**Defining the validity of skin self-examination as a screening test for the detection of suspicious pigmented lesions: A meta-analysis of diagnostic test accuracy**

**Running title:** Determining the accuracy ofskin self-examination – a meta-analysis

Zainab Jiyad*1, 2*(MBBS, MDRes), Elsemieke I. Plasmeijer *3* (PhD), Samantha Keegan*1* (MB BChir), Venura Samarasinghe*1* (MBChB), Adele C Green*4,5* (PhD),Victoria Akhras*1* (MBBS, MDRes)

*1* Department of Dermatology, St. George’s Hospital, London, UK

*2* Population Health Research Institute, St. George’s University of London, UK

*3* The Netherlands Cancer Institute, Amsterdam, The Netherlands

*4* Department of Population Health, QIMR Berghofer, Queensland, Australia

*5* CRUK Manchester Institute and Faculty of Biology, Medicine and Health, University of Manchester, Manchester Academic Health Sciences Centre, Manchester, United Kingdom

**Abbreviations**: AUC: area under curve; CI: confidence interval; DOR: diagnostic odds ratio; FN: false negatives; FP: false positives; LR: likelihood ratio; sROC: summary receiver operator characteristic curve; SSE: skin self-examination; TN: true negatives; TP: true positives.

**Text word count (introduction to discussion):** 2397

**Tables:** 2; **Figures:** 5; **Supplementary tables:** 1.

**Keywords:** skin self-examination, melanoma, diagnostic accuracy, sensitivity, specificity

**Correspondence to:**

Dr Zainab Jiyad, Department of Dermatology, St. George’s Hospital, Blackshaw Road, London, SW17 0QT, UK. E-mail: zainabjiyad@nhs.net

**Key Message:** Skin self-examination is an easily available, minimally-invasive, preliminary screening test which has reasonable screening accuracy.

**ABSTRACT**

**Background:** Skin self-examination (SSE) is widely promoted for the detection of suspicious pigmented lesions. However, determining screening accuracy is essential to appraising the usefulness of SSE.

**Objectives:** To pool estimates from studies of SSE diagnostic accuracy in the detection of suspicious pigmented lesions.

**Methods:** This study was registered with PROSPERO (CRD42021246356) and conducted in accordance with PRISMA-DTA guidelines. A systematic search ofMedline (PubMed) EMBASE, CINAHL and The Cochrane Library was conducted to identify relevant studies. We included studies that examined the accuracy of SSE, either whole-body or site-specific, for detecting change in individual pigmented lesions or detecting an atypical naevus. A univariate random effects model, based on logit-transformed data, was used to calculate a summary diagnostic odds ratio (DOR) as well as pooled sensitivity and specificity. Cochran’s *Q* test and the *I2* statistic were calculated to assess heterogeneity. A proportional hazards model was used to calculate the area under the curve (AUC) and plot the summary receiver operator characteristic curve. We used the Quality Assessment of Diagnostic Accuracy Studies-2 tool to grade study quality.

**Results:** We identified 757 studies, of which three met inclusion criteria for quantitative synthesis. Pooled sensitivity and specificity based on 553 included participants was 59% and 82%, respectively. Summary DOR was 5.88 and the AUC was 0.71. There were some concerns regarding risk of bias in all three studies.

**Conclusions:** SSE can detect suspicious pigmented lesions with reasonable sensitivity and relatively high specificity, with the AUC suggesting acceptable discriminatory ability.

**WORD COUNT: 245**

**INTRODUCTION**

The incidence of melanoma has risen rapidly on a global scale.[1–3] In the UK, it is the fifth most commonly diagnosed cancer, with the incidence increasing 135% since the early 1990s.[4] In the United States melanoma is the leading cause of skin cancer-related deaths.[5] Crucially, early detection of melanoma propends an excellent prognosis.[6] However, the rising incidence places a considerable burden on healthcare services, with lengthy waiting times for dermatologist review posing a particular challenge.[7]

Given that around half of melanoma lesions are detected by patients themselves,[8–10] skin self-examination (SSE) represents an obvious strategy to help meet the growing need for early melanoma diagnosis. Encouraging results were reported from a case-control study in the US which enrolled 1199 participants, 650 of whom had been recently diagnosed with cutaneous melanoma and the remaining 549 were age- and sex-matched controls from the general population.[8] The authors found that SSE, although only practised by 15% of participants, was associated with a reduced risk of melanoma. Based on their estimates, they concluded that SSE could decrease melanoma mortality by 63%. Further, a multicentre cross-sectional study of 685 participants found that routine performance of SSE was associated with thinner superficial spreading melanoma.[11]

Subsequently, several studies have attempted to determine the accuracy of SSE, using a variety of methods. One of the earliest approaches was to compare self-assessment of naevus counts against physician counts, with variable concordance rates reported.[12,13] But such results do not address the clinical application of SSE for the purposes of detecting suspicious pigmented lesions. As such, attempts to provide summaries of sensitivity of SSE, which included naevus counting studies, have reported a widely varying range from 25-93%.[14] However, reported specificity evaluations are higher and more consistent, ranging from 83-97%.[14] However, no quantitative synthesis has integrated the existing diagnostic evidence to produce overall summary estimates.

Confusion regarding accuracy estimates of SSE is reflected in the differing recommendations from various authorities. The National Institute of Clinical Excellence recommends that patients diagnosed with a skin cancer should be given instruction about self-surveillance.[15,16] However, the U.S. Preventive Services Task Force concluded that in terms of counselling adults about SSE to prevent skin cancer, the current evidence is insufficient to recommend it.[17] From a dermatology perspective, both the British Association of Dermatology and the American Academy of Dermatology recommend regular performance of SSE in patients at high risk of melanoma.[18,19]

Determining overall accuracy of SSE in detecting suspicious pigmented lesions is of prime importance for two reasons: firstly, it can better inform and help standardise health authority and guideline recommendations, and secondly, it enables clinicians who recommend SSE to understand the likelihood of its successful conduct. Meta-analyses of diagnostic accuracy are a relatively recent development in meta-analytics which enable simultaneous analysis of a pair of outcome measures, such as sensitivity and specificity.[20] No prior studies have synthesised data to provide an estimate on the accuracy of SSE. Thus, we sought to identify and synthesise results from studies of diagnostic accuracy for suspicious pigmented lesions to provide an enhanced understanding of overall accuracy of SSE.

**METHODS**

This systematic review and meta-analysis was conducted in accordance with published Preferred Reporting Items for Systematic Reviews and Meta-Analyses – Diagnostic Test Accuracy (PRISMA-DTA) guidelines.[21] The protocol was registered with The International Prospective Register of Systematic Reviews (PROSPERO registration number: CRD42021246356).

*Search strategy*

We conducted a systematic search of Medline (PubMed) EMBASE, CINAHL and The Cochrane Library using the following search terms: “skin”, “melanoma”, “self examination”, “naev\*”, “nevi” and “nevus\*” and “mole”, through to May 2021. We also performed reference list searching and citation searching of included studies. We did not apply any language restrictions. All relevant studies were identified by two reviewers (ZJ and VA). ZJ examined titles and abstracts and VA checked included and excluded studies. Subsequently the full-texts of potential articles were assessed against inclusion/exclusion criteria. All excluded studies were recorded with reasons for exclusion. Disagreements and uncertainties were discussed and resolved by consensus agreement between investigators.

*Eligibility criteria*

We included studies that examined the accuracy of SSE, either whole-body or site-specific, for detecting change in individual naevi or detecting an atypical naevus, by any method. We excluded studies which utilised naevi counts as a method of determining accuracy as well as studies which included the detection of keratinocyte cancers in the assessment. The latter were specifically excluded to ensure homogeneity and thus allow pooling of results. The reference standard was either the ‘true’ nature of the lesion, where artificial modification was utilised, or a trained specialist examination. All studies that reported data which enabled a 2 X 2 contingency table of true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN) to be extrapolated, were included in this meta-analysis. Where studies reported both lesion-based and person-based measures, we preferentially selected the former. Further, where studies included a range of artificial modifications, for example a range of artificial changes in naevi size, we selected the minimal change used as being the most challenging to identify. If studies utilised methods to enhance SSE, we selected unenhanced SSE results (without photographic aid/mobile technology) to facilitate data synthesis. Studies that utilised photographs in lieu of performing a skin examination were excluded as photography does not reflect the real-world difficulties of examining the skin in areas that are hard to visualise, for example the back. We also excluded studies published only as abstracts and those involving children under the age of 18 years.

*Data extraction and quality assessment*

The following variables were extracted from each study and recorded in a data extraction table: author and publication year, participant demographics, outcome measure, reference standard, control definition, TN, FN, TP and FP. The methodological quality of selected studies was graded according the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool.[22] The assessment comprised four domains: patient selection, index test, reference standard and flow & timing. Both risk of bias and applicability (in terms of the test, patient sample and reference standard) were assessed. Results were summarised in a quality assessment table and traffic light plot.

*Statistical analysis*

Descriptive summary statistics with per study sensitivity, specificity, diagnostic odds ratio (DOR) and positive and negative likelihood ratios (LR) with 95% confidence interval (CI) were calculated. The DOR is a global measure of diagnostic accuracy and represents the ratio of the odds of positivity in subjects with disease relative to the odds in subjects without disease.[23] Per study sensitivity and specificity were presented as forest plots.

Owing to scarce study numbers, we used a univariate random effects model, based on logit-transformed data to calculate a summary DOR as well as pooled sensitivity and specificity. However, studies have shown that univariate and bivariate analyses produce similar results in meta-analyses of diagnostic accuracy, hence we confirmed our findings by applying a bivariate binomial model.[24] Cochran’s *Q* test and the *I2* statistic were calculated to assess heterogeneity.[25,26] Subgroup analysis of person-based, excluding lesion-based, assessments was performed to further examine heterogeneity. A proportional hazards model approach was used to calculate *θ*, the area under the curve (AUC) and plot the summary receiver operator characteristic curve (sROC). An AUC of 1.0 (100%) indicates perfect discriminatory ability for a diagnostic test. The sROC curve represents an overall summary of test performance and displays the trade-off between sensitivity and specificity. Publication bias was assessed by visual observation of a funnel plot, based on DORs.[27] All statistical analyses were performed in R version 4.0.2 (R Development Core Team 2013).

**RESULTS**

*Study selection*

The systematic search identified a total of 757 studies. After screening, 735 studies were excluded (Figure 1). The full-texts of 22 articles were assessed for eligibility and 19 studies were excluded with reasons for exclusion recorded (Supplementary Table 1). Three studies met inclusion criteria. The total number of participants from all studies included in the meta-analysis was 553.

*Characteristics of included studies*

Of the three studies included, two assessed diagnostic accuracy of SSE by artificially altering naevi using makeup and one assessed the ability to detect atypical naevi on the lower legs in females (Table 1). All three studies were cross-sectional in design.

The majority of participants in all studies were female. In the study by Muhn *et al.*, participants who were at increased risk of melanoma were recruited at a pigmented lesion clinic,[28] as were participants in the study by Oliveria *et al.*, which recruited patients with five or more dysplastic naevi.[29] In the study by Muhn *et al.*, high-risk status was defined as a significantly greater number of naevi than the general population (numbers requiring

monitoring with total body photographs or numbers that required clinic visits for assistance in skin examination), personal or family history of melanoma, and/or multiple dysplastic naevi. Further, participants had been taught SSE and had been practising for at least one year. The study by Titus-Ernstoff *et al.* was a Swedish population-based that recruited women aged 30-49 years who responded to a mailed health questionnaire.[30] The questionnaire included three colour photographs to demonstrate the appearance of atypical naevi.

*Characteristics of excluded studies*

The majority of studies were excluded because they did not assess accuracy of SSE, rather, they examined another feature such as SSE performance [31–37,37–42] (Supplementary Table 1). One study was excluded because it determined accuracy of SSE based on counts of naevi [13], and four studies were excluded because they used photos in-lieu of physical SSE [37,43–45]. The study by Stapleton et al. examined SSE accuracy, however we were unable to extract data to complete a 2 x 2 table [46]. The randomised controlled study by Janda et al. included all histologically diagnosed skin cancers in the sensitivity analysis and considered the identification of atypical naevi as a benign feature, in direct conflict with our other included studies [47].

*Quality assessment*

For all three studies, some aspect of the study design showed either high-risk of bias or insufficient information was given to determine risk (Figure 2). All studies scored better for applicability assessment than risk of bias (Table 2). Overall, in terms of applicability, we found that the included patients and settings matched the review question, as did the index test and its conduct/interpretation, and that the target condition as defined by the reference standard corresponded with the review question.

Certain information about reference standards was missing in all three studies, making it difficult to determine overall assessment in that domain. Information on the index test, however, was generally complete and showed low risk of bias.

*Data synthesis and publication bias*

Sensitivity ranged from 50-60%, whilst specificity varied from 62-96% (Figure 3). Positive LRs varied from 1.5-15.9 whilst negative LRs showed less variation, from 0.4-0.7 (Table 1). Pooled sensitivity was 59% (95% CI 54-63%) with higher pooled specificity of 82% (95% CI 53-95%). DOR estimates varied from 2.2-38.4 in individual studies, with summary DOR of 5.9 (0.7 – 49.5) (Figure 4). Significant heterogeneity was detected between studies using Cochran’s *Q* test and the *I2* statistic (Figures 3 and 4). Subgroup analysis of person-based studies, with the exclusion of the lesion-based study by Titus-Ernstroff *et al*, showed no significant heterogeneity. The AUC was 0.71, with the sROC curve displayed in Figure 5. A funnel plot, based on DORs, indicated the presence of publication bias.

**CONCLUSIONS**

Our synthesis of 553 patients in this meta-analysis showed that skin self-examination for the detection of suspicious pigmented lesions has a pooled sensitivity and specificity of 59% and 82%, respectively. We found the Area Under the Curve was 71%, which by conventional standards, indicates that a test has acceptable discriminatory ability.[48] The summary diagnostic odds ratio was 5.9 which can be interpreted as the odds of a suspicious lesion being detected on SSE being around six times the odds of a benign lesion being incorrectly identified as suspicious on SSE.

There is no accepted standard for validity of a screening test, although the perfect test should have 100% sensitivity and 100% specificity.[49] However, comparisons with other screening tests already deployed in the population show similar results. For example, the guaiac-based faecal occult blood screening test has approximately 50% sensitivity [50], whilst a mammogram has a sensitivity of 75% for the detection of breast cancer, with specificity of 90-95%.[51,52]

Perhaps of greater importance is that specificity was consistently higher in the three studies included. The implication here is that patients are generally good at recognising lesions considered entirely innocuous. However, determining whether a test is ‘good enough’ overall, largely depends on the clinical application. As a screening test for the general population, SSE would seem adequate when compared with similar screening tests already available. However, with a positive likelihood ratio ranging from 1.5-15.9 in included studies, one would be reluctant to suggest SSE for self-monitoring of an indeterminate naevus that has already been identified as requiring surveillance by a clinician.

A particular strength of our synthesis was the exclusion of studies that utilised naevus counts as a measure of SSE diagnostic accuracy, which is key to ensuring clinically translatable results. Although previous studies have used this as a proxy measure for SSE accuracy, we do not believe that simple naevus count assessments can address the issue of accuracy of SSE for detecting suspicious pigmented lesions. However, we recognise there were limitations in this meta-analysis. Firstly, we noted marked heterogeneity between studies, which on subgroup analysis suggests this was likely due to the combination of person- and lesion-based measures. However, we acknowledge there are several other heterogenous elements in the combined studies. Two of the studies looked at high-risk participants [28,29] whilst the third did not include any risk stratification [30]. SSE can be taught in a number of ways and the differing methods used in the included studies, likely add an element of heterogeneity (for example the study by Titus-Ernstoff et al. included photographic examples of atypical naevi, whilst participants in the Muhn et al. study received SSE teaching session). Moreover, we acknowledge heterogeneity introduced by the varying outcomes we have amalgamated, namely identifying change in artificially modified naevi vs. identifying unaltered atypical naevi. The study by Oliveria et al. had a markedly higher specificity compared with the other studies. However, this looked at artificially modified naevi, similar to the study by Muhn et al, which suggests, as per our subgroup analysis, that the difference is largely attributable to this being the only lesion-based study included.

Despite the heterogeneous elements, we believe this is an appropriate synthesis that provides clinically meaningful results, because incorrectly identifying a single lesion or several lesions in practice, constitutes similar overall failures of SSE. Further, we acknowledge the differing combined outcomes. Nonetheless, we would argue that data synthesis is appropriate here as it pragmatically reflects what SSE represents: a heterogenous test with many varying outcomes and clinical utilities including the identification of changing naevi atypical naevi and skin cancers. We also acknowledge the difficulties in devising a study that enables calculation of SSE accuracy. We would recommend further randomised studies such as that by Janda et al.[47], where participants perform SSE to identify abnormal lesions, but with the gold standard being a skin cancer specialist assessment (not histological), in order to provide reliable real-world estimates [47]. We believe that lesion-based studies will enable a better overall assessment and specifically we suggest that in order to evaluate SSE application in a general population, it would be useful to study non-risk-stratified cohorts. Whilst there are differing methods to determine SSE accuracy, we would suggest the identification of unmodified atypical naevi as the ideal method, as this correlates with the most widespread ‘real-world’ utilisation of SSE.

Several studies have examined methods of improving SSE, including a study by Janda *et al*.[47] who found that the use of mobile dermoscopy resulted in lower sensitivity (75%) when compared with naked-eye examination (88%). This study[47] was not included in this meta-analysis because all skin cancers, including keratinocyte cancers, were included in accuracy measures and atypical naevi were excluded. The results of Janda *et al*.[47] contrast with those of Oliveria *et al*. who noted an improvement in diagnostic accuracy when participants had access to baseline photographs.[29] However, the evidence consistently shows that rates of SSE performance can be improved with the aid of photography or smartphones.[53,54]

Our findings aside, the crucial question remains whether a screening test such as SSE should be recommended for use by the general population for detection of suspicious pigmented lesions. SSE, on face value, is a relatively straightforward method that has minimal cost and is acceptable to most. However, SSE, as with other screening tests, can lead to stress associated with the identification of a pigmented lesion that requires further evaluation, attendance at a specialist clinic and surgical morbidity, when a lesion is excised.[55–57] Given that the increase in melanoma diagnosis has shown no concrete translatable reduction in melanoma mortality, many have argued that population-based skin cancer screening ought to be abandoned.[58,59] In a recent publication, Welch *et al.* contended that the increase in melanoma diagnosis reflects diagnostic scrutiny and a tendency towards overdiagnosing melanoma histologically.[58] While we appreciate concerns regarding inflation of an existing trend towards overdiagnosis with the widespread application of SSE, the benefits of an easily available, minimally-invasive, screening test which has relatively reasonable screening accuracy, should be explored further. Additional studies on SSE application are essential to enable comparison of its relative benefits and accuracy in high risk vs low risk cohorts, as well as determining the effects of teaching SSE and methods of maximising this effect.

**AUTHOR CONTRIBUTIONS:**

ZJ – conceptualisation, data collection, statistical analysis, writing.

EP – methodology, writing, reviewing.

SK – conceptualisation, writing.

VS – conceptualisation, writing.

AG - methodology, writing, reviewing.

VA – conceptualisation, data collection, methodology, writing, reviewing.

**DATA AVAILABILITY:**

Publicly available datasets were used in this study.

**STATEMENT OF ETHICS:**

The paper is exempt from ethical committee approval as it is a meta-analysis of published studies

**CONFLICTS OF INTEREST:**

The authors have no conflicts of interest to declare.

**FUNDING SOURCES:**

None**.**

**REFERENCES**

1 Karimkhani C, Green AC, Nijsten T, Weinstock MA, Dellavalle RP, Naghavi M, et al. The global burden of melanoma: results from the Global Burden of Disease Study 2015. Br J Dermatol. 2017 Jul;177(1):134–40.

2 Sacchetto L, Zanetti R, Comber H, Bouchardy C, Brewster DH, Broganelli P, et al. Trends in incidence of thick, thin and in situ melanoma in Europe. Eur J Cancer. 2018;92:108–18.

3 Schadendorf D, van Akkooi ACJ, Berking C, Griewank KG, Gutzmer R, Hauschild A, et al. Melanoma. Lancet. 2018 Sep;392(10151):971–84.

4 Melanoma skin cancer incidence statistics | Cancer Research UK [Internet]. [cited 2021 Apr 17]. Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/melanoma-skin-cancer/incidence#ref-2

5 Melanoma: Statistics | Cancer.Net [Internet]. [cited 2021 Apr 17]. Available from: https://www.cancer.net/cancer-types/melanoma/statistics

6 Weiss S, Hanniford D, Hernando E, Osman I. Revisiting determinants of prognosis in cutanous melanoma. Cancer. 2015;121(23):4108–23.

7 Tsang MW, Resneck JS. Even patients with changing moles face long dermatology appointment wait-times: A study of simulated patient calls to dermatologists. J Am Acad Dermatol. 2006 Jul;55(1):54–8.

8 Berwick M, Begg CB, Fine JA, Roush GC, Barnhill RL. Screening for cutaneous melanoma by skin self-examination. J Natl Cancer Inst. 1996;88(1):17–23.

9 Carli P, De Giorgi V, Palli D, Maurichi A, Mulas P, Orlandi C, et al. Self-detected cutaneous melanomas in Italian patients. Clin Exp Dermatol. 2004 Nov;29(6):593–6.

10 Koh HK, Miller DR, Geller AC, Clapp RW, Mercer MB, Lew RA. Who discovers melanoma? Patterns from a population-based survey. J Am Acad Dermatol. 1992;26(6):914–9.

11 Dessinioti C, Geller AC, Stergiopoulou A, Swetter SM, Baltas E, Mayer JE, et al. Association of Skin Examination Behaviors and Thinner Nodular vs Superficial Spreading Melanoma at Diagnosis. JAMA Dermatol. 2018;154(5):544–53.

12 Lawson DD, Schneider JS, Sagebiel RW. Nevus counting as a risk factor for melanoma: Comparison of self-count with count by physician. J Am Acad Dermatol. 1994 Sep;31(3):438–44.

13 Buettner PG, Garbe C. Agreement between Self-Assessment of Melanocytic Nevi by Patients and Dermatologic Examination. Am J Epidemiol. 2000 Jan;151(1):72–7.

14 Hamidi R, Peng D, Cockburn M. Efficacy of skin self-examination for the early detection of melanoma. Int J Dermatol. 2010;49(2):126–34.

15 NICE. Improving outcomes for people with skin tumours including melanoma. 2006

16 Melanoma: assessment and management NICE guideline. 2015; [cited 2021 Apr 17].Available from: www.nice.org.uk/guidance/ng14

17 Grossman DC, Curry SJ, Owens DK, Barry MJ, Caughey AB, Davidson KW, et al. Behavioral counseling to prevent skin cancer: US preventive services task force recommendation statement. JAMA - J Am Med Assoc. 2018 Mar;319(11):1134–42.

18 Swetter SM, Tsao H, Bichakjian CK, Curiel-Lewandrowski C, Elder DE, Gershenwald JE, et al. Guidelines of care for the management of primary cutaneous melanoma. J Am Acad Dermatol. 2019 Jan;80(1):208–50.

19 Marsden JR, Newton-Bishop JA, Burrows L, Cook M, Corrie PG, Cox NH, et al. Revised U.K. guidelines for the management of cutaneous melanoma 2010. Br J Dermatol. 2010 Aug;163(2):238–56.

20 Reitsma JB, Glas AS, Rutjes AWS, Scholten RJPM, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol. 2005 Oct;58(10):982–90.

21 McInnes MDF, Moher D, Thombs BD, McGrath TA, Bossuyt PM, Clifford T, et al. Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies The PRISMA-DTA Statement. JAMA - J Am Med Assoc. 2018 Jan;319(4):388–96.

22 Whiting PF. QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies. Ann Intern Med. 2011 Oct;155(8):529.

23 Glas AS, Lijmer JG, Prins MH, Bonsel GJ, Bossuyt PMM. The diagnostic odds ratio: A single indicator of test performance. J Clin Epidemiol. 2003;56(11):1129–35.

24 Foxlee N, Stone JC, Doi SAR. A comparison of univariate and bivariate models in meta-analysis of diagnostic accuracy studies. Int J Evid Based Healthc. 2015;13(1):28–34.

25 Cochran WG. The Combination of Estimates from Different Experiments. Biometrics. 1954;10(1):101–29.

26 Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. Bmj. 2003;327(7414):557–60.

27 Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. J Clin Epidemiol. 2005 Sep;58(9):882–93.

28 Muhn CY, From L, Glied M. Detection of artificial changes in mole size by skin self-examination. J Am Acad Dermatol. 2000 Jan;42(5):754–9.

29 Oliveria SA, Chau D, Christos PJ, Charles CA, Mushlin AI, Halpern AC. Diagnostic accuracy of patients in performing skin self-examination and the impact of photography. Arch Dermatol. 2004;140(1):57–62.

30 Titus-Ernstoff L, Thörn M, Tosteson TD, Brahme EM, Yuen J, Baron JA, et al. The accuracy of skin self-examination for atypical nevi. Epidemiology. 1996;7(6):619–23.

31 SB G, GC R, RL B. Sensitivity and specificity of self-examination for cutaneous malignant melanoma risk factors. Am J Prev Med. 1993 [cited 2021 Oct 17]. ;9(1):50–4.

32 DL P, SA O, PJ C, SW D, AC H. Skin self-examination in patients at high risk for melanoma: a pilot study. Oncol Nurs Forum. 2003;30(6):1029–36.

33 MA W, FQ N, RA M. Enhancing skin self-examination with imaging: evaluation of a mole-mapping program. J Cutan Med Surg. 2004 Jan;8(1):1–5.

34 MA W, RA M, PM R, M B, T L, W R, et al. Thorough skin examination for the early detection of melanoma. Am J Prev Med. 1999 Oct;17(3):169–75.

35 FM W, MM P, ME B, K M, CL S, P M, et al. Effect of a Skin Self-monitoring Smartphone Application on Time to Physician Consultation Among Patients With Possible Melanoma: A Phase 2 Randomized Clinical Trial. JAMA Netw open. 2020 Feb;3(2). DOI: 10.1001/JAMANETWORKOPEN.2020.0001

36 RA P, AC G, DR B, TM J, ER P, SM S. Efficacy of skin self-examination practices for early melanoma detection. Cancer Epidemiol Biomarkers Prev. 2009 Nov;18(11):3018–23.

37 JK R, R G, B H, S E, H B, J S, et al. Skin self-examination education for early detection of melanoma: a randomized controlled trial of Internet, workbook, and in-person interventions. J Med Internet Res. 2014;16(1). DOI: 10.2196/JMIR.2883

38 JK R, JD W, MC M, BA H, KA M, R T. Early Detection of New Melanomas by Patients With Melanoma and Their Partners Using a Structured Skin Self-examination Skills Training Intervention: A Randomized Clinical Trial. JAMA dermatology. 2016 Sep;152(9):979–85.

39 CR R, A V, G B, G M, MC M, S M, et al. Early detection of melanoma: an educational campaign in Padova, Italy. Melanoma Res. 2000 [cited 2021 Oct 17]. ;10(2):181–7.

40 D R, KH N, P B, M J. Evaluation of a Skin Self-Examination Programme: a Four-Stage Recursive Model. Asian Pac J Cancer Prev. 2017 Apr;18(4):1063–7.

41 MA W, PM R, RA M, W R, C D, M B, et al. Melanoma early detection with thorough skin self-examination: the “Check It Out” randomized trial. Am J Prev Med. 2007 Jun;32(6):517–24.

42 MA W, PM R, RA M, W R, KJ S, M B, et al. Reliability of assessment and circumstances of performance of thorough skin self-examination for the early detection of melanoma in the Check-It-Out Project. Prev Med (Baltim). 2004 Jun;38(6):761–5.

43 V C, E W, M M, MA W. The use of mole-mapping diagrams to increase skin self-examination accuracy. J Am Acad Dermatol. 2006 Aug;55(2):245–50.

44 John KK, Jensen JD, King AJ, Ratcliff CL, Grossman D. Do Pattern-Focused Visuals Improve Skin Self-Examination Performance? Explicating the Visual Skill Acquisition Model. J Health Commun. 2017;22(9):732–42.

45 King AJ, Gehl RW, Grossman D, Jensen JD. Skin self-examinations and visual identification of atypical nevi: Comparing individual and crowdsourcing approaches. Cancer Epidemiol. 2013;37(6):979–84.

46 JL S, R T, KA M, JK R. Correspondence between pigmented lesions identified by melanoma patients trained to perform partner-assisted skin self-examination and dermatological examination. Cancer Epidemiol Biomarkers Prev. 2015 Aug;24(8):1247–53.

47 Janda M, Horsham C, Vagenas D, Loescher LJ, Gillespie N, Koh U, et al. Accuracy of mobile digital teledermoscopy for skin self-examinations in adults at high risk of skin cancer: an open-label, randomised controlled trial. Lancet Digit Heal. 2020;2(3):e129–37.

48 Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. J Thorac Oncol. 2010 Sep;5(9):1315–6.

49 Power M, Fell G, Wright M. Principles for high-quality, high-value testing. Evid Based Med. 2013 Feb;18(1):5–10.

50 Lansdorp-Vogelaar I, Van Ballegooijen M, Boer R, Zauber A, Habbema JDF. A novel hypothesis on the sensitivity of the fecal occult blood test: Results of a joint analysis of 3 randomized controlled trials. Cancer. 2009 Jun;115(11):2410–9.

51 Elmore JG, Armstrong K, Lehman CD, Fletcher SW. Screening for breast cancer. J Am Med Assoc. 2005 Mar;293(10):1245–56.

52 Mushlin AI, Kouides RW, Shapiro DE. Estimating the accuracy of screening mammography: A meta-analysis. Am J Prev Med. 1998 Feb;14(2):143–53.

53 Marek AJ, Chu EY, Ming ME, Khan ZA, Kovarik CL. Impact of a smartphone application on skin self-examination rates in patients who are new to total body photography: A randomized controlled trial. J Am Acad Dermatol. 2018 Sep;79(3):564–7.

54 Oliveria SA, Dusza SW, Phelan DL, Ostroff JS, Berwick M, Halpern AC. Patient adherence to skin self-examination: Effect of nurse intervention with photographs. Am J Prev Med. 2004;26(2):152–5.

55 Johansson M, Jørgensen KJ, Brodersen J. The benefits of screening–and its harms. Lancet. 2016 Aug;388(10044):563–4.

56 Bibbins-Domingo K, Grossman DC, Curry SJ, Davidson KW, Ebell M, Epling JW, et al. Screening for skin cancer US preventive services task force recommendation statement. JAMA - J Am Med Assoc. 2016 Jul;316(4):429–35.

57 Results - Screening for Skin Cancer in Adults - NCBI Bookshelf [Internet]. [cited 2021 Aug 16]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK379859/#ch3.s23

58 Welch HG, Mazer BL, Adamson AS. The Rapid Rise in Cutaneous Melanoma Diagnoses. N Engl J Med. 2021 Jan;384(1):72–9.

59 Rubin R. Melanoma Diagnoses Rise while Mortality Stays Fairly Flat, Raising Concerns about Overdiagnosis. JAMA - J Am Med Assoc. 2020 Apr;323(15):1429–30.

**Figure 1: Flowchart of study selection**

**Figure 2: Risk of bias in studies included in the quantitative analysis, assessed using the QUADAS-2 tool**

**Figure 3: Forest plots of a) sensitivity and b) specificity of skin self-examination**

**Figure 4: Forest plot of diagnostic odd ratios (DORs) of included studies, with summary DOR**

**Figure 5: Summary receiver operator characteristic curve (sROC)**