Department of Veterans Affairs
Health Services Research & Development Service | Evidence-based Synthesis Program

HSR&D

Teledermatology for Diagnosis and Management of Skin Conditions: A Systematic Review of the Evidence

EXECUTIVE SUMMARY

January 2010

Prepared for:

Department of Veterans Affairs
Veterans Health Administration
Health Services Research &
Development Service
Washington, DC 20420

Prepared by:

VA Evidence Synthesis Program
Center for Chronic Disease
Outcomes Research
Minneapolis VA Medical Center
Minneapolis, MN

Principal Investigator:

Erin Warshaw, MD, MS

Research Associates

Nancy Greer, PhD
Yonatan Hillman, BA
Emily Hagel, MS
Roderick MacDonald, MS
Indulis Rutks, BS

ESP Program Director

Timothy J. Wilt, MD, MPH

PREFACE

VA's Health Services Research and Development (HSR&D) Service works to improve the cost, quality, and outcomes of healthcare for our nation's veterans. Collaborating with VA leaders, managers, and policy makers, HSR&D focuses on important healthcare topics that are likely to have significant impact on quality improvement efforts. One significant collaborative effort is HSR&D's Evidence-based Synthesis Program (ESP). Through this program, HSR&D provides timely and accurate evidence syntheses on targeted healthcare topics. These products will be disseminated broadly throughout VA and will: inform VA clinical policy, develop clinical practice guidelines, set directions for future research to address gaps in knowledge, identify the evidence to support VA performance measures, and rationalize drug formulary decisions.

HSR&D provides funding for four ESP Centers. Each Center has an active and publicly acknowledged VA affiliation and also serves as an Evidence Based Practice Center (EPC) supported by the Agency for Healthcare Research and Quality (AHRQ). The Centers will each generate three evidence syntheses annually on clinical practice topics of key importance to VHA leadership and policymakers. A planning committee with representation from HSR&D, Patient Care Services (PCS), Quality Enhancement Research Initiative (QUERI), Office of Quality and Performance (OQP), and the VISN Clinical and Quality Management Officers, has been established to identify priority topics and key stakeholder concerns and to ensure the quality of final reports. Comments on this evidence report are welcome and can be sent to Nicole Floyd, ESP Coordinating Center Program Manager, at nicole.floyd@va.gov.

Recommended citation: Warshaw E, Greer N, Hillman Y, Hagel E, MacDonald R, Rutks I and Wilt TJ. Teledermatology for Diagnosis and Management of Skin Conditions: A Systematic Review of the Evidence. VA-ESP Project #09-009; 2009

This information is distributed solely for the purposes of pre-dissemination peer review. It has not been formally disseminated by the Department of Veterans Affairs. It does not represent and should not be construed to represent a Department of Veterans Affairs determination or policy.

Financial disclosure: No investigators have any affiliations or financial involvement (e.g., employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties) that conflict with material presented in the report.

EXECUTIVE SUMMARY

BACKGROUND

Telemedicine uses telecommunication technology to transfer medical information. Due to the visual nature of a skin examination, telemedicine, specifically, teledermatology, may be a valuable tool in the diagnosis and management of dermatologic diseases for patients in rural areas (including rural Veterans Affairs Medical Centers and Community Based Outpatient Clinics) who may not have ready access to a dermatologist. Teledermatology may also be useful in primary care settings to triage cases and limit unnecessary dermatology clinic referrals. Although not the focus of this review, teledermatology may also be used to provide follow-up care or monitoring after an in-person dermatology visit. The objectives of this evidence synthesis project were to systematically review and summarize the scientific literature addressing: 1) teledermatology for the diagnosis of skin conditions, 2) teledermatology for the management of skin conditions, 3) clinical outcomes when teledermatology is used, 4) the cost of teledermatology compared with usual care (in-person dermatology), and 5) key elements of, and barriers to, successful teledermatology implementation. Specifically, the key questions were:

KEY QUESTION #1

1a. How does the *accuracy* of teledermatology compare to usual care (in-person dermatology) for the *diagnosis* of skin conditions?

1b. How does the *concordance* of teledermatology compare to usual care (in-person dermatology) for the *diagnosis* of skin conditions?

KEY QUESTION #2

2a. How does the *accuracy* of teledermatology compare to usual care (in-person dermatology) for clinical *management* of skin conditions?

2b. How does the *concordance* of teledermatology compare to usual care (in-person dermatology) for clinical *management* of skin conditions?

KEY QUESTION #3

3. How do clinical outcomes (clinical course, satisfaction, quality of life, visits avoided) of teledermatology compare to usual care (in-person dermatology) for skin conditions?

KEY QUESTION #4

4. How does the cost of teledermatology compare to usual care (in-person dermatology)?

KEY QUESTION #5

5. What are the key structural and process elements associated with successful implementation of teledermatology and what are the barriers?

METHODS

We searched MEDLINE (OVID) and PubMed for controlled clinical trials, systematic reviews, cost studies, and implementation papers from 1990 to June, 2009 using standard search terms (Appendix A). We limited the search to peer-reviewed articles involving human subjects and published in English language. For key questions 1 and 2, inclusion was limited to controlled trials. Search terms included: remote consult/consultation, electronic mail, telecommunications, telemedicine, telepathology, dermatology, and teledermatology.

Titles and abstracts identified from the search were reviewed by physicians and research associates trained in the critical analysis of literature to identify peer-reviewed articles related to one or more of the key questions. We included studies of store and forward (SAF) and live interactive (LI) technologies.

Study characteristics, patient characteristics, and outcomes were extracted by a trained research associate under the supervision of the Principal Investigator, a Veterans Affairs (VA) dermatologist (Appendix B). We assessed study quality according to Quality Assessment of Diagnostic Accuracy Studies (QUADAS) criteria. We identified additional citations from reference lists of related articles. We performed pooled analyses where feasible and clinically appropriate. All other data were narratively summarized.

DATA SYNTHESIS

We constructed evidence tables showing study, patient, and intervention characteristics; methodological quality; and outcomes, organized by key question and teledermatology technology. We analyzed studies to compare their characteristics, methods, and findings. We compiled a summary of findings for each key question based on qualitative and semi-quantitative synthesis of the findings. Variability in patient, intervention, and outcome reporting limited pooling of findings across all studies, however, and when appropriate (similar technology and/or skin conditions), weighted pooled averages were calculated. We identified and highlighted findings from VA or Department of Defense (DoD) populations.

PEER REVIEW

A draft version of this report was sent to two peer reviewers in addition to our Technical Expert Panel. Reviewer comments were addressed and our responses incorporated in the final report (Appendix C).

RESULTS

LITERATURE FLOW

The literature search yielded 658 citations, of which 185 articles met initial criteria for full-text level review. From these, we identified 85 references (including 69 controlled clinical trials) that addressed at least one of the key questions and met inclusion criteria. Among studies included to assess diagnostic accuracy and concordance (KQ1 and 2), most utilized methods to reduce sources of bias; particularly related to appropriate use of index and reference tests. However, the majority of studies using store and forward technology did not clearly address patient selection

biases such as enrolling a representative spectrum of general dermatological patients or clearly describing exclusion criteria. Both store and forward and live interactive studies generally did not account for all patients at the end of the study or include patients with uninterpretable results.

OVERVIEW OF STUDIES FOR QUESTIONS 1 AND 2 (TABLE 2)

SAF: Forty-one SAF teledermatology studies met inclusion requirements and enrolled between 12 and 882 individuals. Of studies reporting specific subject characteristics, the mean age of enrollees was 53 years, 43% were female, and 93% were Caucasian. Twelve studies were conducted in the United States (5 involving veterans or military personnel). Eleven studies enrolled subjects specifically with pigmented lesions and one study enrolled subjects with non-pigmented lesions. Nineteen studies reported diagnostic accuracy (defined as a comparison of the diagnosis against histopathology/other laboratory test) and 27 studies reported diagnostic concordance (a comparison of the diagnoses made with teledermatology and clinical dermatology without verification by histopathology or laboratory test). Two studies reported management accuracy (a comparison of the teledermatology management plan against a management plan based on histopathology/other laboratory test) and 14 studies reported management concordance (a comparison of the management plans based on teledermatology and clinical dermatology).

LI: Ten LI teledermatology studies met inclusion requirements and enrolled between 51 and 351 individuals. The mean age of enrollees was 40 years, 54% were female, and 72% were Caucasian. Five of the studies were conducted in the United States and one study evaluated U.S. military personnel or veterans. One study enrolled subjects with isolated skin lesions; the remainder included subjects with both rashes and lesions. One LI study reported diagnostic accuracy, 10 reported diagnostic concordance, none reported management accuracy, and 4 reported management concordance.

Key Question #1a How does the accuracy of teledermatology compare to usual care (in-person dermatology) for the *diagnosis* of skin conditions? *(Table 3)*

Conclusion: Two-thirds of studies comparing teledermatology and usual care found better diagnostic accuracy with usual care (in-person dermatology visit) as compared to teledermatology. Estimates from subsamples of studies providing sufficient evidence for pooling suggested the magnitude of difference between the accuracy rates was approximately 11% and 19% for primary and aggregated diagnostic accuracy, respectively, and 5% for pigmented lesions. When dermatoscopy-trained teledermatologists were available, teledermatoscopy increased diagnostic accuracy of isolated skin lesions, although overall, rates were still not superior to usual care. One study found that the diagnostic accuracy of teledermatology was significantly worse for eleven common skin neoplasms including melanoma, squamous cell carcinoma, and basal cell carcinoma.

Summary of Studies: Twenty studies (19 SAF and 1 LI) reported diagnostic accuracy defined as matching of teledermatology diagnosis with histopathology diagnosis or other laboratory test. Results were reported as one or more of the following: 1) accuracy rate (percent match between the primary diagnosis and/or aggregated diagnoses [primary plus differential]

and histopathology/laboratory test), 2) kappa statistic, and/or 3) sensitivity and specificity. Fifteen studies also assessed diagnostic accuracy of usual care (in-person dermatology diagnoses), allowing for direct comparisons between these two methods of care. Ten of these 15 studies reported better diagnostic accuracy for usual care (in-person dermatology visit) than teledermatology, 3 studies reported better accuracy for teledermatology, and 2 studies reported mixed results. Statistical pooling of the 6 SAF studies reporting aggregated diagnostic accuracy rates found that the weighted mean absolute difference between accuracy rates was 19% better for usual care than teledermatology. For the 11 SAF studies that reported primary diagnostic accuracy rates, the weighted mean absolute difference between accuracy rates was 11% better for usual care than teledermatology. Similarly, the weighted mean difference for primary diagnostic accuracy rates for six pigmented skin lesion studies was also better (5%) for usual care than teledermatology. Four studies evaluated teledermatology with standard macro images and teledermatoscopy. In general, teledermatology accuracy rates improved with teledermatoscopy (up to 15% absolute difference), although, overall, accuracy of teledermatoscopy was still not superior to usual care.

Key Question #1b How does the concordance of teledermatology compare to usual care (in-person dermatology) for the *diagnosis* of skin conditions? *(Table 4)*

Conclusion: Analysis from a limited subsample of studies providing sufficient evidence for pooling suggested that the aggregated diagnostic concordance rates for SAF teledermatology were similar for lesion studies (64%) and general studies (65%); the rate for LI (87%) was higher, but based on fewer patients. The weighted mean primary diagnostic concordance for SAF teledermatology was also similar for lesion studies (62%) and general studies (66%); the rate for LI studies was higher (71%) but based on fewer patients. In summary, diagnostic concordance of SAF was good and may be better for LI, possibly due to the ability to obtain additional history in the LI setting.

Summary of Studies: Thirty-seven (27 SAF, 9 LI, 1 SAF+LI) studies reported diagnostic concordance (simple agreement without verification by histopathology or laboratory test) between usual care (in-person dermatology visit) and teledermatology. Thirty-five studies (25 SAF, 9 LI, 1 SAF+LI) reported concordance as percent agreement for diagnosis, malignant/benign status, or diagnostic category. Seven studies reported kappa statistics and three studies reported sensitivity and specificity.

Percent Concordance - SAF Studies: Weighted average diagnostic concordance rates for studies involving subjects with isolated skin lesions were 64% (aggregated, number of studies=4) and 62.3% (primary, n=6). Nineteen studies involving a range of dermatologic conditions (lesions and rashes) evaluated diagnostic concordance; rates ranged from 60-100% for aggregated diagnostic concordance (n=10) and 46-88% for primary diagnostic concordance (n=14). Excluding studies in which the same dermatologist served as both clinic dermatologist and teledermatologist, weighted average diagnostic concordance rates were 65% (aggregated, n=7) and 66% (primary, n=11).

Percent Concordance - LI Studies: Diagnostic concordance rates of LI ranged from 78-99% (aggregated, n=6) and 57-78% (primary, n=8). Excluding studies in which the same dermatologist served as both clinic dermatologist and teledermatologist, weighted average

diagnostic concordance rates were 87% (aggregated, n=3) and 71% (primary, n=5).

Kappa, Sensitivity/Specificity: Kappa values ranged from 0.71-0.93 for the three SAF teledermatology studies reporting this statistic. Excluding the one study with likely bias (same dermatologist served as both clinic dermatologist and teledermatologist), kappa values indicated substantial agreement. Three LI studies reported kappa values that ranged from 0.32-0.79. Sensitivity and specificity was reported in three studies (utilizing the clinic dermatologist's assessment as the gold standard and agreement for benign or malignant status, not exact diagnosis). Sensitivity ranged from 0.88-1.0 and specificity ranged from 0.39-0.98.

Key Question #2a How does the accuracy of teledermatology compare to usual care (in-person dermatology) for clinical *management* of skin conditions? *(Table 5)*

Conclusion: While overall rates of management accuracy were equivalent (±10% absolute difference) for teledermatology and usual care, for malignant and premalignant lesions, rates for teledermatology and teledermatoscopy were inferior to usual care; caution is recommended when using teledermatology in these cases.

Summary of Studies: Only two studies, both by the same authors and using store and forward technology, evaluated management accuracy, defined as the percent agreement with an expert panel management plan based on histopathology. One study evaluated pigmented lesions and one evaluated non-pigmented lesions. The range for reported management accuracy was 70-80% for teledermatology compared to 66-84% for usual care. Clinical dermatology management accuracy rates were worse for pigmented lesions (66%) than non-pigmented lesions (84%).

Key Question #2b How does the concordance of teledermatology compare to usual care (in-person dermatology) for clinical *management* of skin conditions? *(Table 6)*

Conclusion: Concordance rates for management were moderate to excellent for both SAF and LI teledermatology (55-100%). The range for kappa statistic values was 0.47-0.71 (3 SAF studies and 1 LI) indicating fair to good agreement.

Summary of Studies:

SAF: Fifteen SAF teledermatology studies reported management concordance (percent agreement n=13, kappa n=3, sensitivity and specificity n=2). Two studies evaluated concordance of the triage management decision of "refer or not refer" for isolated skin lesions, yielding a weighted average rate of 75%. Two studies evaluated concordance rates for three different management options; these rates were 72% and 96%. Three studies evaluated concordance for the diagnostic procedure decision "biopsy or no biopsy" and found concordance rates of 76-100%. Several studies did not describe management options but reported percent concordance rates from 55-94%.

LI: Four LI teledermatology studies reported management concordance. The concordance rate for the decision "biopsy vs. no biopsy" for skin lesions was 86%. Three other studies involving a

wide variety of skin conditions found concordance rates of 64%, 72%, and 75%.

Key Question #3 How do clinical outcomes (clinical course, satisfaction, quality of life, visits avoided) of teledermatology compare to usual care (in-person dermatology) for skin conditions? *(Tables 7, 8, and 9)*

Conclusion: There was insufficient evidence to conclude whether teledermatology had an effect on clinical course, although a large VA/DoD study reported comparable outcomes. Patient overall satisfaction with and preference for teledermatology or usual care were comparable in VA/DoD and other studies. Time to treatment was shorter and in-person visits can be avoided when patients are seen by teledermatology.

Summary of Studies: We identified 29 studies that reported clinical course, satisfaction, and/or visits avoided (17 SAF, 11 LI, and 1 SAF+LI). No studies reported quality of life. Among the SAF studies, two reported clinical course, nine reported patient satisfaction, and nine reported visits avoided. Three studies reported an additional outcome - time to treatment. Among the LI studies, one reported clinical course, nine reported patient satisfaction, and three reported visits avoided. The study that included both SAF and LI technologies reported only patient satisfaction.

Although two of three studies reporting clinical course suggested a more favorable outcome following teledermatology, these three studies used different methods for determining clinical course and assessed clinical course at different time points. The largest study, a VA and DoD study with over 500 patients, found no difference in the percentage of patient considered "improved" at 4 months after initial evaluation. Patients expressed comparable levels of satisfaction with teledermatology and usual care in three randomized, controlled trials (including one VA-based study). One non-randomized study reported greater satisfaction with teledermatology and one repeated measures study reported greater satisfaction with usual care. Response rates for the satisfaction assessments ranged from 58-100%. With the exception of one study which reported that 76% of subjects preferred teledermatology over waiting for a dermatology clinic appointment, preferences for teledermatology or usual care were similar. In one VA study, 42% preferred teledermatology over usual care while 37% preferred usual care over teledermatology. In five SAF studies that reported time to in-person consult or treatment, the time was shorter for patients who were initially seen by teledermatology. Teledermatology also reduced waiting times for clinic appointments and reduced the need for a in-person appointment by 14-66%.

Key Question #4 How does the cost of teledermatology compare to usual care (in-person dermatology)? *(Table 10)*

Conclusion: Cost analyses were limited by broad variations in cost assessment parameters and perspectives. Most studies found teledermatology to be cost effective if certain critical assumptions were met particularly patient travel distance, teledermatology volume, and the costs of usual care.

Summary of Studies: Three studies reported cost outcomes comparing SAF teledermatology to usual care. Six studies compared LI teledermatology to usual care. One study reported data from patients evaluated with both SAF and LI teledermatology. Wide differences existed in cost assessment parameters and perspectives evaluated (societal, health service, or patient). The majority of studies of SAF and LI found teledermatology to be cost effective if certain assumptions regarding patient travel distance, volume of teledermatology, and costs of usual dermatology care were met.

A micro-costing approach using a VA perspective found SAF teledermatology to be cost-effective, but not cost-saving, for decreasing time to initial definitive dermatologic care assuming that VA centers had both on-site primary care and dermatology clinics. The long-duration to achieve definitive dermatologic care, particularly for the usual care population (137.5 days vs. 50 days for the teledermatology group), however, is not consistent with current VA practice (all appointments within 30 days) and may result in an overly favorable estimate of teledermatology. A DoD study reported cost savings of $32 per patient if lost productivity was considered.

Key Question #5 What are the key structural and process elements associated with successful implementation of teledermatology and what are the barriers? *(Table 11)*

Conclusion: Key elements include: defining the setting for implementation, defining the objectives of the program, determining the organizational structure, identifying the resources available, considering all costs associated with teledermatology, determining the business model, procuring organizational support, and determining the training needs.

Summary of Studies: We attempted to categorize success facilitators using previously established definitions. We categorized implementation barriers according to administrative, clinical, patient, and technical factors. We emphasized factors likely to play a role in VA specific settings. We identified 12 descriptive studies that provided information relevant to implementation of a teledermatology program. Key elements included efficient and user-friendly programs, as well as ongoing technical and personnel support.

FUTURE RESEARCH RECOMMENDATIONS

Additional research is needed to determine the long-term effectiveness, feasibility, satisfaction, and cost-effectiveness of teledermatology, especially store and forward methodology. Standardized reporting of diagnostic, management, and outcome accuracy and concordance are important. Research evaluating clinical outcomes and patient management are especially needed. Studies that blind the assessor(s) to the patient/lesion/care method are preferred to reduce bias in outcome assessment. Additional outcomes could assess the impact of teledermatology on primary care practitioners' practice, satisfaction, and follow-up patterns. Barriers to successful implementation need to be identified that incorporate differences in patient populations, skin condition severity, distance traveled, availability of on-site dermatologists, and other clinical setting issues in order to determine the relative feasibility and effectiveness of different teledermatology strategies. Research priorities include comparing teledermatology

with dermatologic care by a VA primary care provider or a dermatology trained nurse practitioner (rather than a dermatologist), assessing patient and primary care provider (as well as dermatologist) satisfaction with teledermatology, and conducting high quality cost effectiveness studies relevant to VA populations and care settings.

CONCLUSIONS

While the concordance of teledermatology and in-person dermatology care for diagnosis and management of skin conditions was generally acceptable, data from studies assessing accuracy indicate that accuracy of teledermatology is inferior to in-person dermatology care, especially for skin malignancies, an important and common condition in the veteran population. Little information exists on the impact of teledermatology on clinical outcomes. Patient and provider satisfaction with teledermatology were relatively high though there were individuals who have strong beliefs for a particular approach. Cost analysis studies were limited in number and relevance to current United States practice. Studies are needed to compare teledermatology with primary care to better understand the most effective way to deliver dermatology care in areas without reliable access to in-person dermatology (e.g., rural areas). . Given the results of this review, the potential benefits of teledermatology (e.g., decreased patient travel, shorter time to intervention, primary care provider education) need to be evaluated in the context of its limitations including inferior diagnostic accuracy and management accuracy, especially for malignant skin neoplasms.