Science was conducted and updated in December 2022 using a combination of key words related to nevus counting studies in general population adults and Medical Subject Heading (MeSH) terms (Appendix 1). The search was limited to studies published after 2000 to assess the quality of recent nevus counting studies.

Covidence [12] was used for screening. Retrieved studies were independently assessed by two reviewers (DJ and DPA), and a decision by a third reviewer (MJ) was sought when needed. Pre-tested tables were used for data extraction by two authors (DJ and NN). Descriptions of nevus counting and reporting methods and any deviations from the IARC protocol for identifying and recording nevi were recorded in detail.

Quality assessment of the selected studies was carried out by two independent reviewers (DJ and NN) using the appropriate Joanna Briggs Institute (JBI) Critical Appraisal Checklist [13-15], based on the type of study. Since the review was mainly focused on the study methods and not their outcomes, we developed ranking criteria to assess the quality of the nevus counting methodologies and their reproducibility based on recommendations of the IARC protocol. The criteria considered whether the study reported: (i) a clear definition for melanocytic nevi; (ii) steps taken to make sure nevi were not confused with other pigmented lesions; (iii) a standard (tool) to measure nevus size; (iv) the setting where nevus counting was carried out (e.g., lighting conditions, furniture used, magnification instruments); (v) followed a set sequence of measurements, e.g., melanocytic nevi, assessing freckling, assessing solar lentigines, and described detailed macroscopic features of a specified sample of melanocytic nevi; (vi) training of observers according to a set of guidelines; (vii) reported reproducibility and inter- and intraobserver variability; and (viii) adherence to the IARC protocol or any other protocol considered.

Results

The PRISMA flow diagram (shown in Figure 1) summarizes the search and screening results. A total of 4,638 articles were identified. After removal of duplicates (n=1,386) and screening of titles and abstracts, 303 articles remained for full text review. Eight articles met all the eligibility criteria and were subjected to complete review.

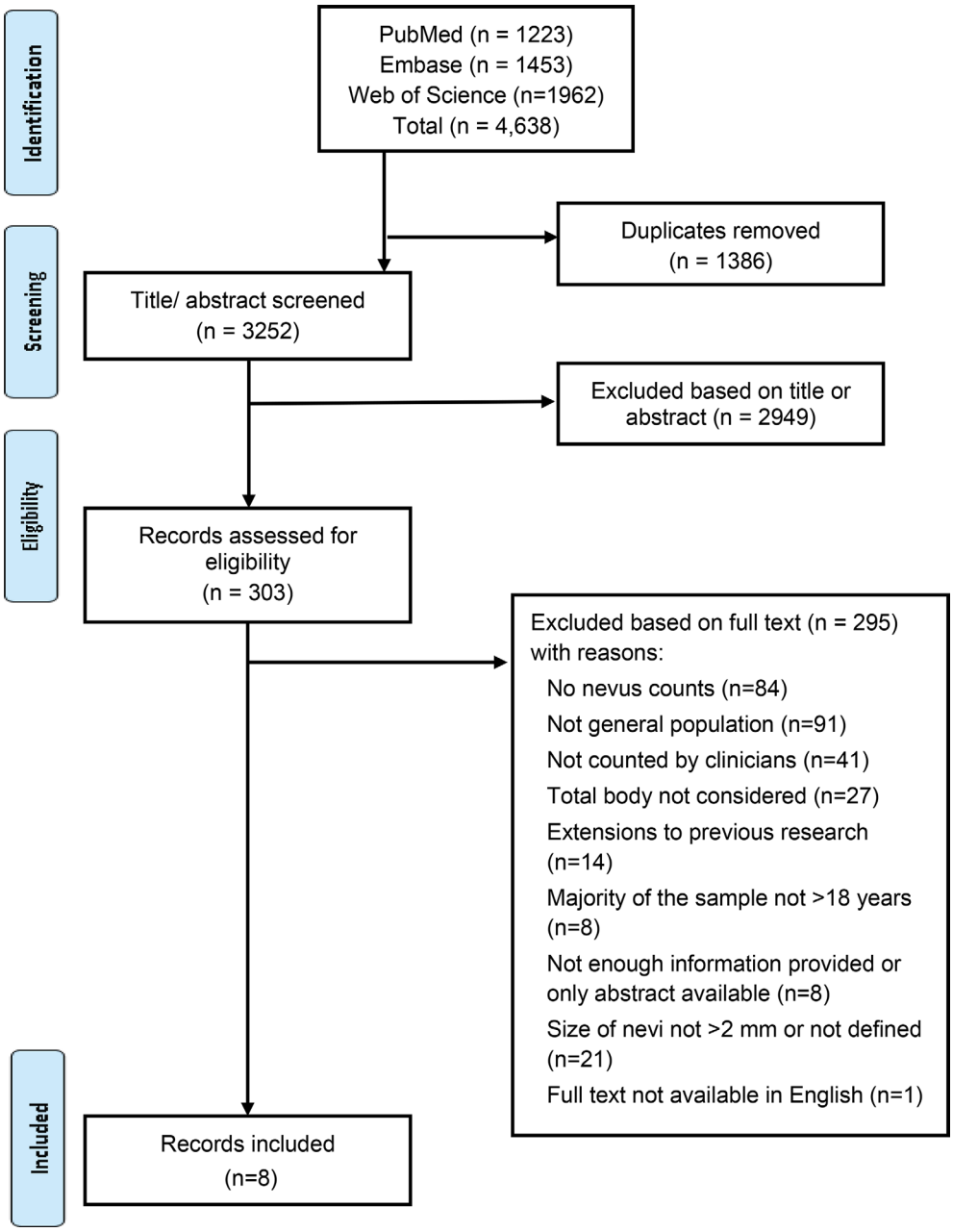

Figure 1. PRISMA flow diagram of the screening process.

Characteristics of Sources of Evidence

All selected studies were from Europe (Table 1). Study designs included four case-control studies [16-19] and four cross-sectional studies [20-23]. The eight studies reported data from 375,464 participants, whose age ranged from 14-77 years. The four case-control studies matched controls for sex and age of the cases. One study [23] was a population-based mass skin cancer screening study in Germany, while five others [16,18,20-22] used a sample from the general population. Two case-control studies [17,19] used patients in a hospital setting other than dermatology clinics (for this review, persons attending a non-dermatology clinic were considered representative of the general population) and spouses or friends of the cases as the control.

Methodologies for Nevus Identification and Reporting

Standardized protocol: Of the eight studies considered, none specifically reported that the investigators followed the IARC protocol (Table 2). One study [22] used a counting protocol validated in two previous melanoma studies from the same group [24, 25], while one study reported that the number of nevi was counted according to standardized criteria which the authors did not describe or reference [19], and one study had a well-defined description of the screening procedure to train the observers [23].

Defining and identifying nevi: None of the studies reported a well-defined criterion for identifying a lesion as a nevus. Three studies explicitly mentioned that they considered "melanocytic nevi" [18, 23] or "pigmented macules or papules" [17], while the remaining five studies did not specify what lesions were considered as nevi. To ensure that nevi were not confused with other pigmented lesions, Breitbart et al. (2012) tested the ability of examiners to differentiate between various skin lesions after the completion of a training program. Three studies also provided clear definitions of other lesions such as freckles and solar lentigines [17-19] which may also have prevented confusion with nevi.

Size of nevi: While six studies did not categorize the size of nevi for analysis (other than ≥2 mm), Newton-Bishop et al. [16] analyzed >5 mm nevi separately, and Bataille et al. [22] considered three categories of size: 2<x<5 mm, 5<x<10 mm, and >10 mm. With regards to tools, only Naldi et al. (2000) [19] reported that they used a simple instrument called a "nevometer" to measure the sizes of nevi.

Atypical nevi: All studies except one [18] separately recorded atypical nevi or considered those >5 mm as a proxy for atypical nevi [22].

Body sites: Three studies [16, 20, 22] reported nevus counts according to predefined body sites and one with assistance of a photographic atlas [19]. One study [18] did not specify body sites excluded for examination, while all other studies did not count nevi on genitalia. Two studies [19, 22] also excluded nevi on the scalp, and two studies [16, 21] excluded nevi on the breasts in females.

Nevus Count-Related Outcomes

Three studies [20-22] reported the mean and/ or median total number of nevi of all participants, while one study [16] reported the median number of nevi for participants categorized by sun exposure characteristics and genotypes (Table 3). The median number of nevi ≥2 mm varied from 11 to 45, while their ranges varied from 0 to355.

Four studies [17-19, 23] did not provide a single measure to summarize average nevus count but reported frequency and percentage of participants belonging to nevus count category subgroups (e.g., 0–20, 20–50, and >50 nevi) (Table 3). Categories varied by study and are summarized in Table 3, with results for the most common subgroup also varying greatly: >40 (56.2%) [23], 35–51 (22.4%) [17], <20 (89%) [18], and 0–5 (48.7%) [19]. One study [20] reported nevus density per body site based on body surface area calculations proposed by Lund and Browder [26].

Inter- and Intraobserver Variation

Expertise of the observers: Four studies [17-20] reported that counting was done only by dermatologists, one study [23] by both dermatologists and physicians, and three studies [16, 21, 22] by nurses (Table 4).

Training of observers: All studies where dermatologists counted nevi did not mention any training for the observers [17-20], and one study [19] did not provide the number of examiners. In the skin cancer screening study in Germany [23], a total of 116 dermatologists and 1,673 non-dermatologist physicians were involved in nevus counting, where each was given an 8-hour training course with a detailed description of the screening procedure and standard whole-body examination with practical examples. When nurses were used to count nevi [16, 21, 22], they were trained by dermatologists, but no training criteria were specifically reported.

Validity and reproducibility of nevus counts: While only one dermatologist examined all participants in the study by Karlsson et al. (2000) [20], counting was validated using independent examination of 48 participants by another observer who had participated in nevus counting in a previous study [27], and their agreement was substantial

Table 1. Characteristics of the Included Studies (n=8).

First Author (Year) & Country	Type of Study	Main Objective of the Study	Sample Size (Female Percentage)	Age (Average Age and Range)	Period of Study
Bataille (2000), UK	Cross-sectional	To investigate the relative contribution of genetic and environmental effects on the expression of nevi and freckles, and to determine if age and sun exposure influence the heritability of nevi.	900 (100%)	Adults (for the monozygotic twin pairs average 48 and range 19-73, for the dizygotic pairs average 44 and range 18-69)	Jan 1997-Dec 1997
Karlsson (2000), Sweden	Cross-sectional	To assess the variation in naevus profile in regions with different levels and patterns of sun exposure and different melanoma incidence, and to investigate the prevalence of common naevi and dysplastic naevi in a Swedish population with a low incidence of melanoma	201 (53%)	Adults (39.8, 30-50)	ND
Naldi (2000), Italy	Case-control	To identify the risk factors of melanoma	542 cases (58%) and 538 controls (ND)	Adults (Median 54, ND)	1992-1994
Landi (2001), Italy	Case-control	To identify risk factors for non-familial melanoma	183 cases (53%) and 179 controls (50%)	Children and adults (Cases average 48.5, controls average 45.8, 17-77)	Not defined specifically - recruited newly diagnosed melanoma cases between Dec 1994 and Jan 1999
Silva (2009), UK	Cross-sectional	Investigating the effect of foreign sun exposure on nevus counts and skin aging among young white women living in temperate climates like the UK	754 (100%)	Adults (36.2, 18-46)	ND
Nagore (2010), Spain	Case-control	To elucidate what factors are associated with melanoma patients older than 60 years	160 cases (46%) and 318 controls (48%)	Adults (68.6, >=60)	ND
Newton-Bishop (2010), UK	Case-control	The risk associated with nevus phenotype was investigated in relation to patterns of sun exposure and the inheritance of 3 single nucleotide polymorphisms (SNPs) on chromosomes 6, 9 and 22	960 cases (ND), 513 controls (ND)	Adults (ND, 18-76)	Cases: Sep 2000-Dec 2005
Breitbart (2012), Germany	Cross-sectional	Description of the implementation of a skin cancer screening project and its feasibility in the existing setting of primary care in Germany	360,288 (75%)	Adults (49.7, >20)	July 2003-June 2004

ND: not disclosed.

Table 2. Methodologies Used for Identifying and Counting Nevi.

First Author (Year) & Country	Protocols Considered		Methodologies Used for the Identification and Counting Of Naevi								Terminology for Nevi	Body Sites Considered
	Was IARC Used	Were Any Other Protocols Used	Types of Nevi Considered	Criteria for Identifying a Lesion as a (Particular Type of) Nevi	Categorization of Size for Analysis	Colors Of Lesions Considered as Nevi	Tools Used to Measure Nevi	Tools Used to Record Location of Nevi	Steps Taken to Make Sure Nevi were not Confused with Other Similar Lesions	Atypical Nevi Counted and Analyzed Separately		
Bataille (2000), UK	ND	A prevalidated protocol (Grulich 1996)	ND	ND	3 categories: $2<x\leq5$, $5<x\leq10$ and >10 mm	ND	ND	Predefined body sites	ND	Yes	Nevi	Whole body, excluding genitalia and posterior scalp
Karlsson (2000), Sweden	ND	ND	ND	ND	ND	ND	ND	Predefined body sites	ND	Yes	Common nevi	Whole body, excluding genitalia
Naldi (2000), Italy	ND	Yes - but not specified	ND	ND	ND	ND	Nevometer	ND	Providing clear definitions for other lesions such as freckles and solar lentigines	Yes	Nevi	Whole body, excluding genitalia and scalp
Landi (2001), Italy	ND	ND	ND	ND	ND	Pigmented	ND	Digital records using photographs	Providing clear definitions for other lesions such as freckles and solar lentigines	Yes	Nevi	Whole body, excluding genitalia
Silva (2009), UK	ND	ND	ND	ND	ND	ND	ND	ND	ND	Yes	Nevi	Whole body, excluding breasts in females and genitalia

Nagore (2010), Spain	ND	ND	ND	ND	ND	ND	ND	ND	Providing clear definitions for other lesions such as freckles and solar lentigines	No	Melanocytic	ND
Newton-Bishop (2010), UK	ND	ND	ND	ND	≥ 5 mm	ND	ND	Predefined body sites	ND	Yes	Nevi	Whole body, excluding breasts in females and genitalia
Breitbart (2012), Germany	ND	Yes	ND	ND	ND	ND	ND	ND	Examiners' ability to correctly distinguish between benign and malignant lesions was tested after the training	Yes	Melanocytic	Whole body

ND: not disclosed.

Table 3. Nevus-related Outcomes Reported.

First Author (Year) & Country	Outcome Reported					Nevus Prevalence	
	Nevus Count Reported	Categorized Nevus Count	Nevus Count per Unit Body Surface Area	Nevus Count per Body Site	Anatomical Site Area Calculations	Average	Dispersion
Bataille (2000), UK	Yes	Yes	ND	ND	ND	Mean=35, Median=22	Range=0-324
Karlsson (2000), Sweden	Yes	Yes	Yes, for each site surface area separately	Yes	Yes, size of each area expressed as a percentage of the total body surface, calculated according to Lund & Browder 1944	Median=15	Range=0-332
Naldi (2000), Italy	No	Yes	ND	ND	ND	0-5 (48.7%), 6-15 (27.9%), 16-30 (14.3%), 31-45 (5.8%), >= 46 (3.3%)	ND
Landi (2001), Italy	No	Yes	ND	ND	ND	0-12 (21.2%), 13-20 (18.8%), 21-34 (17.7%), 35-51 (22.4%), 52-190 (20%)	ND
Silva (2009), UK	Yes	Yes	ND	ND	ND	Median=45, Mean=57.6	Range=0-355
Nagore (2010), Spain	No	Yes	ND	ND	ND	< 20 (89%), 20-50 (8.8%), >50 (2.2%)	ND
Newton-Bishop (2010), UK	Yes	No	ND	ND	ND	Median*	Interquartile range*
Breitbart (2012), Germany	No	No	ND	ND	ND	>=40 nevi=56.2%	ND

*Reported separately for 11 sun exposure characteristics and 3 nevus genotypes. ND: not disclosed.

(Kappa: 0.79, 95% CI: 0.59-0.99). Landi et al. (2001) [17] used only one dermatologist to count nevi and an oncologist to validate diagnosis of nevi using photographs. Further, due to the difficulty in diagnosing nevi in older participants, the authors performed a sensitivity analysis removing participants older than 60 years and observed no change in their results. The Landi et al. study (2001) [17] was also the only one to digitally record nevi using photographs, hence increasing the reproducibility of nevus counts.

Assessment and/or quantifying inter- and intraobserver variability: The inter- and intraobserver variation of nevus counts was assessed only in three studies [19, 21, 22]. In the study by Silva et al. (2009) [21], each nurse independently examined five participants at 6-month intervals and reviewed their counts against the dermatologist's but did not report an inter- or intraobserver variation measure. Naldi et al. (2000) [19] assessed interobserver variation using a correlation coefficient reported as >0.75, but did not provide details of assessment criteria. Bataille et al. (2000) also used the correlation coefficient to assess interobserver variation using only two of the five research nurses (reported as 0.95 between the two nurses); they counted nevi of 20 participants at two different visits.

Table 4. Intra- and Interobserver Variation.

First Author (Year) & Country	Personnel Involved in Nevus Counting	Were Observers Trained against a Standard Criterion to Identify Nevi?	Validity of Counts and Reproducibility of the Study	Measuring Inter- and Intraobserver Variation		
				Assessment Criteria for Inter- and Intraobserver Variation	Measure Used	Value
Bataille (2000), UK	5 research nurses	Trained by a dermatologist, but no criteria mentioned	ND	Interobserver bias was assessed for only 2 observers, by counting nevi of 20 participants in 2 different visits.	Correlation coefficient	0.95
Karlsson (2000), Sweden	1 dermatologist	ND	One observer examined all participants and another observer (who was in the (Augustsson, Stierner et al. 1991 study) independently examined a subset of 48 subjects for validation. The agreement was: Kappa 0.79 (95% CI 0.59-0.99)	ND	ND	ND
Naldi (2000), Italy	Dermatologists, number undefined	Trained, but no protocol mentioned	ND	Assessed, but no details provided	Correlation coefficient	>0.75
Landi (2001), Italy	1 dermatologist	ND	Having only 1 dermatologist perform skin checks and 1 oncologist assess DN and nevi diagnoses using photographs.Sensitivity analysis excluding participants >60 years old	ND	ND	ND
Silva (2009), UK	3 nurses	Trained by a dermatologist, but no criteria mentioned	ND	Monitored at 6-month intervals. Each nurse independently counted nevi in the same 5 patients at 6-month intervals and reviewed their counts together with the dermatologist	Models adjusted for each observer	ND
Nagore (2010), Spain	2 dermatologists	ND	ND	ND	ND	ND
Newton-Bishop (2010), UK	3 nurses	Trained by a dermatologist, but no criteria mentioned	ND	ND	ND	ND

ND: not disclosed.

Critical Appraisal and Quality of the Studies

Based on the JBI Critical Appraisal Checklists for case-control and cross-sectional studies, all considered studies had good quality with respect to general research methodology, except for obtaining reliable and valid exposure measurements and outcome measures (Table S1). Exposure measurements reported in the considered studies included phenotypic characteristics, sun exposure habits, and personal and family history of melanoma (Table S2). Only two studies [16, 21] considered possible temporal changes in hair color and recorded the color at age 21. Recollection of the level of sun exposure during leisure time [16-18, 20, 21] or vacation [16, 18, 21] and the number of sunburns [17-19] throughout life may also not be reliable.

In the ranking criterion we developed to assess quality of nevus counting methodologies of studies and their reproducibility based on the recommendations of the IARC protocol (Table S3), the study by Landi et al. [17] received the highest score (64%), followed by Naldi et al. [19] (59%), Bataille et al. [22] (58%), and Nagore et al. [18] (50%). All other studies scored less than 50% in quality of nevus counting and reporting methodology and in reproducibility.

Discussion

This systematic review evaluated variations in nevus counting and reporting methods in studies published over the past 22 years and assessed the reproducibility and comparability of results. Given the previously observed significant heterogeneity in nevus counting methods [4], we limited our review to studies using skilled examiners for nevi counting, including nevi at least 2 mm in diameter on the total body in adults sampled from the general population. Only eight studies were eligible for evaluation, and even among them there was a considerable variation in counting methods. Overall, the included studies did not often follow a standard counting protocol and differed in their handling of inter- and intraobserver variation.

During the full text screening we found four skin cancer screening studies–in Italy [28], Germany [23, 29], and the whole of Europe [30]– of which we included only one [23] as the other studies did not mention the size of nevi counted. It has been reported that nevus count estimates obtained from hospital-based controls are more reliable because of more precise assessment of nevi in these studies [4]. However, for evaluated nevus counting and reporting criteria, we did not find a distinctly higher quality and reproducibility of methodology in the the hospital-based studies [17-19].

Future studies assessing the risk of melanoma or the prevalence/incidence of nevi should clearly specify the steps followed to count nevi. Most importantly, the reviewed studies were inconsistent in how they defined nevi considered for counting, which could significantly affect validity of nevus counts.

The IARC protocol defines a melanocytic nevus contributing to total body nevus counts as: (i) brown to black pigmented, (ii) macule or papule, (iii) reasonably well-defined, (iv) darker in color than the surrounding skin, and (iv) does not have the features of freckles, solar lentigines, seborrheic keratosis, café-au-lait spots, or non-melanocytic nevi. Only one study [17] reported or defined nevi according to criteria (i), (ii), and (iv). Two studies reported well-defined criteria for identifying freckles or solar lentigines; hence it could be assumed that they followed criterion (iv). The IARC protocol lists skin- or red-colored nevi, halo nevi, nevus spilus, congenital nevus-like nevi, blue nevi, café-au-lait spots, and atypical nevi as other melanocytic nevi not to be considered as countable nevi. None of the studies mentioned exclusion of these from the total nevus counts. While studies reported atypical nevi separately for analysis, they did not mention whether atypical nevi were excluded from the total number of nevi. A recent Delphi study, reporting recommendations from the International Dermoscopic Society for digital monitoring of patients with multiple nevi, defined a 'common nevus' as "being a macular or papular symmetrical lesion, 2-6 mm in diameter, uniform in colour, with well-defined borders and regular overall architecture in dermoscopy." While both common and atypical nevus counts are considered to be independent risk factors for melanoma [4], there is still no clear consensus on whether common nevi and atypical (≥5 mm) nevi should be pooled.

Some studies noted that restricting nevi to only a subset based on size may reduce validity and, with arbitrary cutoffs, reproducibility [31]. Counting nevi of all sizes might be acceptable for children, where confusion with other lesions is limited. However, considering very small <2 mm lesions in adults, specifically people with sun damage, could increase misclassifications [3].

While inter- and intraobserver variation and validity of nevus counts was reported or assessed in most studies, only one study [20] correctly used the kappa statistic to quantify level of agreement in nevus counts. Others used the correlation coefficient, which is not a good measure for quantifying level of agreement between two raters/counters as correlation refers to the existence of an association between two different variables [32]. Cohen Kappa for nominal variables and Bland-Altman plots and intra-class correlation coefficient for continuous variables should be used in future studies of nevus counts.

Further, in most cases when counting was done by an expert and/or a dermatologist, the method of identification of nevi was not specified. Yet, for results to be comparable,

nevus identification, counting, and reporting must be carried out according to a standard protocol or, at the very least, described in detail. For these reasons the IARC protocol was introduced in 1990. However, none of the studies evaluated in this review specifically reported that they had followed the IARC protocol, while one later publication [33] in 2006, based on one of the studies we included in this review [19] (published in 2000), reported that the IARC protocol had been considered for counting and reporting nevi. Only three studies had used some sort of standardized protocol for nevus counting. While the importance of following a standardized protocol for nevus counting is obvious, the lack of popularity of the 1990 IARC protocol for nevus counting is also evident. The reasons are unclear, but the protocol may not have been widely known, might have been too detailed to be deemed feasible, or might have been outdated with the increasing availability of imaging and digital technologies.

In the considered studies, published over the past 22 years, none used modern digital imaging systems to identify or record nevi. However, one study [17] used photographs taken of the back to validate counts of the examiner. Recent advances in digital imaging (e.g., 3D total body imaging, total body skin mapping) and artificial intelligencebased, computer-assisted, diagnosis software have afforded new opportunities for obtaining objective nevus counts that are not limited by the inter- and intraobserver variation inevitable with human observers. While there is the potential for nevus counting to be automated in future studies [34], once again the accuracy of these automated counts depends on the training data sets these algorithms are based on, and therefore the protocol used for the image labelling.

Conclusion

This review highlights the necessity of an updated protocol for identifying, counting, and reporting nevi in future studies. Incorporating emerging medical technologies harnessing artificial intelligence should be considered in an updated protocol for reproducible and objective nevus counting.

References

1. Arnold M, Singh D, Laversanne M, et al. Global Burden of Cutaneous Melanoma in 2020 and Projections to 2040. *JAMA Dermatology*. 2022;158(5):495-503. DOI: 10.1001/jamadermatol.2022.0160.
2. Jayasinghe D, Koh U, Plasmeijer EI, et al. The dynamic nature of naevi in adulthood: prospective population-based study using three-dimensional total-body photography. *Br J Dermatol*. 2023;188(3):437-9. DOI: 10.1093/bjd/ljac095.
3. English DR, Rivers J, Kelly JW, Armstrong BK. Epidemiological studies on melanocytic naevi: Protocol for identifying and recording naevi. Lyon France: International Agency for Research on Cancer; 1990. IARC Internal Report no.90/002 .
4. Gandini S, Sera F, Cattaruzza MS, et al. Meta-analysis of risk factors for cutaneous melanoma: I. Common and atypical naevi. *Eur J Cancer*. 2005;41(1):28-44. DOI: 10.1016/j.ejca.2004.10.015.
5. Jayasinghe D, Nufer KL, Betz-Stablein B, Soyer HP, Janda M. Body Site Distribution of Acquired Melanocytic Naevi and Associated Characteristics in the General Population of Caucasian Adults: A Scoping Review. *Dermatology and Therapy*. 2022. DOI: 10.1007/s13555-022-00806-x.
6. Walter SD, Marrett LD, Hertzman C. Reliability of interviewer and subject assessments of nevus counts in a study of melanoma. J Clin Epidemiol. 1991;44(7):633-40. DOI: 10.1016/0895-4356(91)90024-4.
7. English DR, Armstrong BK. Melanocytic Nevi in Children: II. Observer Variation in Counting Nevi. A*merican Journal of Epidemiology*. 1994;139(4):402-7. DOI: 10.1093/oxfordjournals.aje.a117012.
8. Green A, Swerdlow AJ. Epidemiology of melanocytic naevi. *Epidemiol Rev*. 1989;11(1):204-21. DOI: 10.1093/oxfordjournals.epirev.a036037.
9. Gallagher RP, McLean DI. The epidemiology of acquired melanocytic nevi. A brief review. *Dermatol Clin*. 1995;13(3):595-603. DOI: 10.1016/S0733-8635(18)30065-2.
10. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med*. 2009;151(4):264-9, w64. DOI: 10.7326/0003-4819-151-4-200908180-00135.
11. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev*. 2015;4(1):1. DOI: 10.1186/2046-4053-4-1.
12. Veritas Health Innovation M, Australia. Covidence systematic review software. 2021.
13. Joanna Briggs Institute. Checklist for Analytical Cross Sectional Studies. 2020. https://jbi.global/critical-appraisal-tools
14. Joanna Briggs Institute. Checklist for Case Control Studies. 2020. https://jbi.global/critical-appraisal-tools
15. Joanna Briggs Institute. Checklist for Cohort Studies. 2020. https://jbi.global/critical-appraisal-tools
16. Newton-Bishop JA, Chang YM, Iles MM, 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(8):2043-54. DOI: 10.1158/1055-9965.EPI-10-0233.
17. Landi MT, Baccarelli A, Calista D, et al. Combined risk factors for melanoma in a Mediterranean population. *British Journal of Cancer*. 2001;85(9):1304-10. DOI: 10.1054/bjoc.2001.2029.
18. Nagore E, Hueso L, Botella-Estrada R, et al. Smoking, sun exposure, number of nevi and previous neoplasias are risk factors for melanoma in older patients (60 years and over). *Journal of the European Academy of Dermatology and Venereology*. 2010;24(1):50-7. DOI: 10.1111/j.1468-3083.2009.03353.x.
19. Naldi L, Imberti GL, Parazzini F, et al. Pigmentary traits, modalities of sun reaction, history of sunburns, and melanocytic nevi as risk factors for cutaneous malignant melanoma in the Italian population - Results of a collaborative case-control study. *Cancer*. 2000;88(12):2703-10. DOI: 10.1002/1097-0142(20000615)88:12<2703::aid-cncr8>3.0.co;2-q.
20. Karlsson P, Stenberg B, Rosdahl I. Prevalence of pigmented naevi in a Swedish population living close to the arctic circle. *Acta Dermato-Venereologica*. 2000;80(5):335-9. DOI: 10.1080/000155500459259.

21. Silva ID, Higgins CD, Abramsky T, et al. Overseas Sun Exposure, Nevus Counts, and Premature Skin Aging in Young English Women: A Population-Based Survey. *J Invest Dermatol*. 2009; 129(1):50-9. DOI: 10.1038/jid.2008.190.
22. Bataille V, Snieder H, MacGregor AJ, Sasieni P, Spector TD. Genetics of risk factors for melanoma: an adult twin study of nevi and freckles. *J Natl Cancer Inst*. 2000;92(6):457-63. DOI: 10.1093/jnci/92.6.457.
23. Breitbart EW, Waldmann A, Nolte S, et al. Systematic skin cancer screening in Northern Germany. *J Am Acad Dermatol*. 2012;66(2):201-11. DOI: 10.1016/j.jaad.2010.11.016.
24. Grulich AE, Bataille V, Swerdlow AJ, et al. Naevi and pigmentary characteristics as risk factors for melanoma in a high-risk population: A case-control study in New South Wales, Australia. *Int J Cancer*. 1996;67(4):485-91. DOI: 10.1002/(SICI)1097-0215(19960807)67:4<485::AID-IJC4>3.0.CO;2-O.
25. Bataille V, Bishop JA, Sasieni P, et al. Risk of cutaneous melanoma in relation to the numbers, types and sites of naevi: a case-control study. *Br J Cancer*. 1996;73(12):1605-11. DOI: 10.1038/bjc.1996.302.
26. Lund C. The estimation of areas of burns. *Surgery Gynecology and Obstetrics*. 1944;79:352-8.
27. Augustsson A, Stierner U, Suurkula M, Rosdahl I. Prevalence of common and dysplastic naevi in a Swedish population. *Br J Dermatol*. 1991;124(2):152-6. DOI: 10.1111/j.1365-2133.1991.tb00424.x.
28. Seidenari S, Benati E, Ponti G, et al. Italian Euromelanoma Day Screening Campaign (2005-2007) and the planning of melanoma screening strategies. *Eur J Cancer Prev*. 2012;21(1):89-95. DOI: 10.1097/CEJ.0b013e3283498e14.
29. Schmitt J, Seidler A, Heinisch G, Sebastian G. Effectiveness of skin cancer screening for individuals age 14 to 34 years. *JDDG - Journal of the German Society of Dermatology*. 2011;9(8): 608-16. DOI: 10.1111/j.1610-0387.2011.07655.x.
30. Suppa M, Gandini S, Njimi H, et al. Association of sunbed use with skin cancer risk factors in Europe: an investigation within the Euromelanoma skin cancer prevention campaign. *Journal of the European Academy of Dermatology and Venereology*. 2019; 33:76-88. DOI: 10.1111/jdv.15307.
31. Green A, Siskind V, Hansen ME, Hanson L, Leech P. Melanocytic nevi in schoolchildren in Queensland. *J Am Acad Dermatol*. 1989;20(6):1054-60. DOI: 10.1016/s0190-9622(89)70131-6.
32. Ranganathan P, Pramesh CS, Aggarwal R. Common pitfalls in statistical analysis: Measures of agreement. *Perspectives in clinical research*. 2017;8(4):187-91. DOI: 10.4103/2229-3485.342656.
33. Randi G, Naldi L, Gallus S, Di Landro A, La Vecchia C. Number of nevi at a specific anatomical site and its relation to cutaneous malignant melanoma. *J Invest Dermatol*. 2006;126(9):2106-10. DOI: 10.1038/sj.jid.5700334.
34. Betz-Stablein B, D'Alessandro B, Koh U, et al. Reproducible Naevus Counts Using 3D Total Body Photography and Convolutional Neural Networks. *Dermatology*. 2021. DOI: 10.1159/000517218.