Title: A systematic review of the use of teledermatology for the diagnosis and management of skin cancer

Corresponding Author

Dr Anna Finnane PhD
Dermatology Research Centre
The University of Queensland, School of Medicine
Level 5, Translational Research Institute
37 Kent Street
Woolloongabba QLD 4102
Phone: +617 3443 7000
Email: a.finnane@uq.edu.au
Disclosure statement: Nil disclosures

Authors

Name & qualifications: Ms Kathy Dallest, GradCertBusAdmin
Affiliation/s: Centre for Online Health, The University of Queensland, School of Medicine, Princess Alexandra Hospital, Brisbane, Australia
Disclosure statement: Nil disclosures

Name & qualifications: Prof Monika Janda PhD
Affiliation/s: School of Public Health and Social Work, Institute of Health and Biomedical Innovation, Queensland University of Technology, Brisbane, Australia
Disclosure statement: Dr Monika Janda was funded by NHMRC Career development Fellowship 1045247

Name & qualifications: H. Peter Soyer MD FACD
Affiliation/s: Dermatology Research Centre, The University of Queensland, School of Medicine, Translational Research Institute, Brisbane, Australia
Dermatology Department, Princess Alexandra Hospital, Brisbane, Australia
Disclosure statement: H. Peter Soyer is a shareholder of e-derm consult GmbH and MoleMap by Dermatologists Ltd Pty. He provides teledermatological reports regularly for both companies.

Word count (Manuscript): 3497 words
Number of Tables: 5
Number of Figures: 0 (1 supplementary figure)

Abstract

Importance

As technology becomes more commonplace in dermatological practice, it is essential to continuously review the accuracy of teledermatology devices and services compared with in-person care. The last systematic review was conducted over 5 years ago.

Objective

To synthesise and assess the quality of the evidence to address three research questions: 1) How accurate is teledermatology for skin cancer diagnosis compared to usual care (FTF diagnosis)?; 2) Does teledermatology save clinician and/or patient time, compared with usual care?; and 3) What are the enablers and barriers to adoption of teledermatology in clinical practice, for the diagnosis of skin cancer?

Evidence Review

The review protocol was registered on PROSPERO database. Six databases were searched for studies investigating the diagnostic accuracy and concordance, management accuracy and concordance, measures of time (waiting times, delay to diagnosis), and enablers and barriers to implementation. Potentially eligible articles were screened by 2 reviewers. The Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool was used to evaluate the risk of bias and applicability of individual studies assessing diagnostic accuracy.

Findings

21 studies were reviewed. The diagnostic accuracy (defined as agreement with histopathology for excised lesions or clinical diagnosis for non-excised lesions) of face to face (FTF) dermatology consultation remains higher (67-85% agreement with reference

standard, k=0.90) when compared with teledermatology (51%-85% agreement with reference standard, k=0.41-0.63), for the diagnosis of skin cancer. However, some studies do report high accuracy of teledermatology diagnoses. The majority of the studies of diagnostic accuracy and concordance had significant methodological limitations. Studies of health service outcomes found teledermatology reduced waiting times and could result in earlier assessment and treatment. Patients reported high satisfaction and were willing to pay out of pocket for access to such services.

Conclusions and Relevance

Robust implementation studies of teledermatology are needed, paying careful attention to reducing risk of bias when assessing diagnostic accuracy. Teledermatology services consistently reduced waiting times to assessment and diagnosis and patient satisfaction was high; it is now time to explore optimal pathways to integrate teledermatology into clinical practice.

Introduction

Early diagnosis and treatment of melanoma and non-melanoma skin cancers improves prognosis [1]. As rates of skin cancers increase, there is greater pressure on the dermatology workforce in both rural and urban areas.. Different forms of teledermatology have been explored as a solution to this growing problem[2,3] The two most common types of teledermatology are: store and forward (SAF) involving transfer of images and clinical information to a dermatologist for review at another time and location; and live interactive (LI), usually video-conferencing, which allows real-time interaction between the clinician and patient.

Before implementing a new mode of medical care like teledermatology, it is important to ensure the diagnostic accuracy is comparable to that of face to face (FTF) consultations, and that patient care is not compromised. For the diagnosis of skin lesions, this is not straightforward; when a biopsy is taken the reference standard is the histopathology result but when the lesion is considered benign, the clinical diagnosis by the dermatologist is accepted as the reference standard. In 2010, a US study of histopathology discordance in melanoma diagnosis reported discordant results in to 14.3% of cases (n=392) [4]. An earlier review of diagnostic accuracy in non-melanoma skin cancers reported discordant histopathology results (studies included 2-77 pathologists) in 2-7% of cases [5]. These discrepancies in the 'reference standard' have implications for studies of diagnostic accuracy and can also have clinical consequences for patients[4–6]. While histopathology is still the most accurate diagnostic method for skin cancers and remains the gold standard, it is important to take this margin of diagnostic discordance into account when setting thresholds for acceptable levels of diagnostic accuracy for new diagnostic tests.

A systematic review published in 2011 [7] found the accuracy of FTF dermatology was better than SAF teledermatology for diagnosing skin conditions (weighted mean difference = 11% for primary, 19% for aggregated diagnostic accuracy). However, the authors concluded the levels of diagnostic accuracy and concordance of both SAF and LI teledermatology were still acceptable compared with FTF dermatology [7]. Since that time, there has been significant growth in the number of devices, software and systems marketed for use by dermatologists, ranging from small dermoscopic attachments for mobile phones and digital cameras and associated 'mobile applications', to 3D imaging systems for high resolution full body photography. These technologies have the potential to improve access to specialist services, enable earlier diagnosis of skin cancers, and provide consumers and clinicians with a way of storing high quality images of lesions to support monitoring of any changes over time.

Five years on, this review is an updated synthesis and critique of the most recent studies of the use of teledermatology specifically for the diagnosis and management of skin cancer. We conducted a review of all studies published since June 2009 (cut-off date for inclusion in the previous review) addressing the following research questions: 1) How accurate is teledermatology for skin cancer diagnosis compared to usual care (FTF diagnosis)? 2) Does teledermatology save clinician and/or patient time, compared with usual care? and 3) Are there barriers to adoption of teledermatology in clinical practice, for the diagnosis of skin cancer?

Methods

The protocol was registered on PROSPERO International prospective register of systematic reviews and can be accessed at http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42015014295. The review adheres to the principles of the PRISMA Statement [8].

Search Strategy

We searched Cochrane, PubMed, Medline, Science Direct, Embase and Web of Science databases for systematic reviews and original research papers, restricted to human research published in English.

The search terms "remote consult" "remote consultation" "electronic mail" "telecommunications" "telemedicine" "teledermatology" "dermatology" "store and forward" "dermoscopy" "teledermoscopy" "teledermatoscopy" "skin cancer" "melanoma" "carcinoma" were combined using the appropriate methods for each database.

Study Selection

Studies were included if the primary focus was on the use of teledermatology or teledermoscopy interventions for diagnosing or managing melanoma or non-melanoma skin cancers. Only full papers were included. Specific exclusion criteria included:

1. Studies of teledermatology applications using image analysis software
2. Case studies and case series
3. Studies including participants under 18 years

Quality assessment

The Quality Assessment of Diagnostic Accuracy (2nd edition, QUADAS-2) was used to assess risk of bias in studies of diagnostic accuracy and diagnostic concordance, as well as applicability to the review question [9]. The tool has four domains including *patient selection*, *index test*, *reference standard*, and *flow and timing* (i.e., patient flow and timing of outcome assessments). In diagnostic accuracy studies, the "index test" is the test or intervention being

studied, while the "reference standard", otherwise known as the "gold standard" is the best available method to determine whether participants have the condition. Within these four domains, signalling questions are used to assess whether the risk of bias is low, high or unclear, and the applicability to the original review question is assessed [9]. Applicability to the review question is considered an important aspect of the quality assessment process; it is possible that even high quality, well-designed and reported studies with low risk of bias, differ from the review question in some fundamental way which reduces the generalisability of results. Results of the quality assessment process are presented graphically.

Data Synthesis

Two researchers (AF & KD) extracted data. The outcomes related to each review question are defined below:

1) How accurate is teledermatology for skin cancer diagnosis compared to usual care (FTF diagnosis)?

The main outcomes of diagnostic accuracy, diagnostic concordance, management accuracy and management concordance are defined in Table 1, specifying the relevant reference standard (current gold standard) and index test (comparator). Studies of diagnostic accuracy were separated into teledermatology and teledermoscopy subgroups.

2) Does teledermatology save clinician and/or patient time, compared with usual care?

Results relating to any aspects of clinician or patient time were extracted. This included time in days between referral and specialist consultation, time in days between referral and surgical excision of the lesions, and Breslow thickness as a proxy measure of delayed diagnosis.

3) What are the enablers and barriers to adoption of teledermatology in clinical practice, for the diagnosis of skin cancer?

Results from all studies that explored aspects of patient and clinician satisfaction and receptivity, feasibility of teledermatology or barriers to implementation were synthesised.

Results

The study selection process is detailed in Figure 1. Results from all database searches were combined and duplicates were removed. Titles and abstracts were reviewed and 153/241 studies were excluded for the following reasons: teledermatology for diagnosing or managing skin cancer was not a primary focus; study published prior to June 2009; conference abstract only or; full text not available in English. Of 88 full-text articles assessed, 41 were excluded due to criteria defined above, and 25 were published prior to June 2009.

Twenty eight papers met inclusion criteria but 7/28 were excluded due to insufficient reporting of methods to enable quality assessment [10–16] Characteristics of the 21 included studies are presented in Table 2.

Aim 1: Accuracy of teledermatology diagnosis compared to face to face diagnosis

Eight studies (seven including dermoscopy) reported the diagnostic accuracy of teledermatology consultations [17–24] (see Table 3). Three of the eight studies also assessed accuracy of FTF diagnosis compared to histopathological diagnosis, and compared the level of accuracy for FTF diagnoses with teledermatology diagnoses [18–20].

Three studies assessing diagnostic accuracy of TD using 133-188 clinical images (without dermoscopy) reported 68-85% agreement between TD diagnosis and reference standard [18,23], and sensitivity and specificity for detecting melanoma of 98% and 30%, respectively [17]. Five studies including dermoscopic or microscopic images in teledermatology consultations

(n=69-613 lesions) reported agreement between 51-92% [18,19,23], k=0.41-0.63 [20], and sensitivity and specificity for detecting melanoma of 96% and 62%, respectively [24].

One study reported very high sensitivity and specificity of TD for both malignant melanocytic lesions (sensitivity = 100%; specificity= 97-98%, n=6) and malignant non-melanocytic lesions (sensitivity = 97%, n= ; specificity= 92-94%, n=58), for both clinical and dermoscopic images (no significant differences between these methods) [18].

Studies that compared the diagnostic accuracy between teledermatology and FTF diagnoses reported variable results. Levels of agreement between teledermatology diagnoses and histopathological diagnoses were comparable to levels of agreement between FTF diagnoses and histopathological diagnoses in one study; 79-85% for TD (clinical images = 79%, dermoscopic images = 85%) and 85% for FTF [18]. However, in a second study that compared TD and FTF primary diagnoses as well as malignant/benign categorisation, the agreement was lower for TD (51-61% for TD versus 67% FTF for primary diagnoses and 75-80% for TD and 87% FTF for benign versus malignant) [19]. A third study reported much lower agreement between telediagnosis and histopathological diagnosis, compared with FTF diagnosis and histopathological diagnosis (k=0.41-0.63, TD, k=0.90, FTF)[20].

Four other studies reported diagnostic accuracy of teledermatology, without any comparison to FTF diagnoses. These studies reported agreement between telediagnosis and the reference standard of 51-100% [21–24]. When separating malignant and benign lesions, telediagnoses of malignant lesions were histopathologically confirmed in 62-100% cases, depending on the study. Of note, the study reporting 100% agreement between telediagnosis and histopathological diagnosis for malignant lesions only included 8 malignant lesions [22].

Concordance between teledermatology and face to face diagnoses

Ten included studies reported diagnostic concordance between teledermatology and FTF clinical diagnoses (see Table 3). These results were not histologically confirmed, so the FTF diagnosis is considered the reference standard.

Three studies without dermoscopy reported diagnostic concordance between 62-94%. Diagnostic concordance in studies including dermoscopic images ranged from 46-90% for primary diagnoses, or 71-91% for aggregated diagnostic categories. There were significant differences in diagnostic concordance between telederm and FTF clinical diagnoses when the same dermatologist performed both methods (k=0.95, (0.91-0.99) [25]) compared with studies involving different dermatologists (k=0.47-0.51) [19].

Of note, one study examined interobserver reliability between 5 teledermatologists and reported wide variation in levels of agreement (k=0.38-0.97) [26].

Quality of diagnostic accuracy and diagnostic concordance studies

The majority of studies of diagnostic accuracy and concordance of Teledermatology had significant methodological limitations. Many studies did not take (or report) adequate steps to reduce risk of selection bias, which could lead to overestimating the diagnostic accuracy (see Table 4). For example, enrolling high risk or excluding low quality images patients could lead to apparent higher sensitivity of Teledermatology than would be found in a general population group.

Other concerns with increased risk of bias included studies where the same dermatologist provided the FTF and telediagnosis; this may bias the index test if the telediagnosis was provided following the FTF diagnosis, or the reference standard, if vice versa (see Table 4).

There were no systematic differences between the results of studies that appeared to have taken steps to reduce risk of bias, compared to those with higher risk of bias. For example,

there were higher quality studies that reported both higher and lower levels of diagnostic accuracy, and the same was true of studies with significant limitations. This suggests the wide variation in results of diagnostic accuracy and concordance is not due to one consistent, identifiable type of bias but rather is a due to a combination of methodological limitations in the majority of studies in this field. These may be actual limitations, or important aspects of study design omitted in the reporting of studies, as evidenced by the large number of 'unclear' ratings (see Figure 2).

Accuracy of clinical management following teledermatology consultation and face to face consultation

Only two recent studies [19,21] assessed the management accuracy of teledermatology [19,21], measured by the level of agreement between recommended management based on teledermatological diagnosis and histopathological diagnosis (see Table 3). In one study [19], one melanoma *in situ* (1.5%, n=69) would have received no treatment based on the telediagnosis from one of the dermatologists in the study. The second dermatologist in that study made adequate management decisions in 100% of cases. The second study used teledermatology as a triage tool; 100% of patients with invasive melanoma (n=19) and 100% (n=16) of patients with melanoma *in situ* were prioritised appropriately as high, and at least medium priority, respectively. In the same study, 3 of 4 patients with invasive melanomas and 3 out of 5 patients with melanoma *in situ* referred by traditional paper referral were incorrectly given medium or low priority, and low priority, respectively.

Concordance between teledermatology and face to face clinical management

Six studies (3 with dermoscopy and 3 without) examined management concordance between teledermatology and FTF consultations (see Table 3). Agreement between management plans

(ie. decision to excise lesions, review in short term, long term or not review at all) ranged from 66%-85%, k=0.19-0.83 [20,23,27–30].

Aim 2: Time involved in teledermatology consultations compared with face to face consultations

<u>Breslow thickness</u>

One study reported Breslow thickness of melanoma as an indicator of earlier diagnosis and reported lower Breslow thickness in the teledermatology group compared to conventional referral (1.06mm vs 1.64mm, p=0.03, n=201) [31].

<u>Waiting time</u>

Four studies examined the effect of teledermatology consultations on waiting times to FTF appointments, waiting time to diagnosis and/or waiting time to surgery. Reductions in waiting times in teledermatology groups were reported in all studies.

Three studies specifically reporting waiting times for patients with melanoma[21,24,32] found patients in teledermatology groups waited significantly shorter periods than those with conventional referrals. The actual average waiting periods varied significantly between studies, from 9 days (TD) compared to 14 days (paper referral), 9 days (TD) versus 26.5 days (FTF clinic) and 36 (photo triage) compared to 39 days (conventional referral). Patients who were referred using teledermatology triage systems more often received primary treatment in a single dermatology appointment and required fewer appointments for treatment [21,32].

Aim 3: Patient satisfaction and barriers to teledermatology

Eleven included studies explored patient satisfaction, clinician factors and/or a range of barriers to teledermatology. The heterogeneous nature of these studies and the outcome measures allowed only a descriptive summary.

Patient satisfaction

Four studies examined measures of patient satisfaction, including confidence and use of teledermatology, satisfaction with waiting times, preferences and willingness to pay.

In 2/11 studies consumers submitted their own images (with and without dermoscopy) and reported they were satisfied with the ease of use (100%, n= 55) [27] and willing to pay out of pocket costs for teledermatology services[33]. Economic modelling was used to estimate willingness to pay [33], suggesting consumers would pay an average of AUD110 to have teledermoscopy review as a health service option, in addition to the currently available options of skin self-examination, skin cancer clinic or GP clinic. This concept was supported by a US study that reported patients were willing to pay USD20-500 for a mobile dermatoscope (median=USD100) [34].

Other benefits of teledermatology and teledermoscopy reported by patients included shorter waiting times, more frequent monitoring and improved privacy and comfort. One study measuring patient satisfaction using a 5 point-satisfaction scale reported higher satisfaction with waiting times (2.9 FTF, 4.0 TD) and overall satisfaction compared to FTF (3.8 FTF, 4.5 TD)[35].

Clinician-reported diagnostic difficulty and diagnostic confidence

One study[19] reported higher diagnostic difficulty for teledermatology consultations versus FTF consultations. Clinicians using teledermatology reported 61-87% cases as high difficulty compared to the clinician seeing patients face to face, who reported 54% as high difficulty

diagnoses. Another study examining diagnostic confidence of clinicians found diagnostic agreement increased as the clinicians diagnostic confidence increased[30].

Image quality

Studies [20,27,36]reported up to 8% of images as being unevaluable or unacceptable quality , but this proportion increased to 36% of 'bad quality images' when clinicians were asked to rate image quality as good, reasonable or poor. Of note, a large study of 959 images [22] found telediagnosis was possible in 99.7% of cases, and only 1% of dermoscopic and 4% of clinical images were rated as low quality.

Inter-observer reliability

Only one included study [25] reported the inter-observer reliability of teledermatologists as moderate for diagnostic group (k=0.56-0.78) and low for management plans (k=0.31-0.38).

Self-monitoring

An Australian study (n=49) examining the introduction of mobile dermoscopy into current skin self-examination recommendations reported barriers to effective self-monitoring [37]. In this study, the FTF dermatologist identified 40 lesions of concern on 25 people, which had not been identified during skin self-examination. Of these 40 lesions, 24 did not meet the AC rule (Asymmetry and Colour) communicated to consumers as a method for identifying concerning lesions. However, none of these lesions were subsequently diagnosed as melanoma.

Discussion

Five years after the last systematic review of teledermatology for the diagnosis and management of skin conditions, including skin cancer [7], the conclusion remains the same; the

accuracy of FTF dermatology consultation is generally higher than teledermatologyHowever, some studies in this review did report high accuracy of teledermatology diagnoses for skin cancer. Addressing the limitations of previous research will help to determine whether teledermoscopy is a safe and appropriate alternative to in-person assessment, which is particularly important for countries with high rates of skin cancers and geographically dispersed populations, including Australia and the US.

Future researchers in this field should aim to overcome the methodological limitations including lack of histopathology as reference standard, sample and diagnostic bias. A cross-over trial of tele and FTF diagnosis could be beneficial, with different clinicans providing the tele and FTF diagnosis, before switching into the other arm. Carefully designed, rigorous diagnostic studies, could help to identify whether teledermatology is equally or more accurate for diagnosing particular types of lesions, and whether the variation seen in accuracy of teledermatology is due to differences in clinical opinions, or some aspect of the teledermatology technology or process itself. Additional recommendations based on this review are presented in Table 5.

A common challenge in studies of telemedicine interventions is separating the effect of the intervention from other factors influencing the clinical outcomes. The most methodologically sound way to test the diagnostic accuracy of teledermatology would be assigning different clinicians for the tele- and FTF consultations, to prevent bias resulting from recall of the lesions and associated diagnosis the second time they see them. However, it is necessary to first ensure the clinicians have high inter-rater reliability, which may be more likely when clinicians have had similar training and are equally as experienced as dermatologists. Without this, it is very difficult to tell whether the limited agreement in diagnoses is related to the use of the technology itself, or differences in clinical opinion which could ordinarily exist in practice, as suggested by the variation in inter-observer reliability reported in studies

comparing diagnoses and management plans from multiple dermatologists [20,25] and previous studies in dermatopathology[4–6].

The nomenclature used by clinicians and researchers can influence measures of accuracy. This has been recognised and is currently being addressed by the International Skin Imaging Collaboration [39]. This aside, even dermatologists in the same countries who have received different training may describe the same lesion differently. If the researchers are not well versed in dermatology terminology, this variation in nomenclature could inadvertently result in underestimation of diagnostic accuracy.

While diagnostic accuracy is important, its relevance is lessened as long as the patient receives the same treatment, for example if the lesion is recommended for excision regardless. Studies included in this review focusing on agreement between prescribed management plans from teleconsultations versus FTF consultations suggest the management plans prescribed by teledermatologists were appropriate, and only one missed case of melanoma *in situ* was reported in one study. Unfortunately all studies had small sample sizes, requiring confirmation in larger and more diverse samples. Importantly, when teledermatology was used as a triage tool, all melanoma and melanoma *in situ* cases were correctly prioritised as high priority, while a number of melanoma and melanoma *in situ* were inappropriately triaged as medium or low priority using conventional (non-teledermatology) referral pathways.

Very few studies (4/21) assessed health services outcomes measures. Those that did found the use of teledermatology could reduce waiting times and result in earlier assessment and treatment, patients reported high satisfaction and were willing to pay out of pocket for access to such services. On the other hand, clinicians reported higher diagnostic difficulty for teledermatology cases compared to FTF consultations and levels of diagnostic agreement

were found to be related to clinician-reported diagnostic confidence. It is possible that with more experience and exposure to teledermatology cases, clinicians' increased confidence with this mode of health care delivery could result in improved accuracy. There were no studies assessing longer term outcomes like quality of life or workflow that were specific to skin cancer and eligible for inclusion in this review.

Authors of a recent commentary from Spain proposed teledermatology be used primarily to improve referral and triage systems, rather than replacing in-person consultations, due to the current lack of high level evidence to support diagnostic accuracy of teledermatology[38]. The evidence from this review also supports the implementation of teledermatology as a referral and/or triage tool. However, the potential benefits to patients who currently have limited access to dermatological care suggests there is reason to invest more resources to definitively establish the diagnostic accuracy of teledermoscopy for skin cancer diagnosis. Incorporating economic outcome measures into a larger diagnostic study would enable concurrent exploration of the most suitable models of care to integrate teledermatology into the diagnosis and management of melanoma and non-melanoma skin cancers.

Critical revision of the manuscript for important intellectual content: Janda, Soyer and

Dallest

Statistical analysis: N/A

Obtained funding: Nil, this study was unfunded.

Administrative, technical, or material support: Finnane and Dallest

Study supervision: Soyer and Janda

References

1.   Balch C, Soong S, Gershenwald J, et al. Prognostic factors analysis of 17,600 melanoma patients: validation of the American Joint Committee on Cancer melanoma staging system. *J Clin Oncol*. 2001;19(16):3622-3634.

2.   Ekeland AG, Bowes A, Flottorp S. Effectiveness of telemedicine: A systematic review of reviews. *Int J Med Inform*. 2010;79(11):736-771.

3.   Coates SJ, Kvedar J, Granstein RD. Teledermatology: From historical perspective to emerging techniques of the modern era: Part II: Emerging technologies in teledermatology, limitations and future directions. *J Am Acad Dermatol*. 2015;72(4):577-586.

4.   Shoo A, Sagebiel R, Kashani-Sabet M. Discordance in the histopathologic diagnosis of melanoma at a melanoma referral center. *J Am Acad Dermatol*. 2010;62(5):751-756.

5.   Mogensen M, Jemec G. Diagnosis of nonmelanoma skin cancer/keratinocyte carcinoma: A review of diagnostic accuracy of nonmelanoma skin cancer diagnostic tests and technologies. *Dermatologic Surg*. 2007;33:1158-1174.

6.   Heal CF, Weedon D, Raasch BA, Hill BT, Buettner PG. Agreement between histological diagnosis of skin lesions by histopathologists and a dermato-histopathologist. *Int J Dermatol*. 2009;48(12):1366-1369.

7.   Warshaw EM, Hillman YJ, Greer NL, et al. Teledermatology for diagnosis and management of skin conditions: A systematic review. *J Am Acad Dermatol*. 2011;64(4):759-772.

8.   Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009;6(7):e1000097.

9.   Whiting PF, Rutjes AWS, Westwood ME, et al. QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155(8):529-536.

10.   Griffiths W. Improving melanoma diagnosis in primary care: a teledermatoscopy project. *J Telemed Telecare*. 2010;16:185-186.

11.   Chao JT, Loescher LJ, Soyer HP, Curiel-Lewandrowski C. Barriers to mobile teledermoscopy in primary care. *J Am Acad Dermatol*. 2013;69(5):821-824.

12.   Warshaw EM, Gravely AA, Bohjanen KA, et al. Interobserver accuracy of store and forward teledermatology for skin neoplasms. *J Am Acad Dermatol*. 2010;62(3):513-516.

13.   Warshaw EM, Gravely AA, Nelson DB. Accuracy of teledermatology/teledermoscopy and clinic-based dermatology for specific categories of skin neoplasms. *J Am Acad Dermatol*. 2010;63(2):348-352.

14.   Janda M, Loescher L, Banan P, Horsham C, Soyer HP. Lesion selection by melanoma high-risk consumers during skin self-examination using mobile teledermoscopy. *JAMA dermatology*. 2014;150(6):656-658.

15. Janda M, Loescher L, Soyer HP. Enhanced skin self-examination: A novel approach to skin cancer monitoring and follow-up. *JAMA Dermatology*. 2014;149(2):231-236.

16. Piccoli MF, Amorim BDB, Wagner HM, Nunes DH. Teledermatology protocol for screening of skin cancer. *An Bras Dermatol*. 2015;90(2):202-210.

17. Wolf JA, Moreau JF, Akilov O, et al. Diagnostic inaccuracy of smartphone applications for melanoma detection. *JAMA dermatology*. 2013;149(4):422-426.

18. Kroemer S, Frühauf J, Campbell TM, et al. Mobile teledermatology for skin tumour screening: Diagnostic accuracy of clinical and dermoscopic image tele-evaluation using cellular phones. *Br J Dermatol*. 2011;164(5):973-979.

19. Borve A, Terstappen K, Sandberg C, Paoli J. Mobile teledermoscopy-there's an app for that! *Dermatol Pract Concept*. 2013;3(2):41-48.

20. van der Heijden JP, Thijssing L, Witkamp L, Spuls PI, de Keizer NF. Accuracy and reliability of teledermatoscopy with images taken by general practitioners during everyday practice. *J Telemed Telecare*. 2013;19(6):320-325.

21. Borve A, Dahlen Gyllencreutz J, Terstappen K, et al. Smartphone Teledermoscopy Referrals: A Novel Process for Improved Triage of Skin Cancer Patients. *Acta Derm Venereol*. 2015;95(2):186-190.

22. Massone C, Maak D, Hofmann-Wellenhof R, Soyer HP, Fruhauf J. Teledermatology for skin cancer prevention: an experience on 690 Austrian patients. *J Eur Acad Dermatol Venereol*. 2014;28(8):1103-1108.

23. Senel E, Sabancilar E, Mansuroglu C, Demir E. A preliminary study of the contribution of telemicroscopy to the diagnosis and management of skin tumours in teledermatology. *J Telemed Telecare*. 2014;20(4):178-183.

24. Congalton AT, Oakley AM, Rademaker M, Bramley D, Martin RCW. Successful melanoma triage by a virtual lesion clinic (teledermatoscopy). *J Eur Acad Dermatology Venereol*. 2015;29:2423-2428.

25. Tan E, Oakley A, Soyer HP, et al. Interobserver variability of teledermoscopy: an international study. *Br J Dermatol*. 2010;163(6):1276-1281.

26. Tan E, Yung A, Jameson M, Oakley A, Rademaker M. Successful triage of patients referred to a skin lesion clinic using teledermoscopy (IMAGE IT trial). *Br J Dermatol*. 2010;162(4):803-811.

27. Boyce Z, Gilmore S, Xu C, Soyer HP. The remote assessment of melanocytic skin lesions: A viable alternative to face-to-face consultation. *Dermatology*. 2011;223(3):244-250.

28. Lamel SA, Haldeman KM, Ely H, Kovarik CL, Pak H, Armstrong AW. Application of mobile teledermatology for skin cancer screening. *J Am Acad Dermatol*. 2012;67(4):576-581.

29. Borve A, Holst A, Gente-Lidholm A, Molina-Martinez R, Paoli J. Use of the mobile phone multimedia messaging service for teledermatology. *J Telemed Telecare*. 2012;18:292-296.

30. Warshaw EM, Gravely AA, Nelson DB. Reliability of store and forward teledermatology for skin neoplasms. *J Am Acad Dermatol*. 2015;72(3):426-435.

31. Ferrándiz L, Ruiz-de-Casas A, Martin-Gutierrez FJ, et al. Effect of teledermatology on the prognosis of patients with cutaneous melanoma. *Arch Dermatol*. 2012;148(9):1025-1028.

32. Morton CA, Downie F, Auld S, et al. Community photo-triage for skin cancer referrals: an aid to service delivery. *Clin Exp Dermatol*. 2010;36:248-254.

33. Spinks J, Janda M, Soyer HP, Whitty JA. Consumer preferences for teledermoscopy screening to detect melanoma early. *J Telemed Telecare*. 2015;22(1):39-46.

34. Wu X, Oliveria S a, Yagerman S, et al. Feasibility and efficacy of patient-initiated mobile teledermoscopy for short-term monitoring of clinically atypical nevi. *JAMA dermatology*. 2015;151(5):489-496.

35. Lim D, Oakley AMM, Rademaker M. Better, sooner, more convenient: a successful teledermoscopy service. *Australas J Dermatol*. 2012;53(1):22-25.

36. Silveira CEG, Silva TB, Fregnani JHGT, et al. Digital photography in skin cancer screening by mobile units in remote areas of Brazil. *BMC Dermatol*. 2014;14:19-24.

37. Manahan MN, Soyer HP, Loescher LJ, et al. A pilot trial of mobile, patient-performed teledermoscopy. *Br J Dermatol*. 2014;172:1072-1080.

38. Moreno-Ramírez D, Romero-Aguilera G. Teledermatology: From the tempest of debate to calmer waters. *Actas Dermo-Sifiliográficas (English Ed*. 2016;107(5):366-368.

39. International Society for Digital Imaging of the Skin. ISIC Melanoma Project: Implementation. http://isdis.net/isic-project/isic-melanoma-project-implementation/. Published 2016. Accessed July 7, 2016.

40. OCEBM Levels of Evidence Working Group. *The Oxford 2011 Levels of Evidence*.; 2011. http://www.cebm.net/ocebm-levels-of-evidence/.

41. Robinson J, Dellavalle R, Bigby M, Callen J. Systematic Reviews: Grading recommendations and evidence quality. *Arch Dermatol*. 2008;144(1):97-99.

Table 1. Definitions for main review outcomes, reference standards and index tests

Study Outcome	Reference Standard	Index Test
Diagnostic Accuracy	Histopathology results (excised lesions) AND FTF diagnosis (non-excised lesions)	Teledermatologist's diagnosis OR Face to Face Dermatologist's diagnosis
Diagnostic Concordance	Face to Face Dermatologist/s diagnosis	Teledermatologist's diagnosis
Management Accuracy	Management plan based on histopathology results (excised lesions) AND Management plan based on FTF results (non-excised lesions)	Teledermatologist's prescribed management plan OR Face to Face Dermatologist's prescribed management plan
Management Concordance	Face to Face Dermatologist/s prescribed management plan	Teledermatologist's prescribed management plan

Table 2. Study characteristics

Study details	Design	Population	Intervention	Outcome	Quality Rating*
Studies of Diagnostic Accuracy & Concordance and Management Accuracy and Concordance					
Borve 2012	Prospective diagnostic accuracy	Sweden 40 patients, 23 female/17 male (mean age = 49 yrs)	Teledermatology referrals via MMS service using mobile phone camera. Compared to FTF.	Diagnostic accuracy Management concordance Barriers – image quality	3
Borve 2013	Prospective diagnostic accuracy	Sweden 62 patients, 24 female/38 male, mean age = 64 yrs, 69 included lesions.	Clinical info and dermoscopic images uploaded to iDoc24 app and compared to FTF	Diagnostic accuracy Management accuracy Barriers – image quality and diagnostic difficulty	2
Borve 2015	Open, controlled, multicentre prospective observational study	Sweden 20 Primary healthcare centres, an urban hospital and a rural hospital; 346 referrals	iDoc24 teledermoscopy referral compared to paper based referral system	Diagnostic accuracy Management accuracy Time	2
Boyce 2011	Prospective diagnostic accuracy	Australia 55 patients, 22 female/33 male (median age =26 yrs), 157 lesions.	Assessment of mobile phone images compared with FTF assessment.	Management concordance Barriers – ease of use	3
Congalton 2015	Prospective diagnostic accuracy	New Zealand 310 eligible patients, 54% female, (median age=58 yrs), 613 lesions	Macro and dermoscopic images captured at molemap clinic and sent with clinical info.	Diagnostic accuracy Time	3
Kroemer 2011	Prospective diagnostic accuracy	Austria 88 patients (41 men, 47 women, median age = 69 years), 113 lesions	Mobile phone camera and dermatoscope used to take images. Uploaded and compared to FTF diagnosis	Diagnostic accuracy Barriers - Image quality	3
Lamel 2012	Prospective diagnostic concordance	United States 86 patients 58% female, mean age = 45.24 yrs. 137 lesions	Mobile phone images compared with FTF diagnosis at screening event	Diagnostic concordance Management Concordance	2
Manahan 2015	Pilot Randomized Controlled Trial	Australia 49 participants, 50-64 years, 49% male 49% with history of at least one skin cancer.	Mobile teledermoscopy used by consumers to submit images for review. Randomised to receive instructions for self-examination or not.	Management concordance Barriers – instructions for self-examination	2

Table 2 continued.

Study details	Design	Population	Intervention	Outcome	Quality Rating*
Massone 2014	Observational	Austria 690 patients, 48 female, 642 male (mean age = 47 yrs) 962 lesions	GPs performing FTF skin checks at a screening programme took macro and dermoscopic images for Teledermatology assessment	Diagnostic accuracy Barriers – image quality	3
Senel 2014	Retrospective diagnostic accuracy	Turkey 120 consecutive cases, 57% male (mean age = 63 yrs)	Medical records and images from patients archives used to compare retrospectively FTF diagnosis with telediagnosis	Diagnostic accuracy Management concordance	2
Silveira 2014	Prospective diagnostic accuracy	Brazil 460 suspicious lesions from 2592 dermatological examinations	Images taken at Mobile Prevention Unit using digital camera, prior to biopsy, and sent to two oncologists for telediagnosis	Diagnostic concordance Barriers – image quality	2
Tan 2010a	Prospective diagnostic accuracy	New Zealand 200 patients, 126 female/74 male (age =11-94yrs) with 491 lesions.	Panoramic, macroscopic and dermoscopic images uploaded to MoleMap for telediagnosis, compared to FTF diagnosis	Diagnostic concordance Barriers – access	2
Tan 2010b	Prospective diagnostic accuracy	New Zealand 979 lesions from 206 patients	Panoramic, macroscopic and dermoscopic images uploaded to MoleMap for telediagnosis by 5 International dermatologists online.	Interobserver variability	3
Van der Heijden 2013	Prospective diagnostic accuracy	Netherlands 105 patients, 55% female (mean age = 47 yrs) 108 lesions	GPs took macro and dermoscopic images and sent for teledermatology diagnosis, compared with FTF diagnosis by dermatologist.	Diagnostic accuracy Management concordance Barriers – image quality	3
Warshaw 2015	Cross-sectional, repeated measures study	United States 2152 patients, 96.8% male (mean age=68 yrs), 3021 lesions	Macro and dermoscopic images collected and sent to teledermatologist. Diagnosis compared to FTF diagnosis.	Diagnostic concordance Management concordance Barriers – image quality	2
Wolf 2013	Case-control diagnostic accuracy study	United States 188 images of lesions from clinical database	Images from database uploaded to mobile application for Teledermatology diagnosis and compared with histopathological diagnosis.	Diagnostic Accuracy Barriers – image quality	3
Studies of enablers and barriers to implementation					
Ferrandiz 2012	Descriptive longitudinal study	Spain 201 patients (52.4% women, mean age 57.5 years)	Teledermatology system used to triage patients with suspicious lesions. Outcomes compared to conventional referral system.	Time - measured as prognosis	3
Lim 2012	Prospective observational study of patient flow	New Zealand 100 FTF patients (36% male, mean age=62.7yrs), 200 Virtual lesion clinic patients (39% male, mean age-53yrs)	Dermatologist triaged referrals to be seen FTF or through Virtual Lesion Clinic. Patient flow was compared.	Time Barriers	3

26

Study details	Design	Population	Intervention	Outcome	Quality Rating*
Morton 2010	Prospective observational study of patient flow	Scotland 289 patients photo-triaged (171 female, 118 male, mean age=51 yrs), 188 conventional pathway (93 female, 95 male, mean age = 52 yrs)	Patients referred by GP by conventional electronic letter or through a dedicated skin cancer standard referral form with photographs taken at GP practice. Patient flow compared between groups.	Time	3
Spinks 2015	Discrete choice experiment	Australia Participants from a teledermoscopy trial, aged 50-64 35 participants	Participants who had used consumer-driven teledermoscopy completed a survey to investigate preferences for melanoma screening options.	Barriers - Patient choices and willingness to pay	3
Wu 2015	Prospective cohort study	United States 34 patients, 18 men, 16 women, mean age = 43.6 yrs (18-81yrs). 29 patients with 33 lesions completed follow up	Patients took images of lesions at baseline and follow-up and completed surveys assessing skincare awareness and attitudes towards teledermoscopy.	Barriers – difficulty of image acquisition, patient receptivity, confidence	3

*Quality rating scheme is modified from the Oxford Centre for Evidence-Based Medicine ratings of individual studies: (1) Systematic review of cross sectional studies with consistently applied reference standard and blinding (2) Individual cross sectional studies with consistently applied reference standard and blinding (3) Non-consecutive studies, or studies without consistently applied reference standards** (4) Case-control studies, or "poor or non-independent reference standard** (5) Mechanism-based reasoning. *Level may be graded down on the basis of study quality, imprecision, indirectness (study PICO does not match questions PICO), because of inconsistency between studies, or because the absolute effect size is very small; Level may be graded up if there is a large or very large effect size.*[40]

Table 3. Summary table of results for the diagnostic accuracy, diagnostic concordance, management accuracy and management concordance of teledermatology for the diagnosis of skin cancer

Reference standard	Index test	Intervention	Results	References
Diagnostic Accuracy				
Histopathology	Teledermatology	No Dermoscopy	Diagnostic agreement 68% Sensitivity 98.1% (95%CI=88.8-99.9) Specificity 30% (95%CI=22.1-40.3)	Senel 2014 Wolf 2013
		+ Dermoscopy	Diagnostic agreement 51-92% Sensitivity=96%, Specificity=62%	Borve 2013, Borve 2014, Senel 2014 Congalton 2015
	Face to Face	No Dermoscopy	No studies	
		+ Dermoscopy	Diagnostic agreement 67%	Borve 2013
Histopathology (excised lesions) *and* FTF diagnosis (non-excised lesions)	Teledermatology	No Dermoscopy	Diagnostic agreement 85%	Kroemer 2011
		+ Dermoscopy	Diagnostic agreement 79-94% k=0.41-0.63	Kroemer 2011, Massone 2014 van der Heijden 2013
	Face to Face	No Dermoscopy	No studies	
		+ Dermoscopy	Diagnostic agreement 85% k=0.90	Kroemer 2011 Van der Heijden 2013
Diagnostic Concordance				
Face to face	Teledermatology	No Dermoscopy	Diagnostic concordance 62-94%	Lamel 2012, Borve 2012, Silveira 2014
		+ Dermoscopy	Diagnostic concordance 46-90%	Tan 2010a, Borve 2013, Manahan 2015, Warshaw 2015
Management Accuracy				
Histopathology	Teledermatology	No Dermoscopy	Not reported	
		+ Dermoscopy	1 melanoma *in situ* would have received no treatment (according to 1 TD) 100% of invasive MM (n=19) prioritised correctly, 100% of MM (n=16) prioritised correctly (at least medium)	Borve 2013 Borve 2015

Table 3 continued.

Reference standard	Index test	Intervention	Results	References
Histopathology	Face to Face	No Dermoscopy	Not reported	
		+ Dermoscopy	3/4 invasive melanoma and 3/5 melanoma *in situ* incorrectly given medium or low priority triage	Borve 2015
Management Concordance				
Face to face	Teledermatology	No Dermoscopy	Management concordance 69-95% k=0.23-0.57	Boyce 2011, Lamel 2012, Borve 2012
		+ Dermoscopy	Management concordance 66-85% k=0.19-0.83	Van der Heijden 2013, Warshaw 2015, Sene 2014

Table 4. QUADAS-2 Summary Table

Study	RISK OF BIAS				APPLICABILITY CONCERNS		
	PATIENT SELECTION	INDEX TEST	REFERENCE STANDARD	FLOW AND TIMING	PATIENT SELECTION	INDEX TEST	REFERENCE STANDARD
Borve et al 2012	Low	Low	High	Unclear	Low	Low	Low
Borve et al 2013	High	Unclear	Low	Low	Low	Low	Low
Borve et al 2015	Low	Low	Unclear	Low	Low	Low	Low
Congalton et al 2015	Low	Low	Unclear	High	Low	Low	Low
Kroemer el al 2011	Unclear	Unclear	Low	High	Low	Low	Low
Lamel et al 2012	High	Low	Low	Low	Low	Low	Low
Manahan et al 2015	High	Low	Low	Low	Low	Low	Low
Massone et al 2014	High	Low	High	High	High	Low	Low
Senel et al 2015	Low	Low	Low	Low	Unclear	Low	Low
Silveira et al 2014	Unclear	Low	Low	Unclear	Low	High	Low
Tan et al 2010a	Unclear	High	Low	Low	Unclear	Low	Low
Van der Heijden 2013	Unclear	Unclear	Unclear	Unclear	Unclear	Low	Low
Warshaw et al 2015	Unclear	Low	Low	Low	High	Low	Low
Wolf et al 2013*[i]	High	High	Low	Low	High	High	Low

Table 5. Summary of recommendations for teledermatology for the diagnosis and management of skin cancers

Recommendation	Grade of Recommendation*	Quality of Evidence*	Source
Teledermatology should be used for patients where it is not feasible to provide FTF consultation	2A	B	Borve 2012; Borve 2013; Borve 2015; Boyce 2011; Congalton 2015; Kroemer 2011; Lamel 2012; Manahan 2015; Massone 2014; Senel 2014; Silveira 2014; Tan 2010a; Tan 2010b; Van der heijden 2013; Warshaw 2015; Wolf 2013
Teledermatology can be used as a triage tool to reduce waiting times to assessment.	2A	B	Ferrandiz 2012; Lim 2012; Morton 2010; Borve 2015
Currently available technology is suitable for Teledermatology assessment. Training of clinicians and consumers/patients should be considered to improve image quality.	1	B	Boyce 2011; Wolf 2013; Silveira 2014; Borve 2013; Van der heijden 2013; Massone 2014; Manahan 2015; Warshaw 2015; Wu 2015

Abbreviations: FTF, face to face.
*Graded according to criteria by Robinson et al. [41] Grade of recommendation: 1, strong recommendation; high-quality, patient-oriented evidence; 2A, weak recommendation; limited-quality, patient-oriented evidence; 2B, weak recommendation; low-quality evidence. Quality of evidence: A, systematic review/meta-analysis; randomized clinical trials with consistent findings; all-or-none observational studies; B, systematic review/meta-analysis of lower-quality clinical trials or studies with limitations and inconsistent findings; lower-quality clinical trial; cohort study; case-control study; C, consensus guidelines, usual practice, expert opinion, case series.