Teledermatology: An Intraobserver Diagnostic Correlation Study, Part II

MAJ Hon S. Pak, MC, USA; MAJ David Harden, MC, USA; David Cruess, PhD; LTC Mark L. Welch, MC, USA; LTC(P) Ronald Poropatich, MC, USA; and the National Capital Area Teledermatology Consortium

*This is part II of an intraobserver diagnostic correlation study comparing teledermatology with traditional face-to-face evaluation. Part I discussed the methodology and diagnostic correlation results between teledermatology and in-person consultation (*Cutis®*. 2003;71:399-403). This second part reports the diagnostic certainty level between the 2 groups, which are shown to be significantly different (teledermatology, 7/10; in-person, 9/10). This difference held true in every category of skin condition evaluated (P≤.0065). Unlike other studies, we found that teledermatologists recommended biopsies 10% more frequently than clinic-based evaluators. We discuss the reasons for the lower diagnostic certainty level of teledermatologists, as well as the limitations of this study. Despite the limitations, we conclude that teledermatology appears to be an effective method of delivering dermatologic care in the appropriate setting.*

Cutis. 2003;71:476-480.

Background

Recent studies on store and forward (S&F) teledermatology have shown diagnostic agreement similar to that of video teleconference teledermatology, ranging from 54% to 92% depending on how diagnostic agreement is defined.[1] Moreover, most studies evaluated and compared biopsy rates between teledermatologists and in-person dermatologists, reporting no statistical differences. Furthermore, ample data in the literature suggest there is a direct correlation between diagnostic certainty level (diagnostic confidence) and the level of diagnostic agreement (complete, partial, disagreement). In this second part of the article, we report the results of our research on the biopsy rates, comparative diagnostic certainty level, and limitations of this paper.

Results

For a complete discussion of the methodology, see Part I of this article (Cutis. 2003;71:399-403).

Diagnostic Certainty Level—Overall, the diagnostic certainty level was significantly lower with teledermatology (7/10) than with in-person evaluation (9/10)($P=.0001$). As expected, the diagnostic certainty level was significantly higher with the complete agreement group compared with the noncomplete agreement group (partial agreement plus disagreement)(Table 1). Interestingly, we found that eczematous and papulosquamous categories had the lowest teledermatology diagnostic certainty level. The difference in diagnostic certainty level between teledermatology and in-person consultation was statistically significant in every category of skin condition evaluated ($P\leq.0001$).

Biopsy Correlation—When recommendations for biopsy from the teledermatology group were compared with the number of patients who actually underwent a biopsy, the teledermatology group had a 10% higher overall biopsy recommendation rate (160 [39.6%] vs 119 [29.5%])(Table 2). This 10% actually represents 2 subgroups: (1) unnecessary biopsies recommended by teledermatologists (false-positive rate of 17% [69]), and (2) necessary biopsies not recommended by teledermatologists (false-negative rate of 7% [28]).

When the results of the biopsies (histologic diagnoses) were compared with the diagnoses of the teledermatologists and in-person consultants, the in-person consultants were correct more often (complete agreement 73% vs 18.9%). However, the difference in the overall rate of agreement

Accepted for publication February 6, 2003.

Dr. Pak is from Brooke Army Medical Center, San Antonio, Texas. Drs. Harden, Welch, and Poropatich are from Walter Reed Army Medical Center, Washington, DC. Dr. Cruess is from Uniformed Services University of Health Sciences, Bethesda, Maryland.

The authors report no conflict of interest.

The opinions expressed in this article are those of the authors and should not be construed to reflect those of the US Army or the US Department of Defense.

Reprints: MAJ Hon S. Pak, MC, USA, Department of Dermatology, Brooke Army Medical Center, 3851 Roger Brooke Dr, Bldg 3600, San Antonio, TX 78258 (e-mail: hon.pak@amedd.army.mil).

Table 1.

Diagnostic Certainty Level by Category of Skin Diseases (N=404)*

Category of Disease	No. of Patients	Teledermatology CL	In-Person CL
Acneform eruption	28	7.7	9.6
Benign tumor	81	7.5	9.0
Infectious	31	7.5	9.2
Pigmented lesion	34	7.4	8.8
Premalignant/malignant	54	7.2	8.9
Other	27	7.1	9.0
Papulosquamous	105	6.8	9.0
Eczematous	44	5.7	9.0

*CL indicates certainty level (1=absolutely uncertain, 10=absolutely confident).

Table 2.

Correlation of Biopsies Between Teledermatology and In-Person Evaluations (N=404 patients)

	In-Person Evaluators	
	Biopsy Performed, n (%)	No Biopsy Performed, n (%)
Teledermatologist recommended a biopsy (n=160)	91 (22.5%)	69 (17.1%) false positive*
Teledermatologist did not recommend a biopsy (n=244)	28 (6.9%) false negative†	216 (53.5%)

*False positive indicates unnecessary biopsies recommended by a teledermatologist.
†False negative indicates necessary biopsies not recommended by a teledermatologist.

(complete plus partial) was less noticeable (78.4% vs 59.5%) between the 2 groups. A comparison of the differences in the number of biopsies recommended was calculated using the McNemar paired χ^2 test. The differences were statistically significant ($P=.001$).

Of note, 2 of the categories with the highest false-positive biopsy rates (tendency of teledermatologists to recommend unnecessary biopsies) were eczematous and papulosquamous.

Comment

Part I of this article reported a 90.6% overall diagnostic correlation between teledermatology and in-person consultations. Table 3 presents a review of the literature and summarizes the most recent significant articles on teledermatology diagnostic correlation.[1-14]

We found that the diagnostic certainty level correlated well with the type of diagnostic agreement, as did the study by Kvedar et al.[13] Not surprisingly, we found the diagnostic certainty level (1 being absolutely uncertain and 10 being absolutely confident) to be generally lower in the teledermatology group (7) compared with the in-person evaluation (9). We hypothesize that lack of confidence in computer technology, inability to obtain additional patient history, inability to perform in-office tests, and lack of training in teledermatology are some of the major reasons for the lower confidence level shown in teledermatology.

Table 3.

Pertinent Literature Review for Diagnostic Capability of Store and Forward Teledermatology*

Study	No. of Patients	Results	Conclusion(s) Drawn by Authors	Still Images (S&F) or Live Video
Whited et al[1]	126 (168 lesions)	Proportion of agreement among IP examiners for single diagnosis: .54; when differentials were included: 92; digital image consultants had equally reliable proportion of agreement	Digital image consultations resulted in reliable and accurate diagnostic out-comes when compared with clinic-based consultations	Still images (1280× 1000)
Phillips et al[2]	51 (107 skin tumors)	59% absolute diagnostic agreement (VTC vs IP)	Ability to detect skin tumors using TD was not decreased; recommendations to do a biopsy were not significantly affected by telemedicine technology	Interactive VTC
Loane et al[3]	351 (427 diagnoses)	67% diagnostic agreement and 64% management agreement between VTC and IP	High proportion of dermatologic conditions can be successfully managed by real-time TD; diagnostic correlation by same dermatologist (72%) and different dermatologist (64%); management agreement was 62 to 65 (64%); substantial proportion of differences in the TD trial reflects differences between dermatologists rather than problems with VTC	Interactive VTC
Oakley et al[4]	104 (135 skin conditions)	75% correctly diagnosed IP; 82% correlation when differential diagnoses were included	VTC can be used with a reasonable degree of accuracy for dermatologic disease	Interactive VTC
Phillips et al[5]	60 (79 diagnoses)	77.2% absolute diagnostic agreement	There was a reasonable degree of agreement between TDs and IP dermatologists; TDs had a lower degree of confidence in their diagnoses; race, sex, and type of skin lesion did not correlate with diagnostic agreement	Interactive VTC
Lesher et al[6]	60	Complete, partial, disagreement for TD: 78%, 21%, 1%; for IP evaluator: 94%, 6%, 0%	TD is an effective means of diagnosing cutaneous disease; higher probability of partial agreement with TD evaluation	Interactive VTC
Lowitt et al[7]	139 (260 skin conditions)	80% diagnostic agreement	Patient and physicians were satisfied, and diagnostic agreement was high	Interactive VTC
Loane et al[8]	205	64% same or similar management plan	Clinical management of dermatologic conditions possible. 64% found to have appropriate clinical management; 36% had either unable or inappropriate management	Interactive VTC
Oakley et al[9]	100	Live operational system resulted in 20% of patients requiring IP appointment	Operational pilot trial with acceptable results for expansion	Interactive VTC
Gilmour et al[10]	126 (155 diagnoses)	Absolute agreement in 59% of cases; missed a secondary diagnosis in 6%; unable to make a diagnosis in 11%; wrong diagnosis in 4%	Acceptable diagnostic agreement with TD with higher management agreement; follow-up rates similar from both TD and IP evaluation; 50% of patients from TD managed without further intervention from specialist; high satisfaction rate	Interactive VTC

Table 3. (continued)

Study	No. of Patients	Results	Conclusion(s) Drawn by Authors	Still Images (S&F) or Live Video
Harrison et al[11]	657	TD had diagnostic accuracy better than GP (71% vs 49%); TD detected more malignancies than GP (94% vs 70%)	TD is feasible; this study used histologic diagnosis as the gold standard	Conventional photos
Zelickson et al[12]	29 (30 skin conditions)	67% to 88% diagnostic and 70% to 90% treatment agreement (still digital images vs IP 2 days later)	Nursing home TD consults may replace some onsite consultations and appears to be cost effective; overall 85% to cases accurately diagnosed and treated; higher agreement when history and images were both available; no significant morbidity seen with TD	Still image telephone-2 (640×480)
Kvedar et al[13]	116	Agreement 75% when image quality and certainty were high; 61% to 64% overall agreement	Still images can substitute for the dermatologic physical examination in up to 83% of cases; concluded that this study validates S&F concept of telemedicine as applied to dermatology	Still images
Whited et al[14]	12 (13 lesions)	Agreement in 82% of cases	Preliminary data; good agreement between IP and digital examiners on differential diagnosis and biopsy recommendations	Still images

*S&F indicates store and forward; VTC, video teleconferencing; IP, in-person; TD, teledermatology and teledermatologist; and GP, general practitioner.

Given the overall lower diagnostic certainty levels of teledermatologists, higher biopsy rates would be expected in this group. Our study revealed this to be true. Teledermatologists recommended biopsies at an 8.1% higher rate than in-person evaluators. Potentially serious dermatologic conditions, namely cutaneous malignancies, were detected in all cases but one. The significance of this one misdiagnosis using teledermatology is uncertain because the true number of cases in which face-to-face dermatologists missed a cutaneous malignancy is not known. Most studies, however, have reported no differences in the recommendations for biopsies between in-person evaluators and teledermatologists. This increase in the biopsy rate may be clinically significant because it may represent an increase in morbidity, which subsequently can increase the cost of dermatologic care. However, it is more likely that this increase in the biopsy rate is transient and will normalize as S&F teledermatology training improves and becomes more familiar to dermatologists. Moreover, when compared with the potential cost savings to the patient and the healthcare system through effective triage, this small increase in cost and morbidity is easily offset. Lastly, given the small number of biopsy recommendations studied, it would be premature to conclude that S&F teledermatology produces a higher morbidity.

Our study had a few limitations. One is that we used the diagnosis rendered during an in-person dermatologist consultation as the gold standard, as did other studies found in the literature review. Use of a face-to-face consultation as a gold standard has significant limitations; however, there appears to be no other good alternatives. Other studies have utilized histologic diagnosis as a gold standard, but cutaneous pathology also has its limitations—it only can be used as a gold standard in very specific cutaneous neoplasms and dermatoses. Therefore, in the absence of a true gold standard, utilization of diagnostic correlation from teledermatology and in-person consultations represents a comparison of teledermatology to the current standard of care (face-to-face evaluations).

Another limitation to our study is that, theoretically, it may be argued that teledermatology might bias the diagnosis rendered during a subsequent face-to-face consultation. However, because the physician in the face-to-face consultation can obtain more patient history, perform unrestricted cutaneous examinations and in-office laboratory tests, and have a better level of comfort in his or her evaluation, we believe this bias is not clinically significant. Rather, we believe the dermatologist would disregard any possible inconsistent history or physical findings from the teledermatology consultation, and base a diagnosis on the in-person findings.

It also may be argued that our use of multiple dermatologists with varied experiences garnered over the course of 1 to more than 20 years may make interpretation of the diagnostic correlation somewhat difficult. Although the more experienced dermatologists may be expected to perform better in the teledermatology consultation and thus have a higher correlation rate, our study showed no significantly different correlation rates between dermatology residents and board-certified dermatologists.

Lastly, there are certain limitations inherent to S&F teledermatology. These include image quality—a limitation that can be minimized by improving technology and training—the skill of the photographer, limited patient history, lack of interaction with the patient, the inability to palpate lesions, and the inability to perform simple laboratory tests such as a skin scraping (potassium hydroxide). Despite these limitations, the results from our study and others reported in the literature indicate that the S&F teledermatology consult system seems to be clinically effective in the appropriate setting. Although an agreement discord existed in about 30% of cases, most represented no clinical significance (no harm done to the patient). However, the 12.5% of cases that had a clinically significant incorrect diagnosis in the teledermatology consultation when compared with in-person evaluation cannot be ignored and should be regarded as a potential risk when utilizing S&F teledermatology. But it is not known if this 12.5% represents a true error rate when compared with an in-person diagnosis. We believe that this 12.5% may closely represent the diagnostic uncertainty that exists within the daily practice of traditional, clinic-based dermatology. Furthermore, it should be stressed that among the 12.5% of cases, there was only one case in which a teledermatologist failed to detect a cutaneous malignancy (basal cell carcinoma).

Given its low morbidity and the high diagnostic correlation of teledermatology to an in-person consultation, we confirm previously published studies indicating that teledermatology is an acceptable method of providing quality, expert care in an expedient and low-cost manner, especially to patients who do not have ready access to a dermatologist.

Acknowledgment—The authors would like to thank Mrs. Linda McKnight for the organization and the data collection of this study; the staff and residents of the Dermatology Service, Walter Reed Army Medical Center, for their hard work and support in this study; and John D. Whited, MD, for his critical review of and recommendations for this manuscript.

REFERENCES

1. Whited JD, Russell HP, Simel DL, et al. Reliability and accuracy of dermatologists' clinic-based and digital image consultations. *J Am Acad Dermatol*. 1999;41:693-702.
2. Phillips CM, Burke WA, Allen MH, et al. Reliability of telemedicine in evaluating skin tumors. *Telemed J*. 1998;4:5-9.
3. Loane MA, Corbett R, Bloomer SE, et al. Diagnostic accuracy and clinical management by realtime teledermatology: results from the Northern Ireland arms of the UK Multicentre Teledermatology Trial. *J Telemed Telecare*. 1998;4:95-100.
4. Oakley AMM, Astwood DR, Loane M, et al. Diagnostic accuracy of teledermatology: results of a preliminary study in New Zealand. *N Z Med J*. 1997;110:51-53.
5. Phillips CM, Burke WA, Shechter A, et al. Reliability of dermatology teleconsultations with the use of teleconferencing technology. *J Am Acad Dermatol*. 1997;37:398-402.
6. Lesher JL, Davis LS, Gourdin FW, et al. Telemedicine evaluation of cutaneous diseases: a blinded comparative study. *J Am Acad Dermatol*. 1998;38:27-31.
7. Lowitt MH, Kessler II, Kauffman L, et al. Teledermatology and in-person examinations. *Arch Dermatol*. 1998;134:471-476.
8. Loane MA, Gore HE, Corbett R, et al. Preliminary results from the Northern Ireland arms of the UK Multicentre Teledermatology Trial. *J Telemed Telecare*. 1998;4(suppl 1):3-5.
9. Oakley AMM, Duffill MB, Reeve P. Practising dermatology via telemedicine. *N Z Med J*. 1998;111:296-299.
10. Gilmour E, Campbell SM, Loane MA, et al. Comparison of teleconsultations and face-to-face consultations: preliminary results of a United Kingdom multicentre teledermatology study. *Br J Dermatol*. 1998;139:81-87.
11. Harrison PV, Kirby B, Dickinson Y, et al. Teledermatology—high technology or not? *J Telemed Telecare*. 1998;4:31-32.
12. Zelickson BD, Homan L. Teledermatology in the nursing home. *Arch Dermatol*. 1997;133:171-174.
13. Kvedar JC, Edwards RA, Menn ER, et al. The substitution of digital images for dermatologic physical examination. *Arch Dermatol*. 1997;133:161-167.
14. Whited JD, Mills BJ, Hall RP, et al. A pilot trial of digital imaging in skin cancer. *J Telemed Telecare*. 1998;4:108-112.