

Comparing High Definition Live Interactive and Store-and-Forward Consultations to In-Person Examinations

Richard Marchell, MD,¹ Craig Locatis, PhD,²
Gene Burges, MD, PhD,^{3,‡} Richard Maisiak, PhD,⁴ Wei-Li Liu, MBA,²
and Michael Ackerman, PhD²

¹Department of Dermatology and Dermatologic Surgery,
Medical University of South Carolina, Charleston, South Carolina.

²National Library of Medicine, National Institutes of Health,
Bethesda, Maryland.

³Department of Dermatology and Dermatologic Surgery,
Medical University of South Carolina, (Deceased).

⁴University of Alabama at Birmingham, Birmingham, Alabama.

[‡](Deceased).

Abstract

Background: There is little teledermatology research directly comparing remote methods, even less research with two in-person dermatologist agreement providing a baseline for comparing remote methods, and no research using high definition video as a live interactive method. **Objective:** To compare in-person consultations with store-and-forward and live interactive methods, the latter having two levels of image quality. **Methods:** A controlled study was conducted where patients were examined in-person, by high definition video, and by store-and-forward methods. The order patients experienced methods and residents assigned methods rotated, although an attending always saw patients in-person. The type of high definition video employed, lower resolution compressed or higher resolution uncompressed, was alternated between clinics. Primary and differential diagnoses, biopsy recommendations, and diagnostic and biopsy confidence ratings were recorded. **Results:** Concordance and confidence were significantly better for in-person versus remote methods and biopsy recommendations were lower. Store-and-forward and higher resolution uncompressed video results were similar and better than those for lower resolution compressed video. **Limitations:** Dermatology residents took store-and-forward photos and their quality was likely superior to those normally taken in practice. There were variations in expertise between the attending and second and third year residents. **Conclusion:** The superiority of in-person consultations suggests the tendencies to order more biopsies or still see patients in-person are often

justified in teledermatology and that high resolution uncompressed video can close the resolution gap between store-and-forward and live interactive methods.

Keywords: technology, telecommunications, teledermatology, telemedicine

Background

Teledermatology can be live interactive, employing videoconferencing technology for synchronous examination, or store-and-forward, with photographs and histories sent to consulting dermatologists for later asynchronous evaluation.¹⁻⁹ Store-and-forward images can be more than eight times the resolution of live interactive, but diagnoses are delayed and, if images and histories are poor or incomplete, another store-and-forward or in-person consultation may be required. Live interactive examinations are immediate and allow image adjustments, but take longer and constrain consultation time and location.⁵ Diagnostic agreement between examinations done remotely and in-person is considered the most appropriate standard for judging telemedicine interventions, since parity with face-to-face assessments, not superiority, needs to be proven.^{1,2,7-9}

Studies of remote agreement with face-to-face assessments can be complete (with identical primary diagnoses) or partial (with one of the specialists including the primary diagnosis of the other in their differential).¹⁰ Partial agreement is always higher than complete, since the agreement threshold is relaxed. Some studies also report aggregate agreement (a sum of complete and partial).⁶

Teledermatology research of diagnostic concordance with in-person examinations has been criticized because usually the diagnoses of one teledermatologist and one clinical dermatologist are compared. Measuring agreement of two in-person clinical examiners is needed to establish a valid baseline.^{2,6,8} Only two teledermatology research reviews specifically looked at concordance in teledermatology studies having baseline inter-observational agreement for in-person exams.^{7,8} The earliest review⁷ had only one store-and-forward and one live interactive study,^{10,11} while a later review⁸ identified 12 studies with multiple dermatologist evaluations.¹⁰⁻²¹

In-person agreement is reported in only four,¹²⁻¹⁶ however, one of which¹³ appears to be a pilot for another.¹⁰ The other studies had more than one teledermatologist.¹⁴⁻²¹ Of the four studies, measuring inter-observational agreement for in-person exams, two showed significantly better agreement among in-person clinicians than distant teledermatologists,^{11,12} especially for primary diagnoses, and two did not.^{10,13} A PubMed search for teledermatology research done after the latest research reviews published in 2011⁹ identified only one additional study having two in-person dermatologists.²² In-person primary diagnosis agreement was 83.3% and agreement between in-person and remote dermatologists ranged from 78.2% to 83.9%.

Teledermatology research reviews report highly variable rates of agreement in different studies. The reviews differ on agreement ranges depending on when they were conducted, the studies they included and excluded, whether they separate agreement for live interactive and store-and-forward interventions; whether complete, partial, or aggregate agreement is reported; and the statistics used to quantify agreement. Statistics usually reported are raw percentages or kappa coefficients accounting for chance agreement and which method to use depends on specific features of a study's research design. In addition, most reviews do not account for studies having a two in-person consultation baseline and those that do not.

The two most recent reviews published in 2011, covering a broad range of studies indicate complete diagnostic agreement ranges of 48-94%⁴ or 46-88%⁹ for store-and-forward and 57-78% for live interactive studies reporting raw percentages.^{4,9} The way each review accounted for differential agreement, by aggregate agreement for different types of lesion⁹ or partial agreement in individual studies⁴ makes the reviews' overall range comparisons for partial agreement difficult. Moreover, review classifications of studies as either store-and-forward or live interactive do not inform about the specific type of technology employed or the resolution of the images or video. This is understandable since the studies themselves often omit these details.

Most research focuses on either store-and-forward or live interactive interventions independently, with few direct comparisons.⁴ Three studies compared the two modes.^{14,23,24} In one study,²³ patients were not seen in-person, while another found identical diagnoses with in-person for 64% of patients with greater agreement for live interactive than store-and-forward that was not statistically significant.²⁴ A third study showed combining methods significantly increases concordance with in-person exams.¹⁴

Although live interactive and store-and-forward methods have been compared before^{14,23,24} the studies comparing remote methods to in-person^{14,24} used diagnoses of single in-

person dermatologists. This study extends teledermatology research by directly comparing concordance between in-person, live interactive, and store-and-forward methods with two in-person dermatologists establishing a diagnostic comparison baseline, while also addressing confidence and biopsy decisions and effects of video quality in live interactive consultations. With the exception of confidence,²⁴ these variables have not been addressed in direct comparisons of methods and the very high resolution video assessed in this study has never been tested.

Materials and Methods

This study was a quasi-randomized control trial, in that clinics were scheduled whenever the number of dermatology referral patients volunteering for the study exceeded 10. Patients were referred from other clinics at the university where the study was conducted and nearby collaborating clinics and they were compensated for time and travel. The study's 214 patients were evaluated 3 times in a single clinical session; in-person, by either high definition uncompressed or compressed video, and by store-and-forward methods. Uncompressed video was 1920 by 1080 pixels transmitted at almost 1.5 gigabits per second, while compressed video was 1280 by 720 pixels transmitted at about two megabits per second. Each

Table 1. Most Common In-Person Consensus Diagnoses and Frequencies—Total Cases = 162

Benign nevus	76
Seborrheic keratosis	23
Dermatofibroma	8
Dysplastic nevus	8
Lentigo	8
Acne	5
Actinic keratosis	5
Cyst	4
Lichen simplex chronicus	4
Eczematous dermatitis	3
Hand dermatitis	3
Hemangioma	3
Postinflammatory pigment Alteration	3
Psoriasis	3
Scabies	3
Skin Tag	3

COMPARING REMOTE AND IN-PERSON DERMATOLOGY CONSULTATIONS

Table 2. Mean Proportion Agreement on Top 1, 2, and Partial (In Differential) Diagnoses Between In-Person Consensus, In-Person Attending, and In-Person Resident

Diagnosis	IN-PERSON ATTENDING			IN-PERSON RESIDENT		
	Top 1	Top 2	In differential	Top 1	Top 2	In differential
In-person consensus	0.98	1.0	1.0	0.91	0.98	1.0
In-person attending	NA	NA	NA	0.87 ^a	0.96	1.0

All proportions were significantly greater than zero, Chi Square test.

^aTop 1 versus top 2, $p=0.0005$, top 1 versus in differential $p<0.0000001$, top 2 versus in differential, $p=0.06$. McNemar test, exact method.

NA, Not Applicable.

videoconferencing system was installed in a clinic examination room and had pan, tilt, and zoom cameras that could be remotely controlled from a teledermatology consultation room outside the examination area.

Type of video alternated between clinics. Patients were taken to the teleconferencing examination room, introduced to the teledermatologist on screen, and were left alone for examination. Store-and-forward work ups followed a protocol having a standardized form for history taking and required a minimum of three 10 megapixel JPEG images (3648×2736 pixel 24 bit color), each including a ruler and color wheel. The order patients experienced the three methods rotated between clinics as did the dermatology residents assigned each method. An attending, board-certified dermatologist, however, always saw patients in person along with a resident assigned to that method.

The attending and in-person resident reached consensus on the differential and primary diagnoses that were used to de-

termine remote exam concordance. To provide a better baseline, the attending and in-person residents made separate independent differentials and diagnoses before consensus for a subset of 134 patients. These were compared to each other and to the consensus to keep the standard for scoring all cases consistent.

A form was used in each treatment where the primary diagnosis was listed first and alternative diagnoses were listed in order of likelihood. The residents and attending also indicated whether biopsy was needed and they rated their confidence in primary diagnosis and biopsy decisions on the five-point scale with one indicating very certain and five very uncertain. The form also had a place for making comments.

Differences between dichotomous variables were tested using either McNemar Exact tests for related cases or Fisher's Exact test. Differences in interval data were tested using nonparametric tests, including the Friedman test for multiple related groups, the Wilcoxon Signed Rank test for two related groups, or the Mann-Whitney test of independent groups. Kappa coefficients were calculated for biopsy agreement. All tests were done with the statistical package SPSS and had a two-tailed significance threshold of 0.05. The study was approved by the Institutional Review Boards of the Medical University of South Carolina and the National Institutes of Health.

Results

Top diagnoses are listed in *Table 1*, which constitute over 75% of the cases. Concordance between the in-person residents' and the attending's primary, secondary, and entire differential diagnoses and concomitant consensus diagnoses for the 134 patient subsample are shown in *Table 2*. The attending and in-person residents had high agreement, both

Table 3. Mean Proportion Agreement with Consensus Diagnosis for Top 1, 2, and Partial (In Differential) Diagnoses of Residents by Diagnostic Method

Agreement category	DIAGNOSTIC METHOD											
	IN-PERSON			STORE-AND-FORWARD			UNCOMPRESSED VIDEO			COMPRESSED VIDEO		
	Top 1	Top 2	In differential	Top 1	Top 2	In differential	Top 1	Top 2	In differential	Top 1	Top 2	In differential
Proportion	0.91 ^a	0.98 ^b	1.0 ^c	0.76	0.85	0.87	0.76	0.83	0.89	0.72	0.87	0.88
N	134	134	134	213	213	213	101	101	101	112	112	112
Std. Dev.	0.29	0.15	0	0.43	0.36	0.33	0.43	0.38	0.31	0.45	0.34	0.32

^aSignificantly higher than overall store-and-forward ($p=0.002$), compressed video ($p=0.003$) or uncompressed video ($p=0.03$), McNemar test.

^bSignificantly higher than overall store-and-forward ($p=0.004$), compressed video ($p=0.02$) and uncompressed video ($p=.03$), McNemar test.

^cSignificantly higher than overall store-and-forward ($p=0.001$), compressed video ($p=0.004$) and uncompressed video ($p=0.03$), McNemar tests.

Differences between the remote methods were not significant ($p<0.05$), Friedman test.

Table 4. Mean Proportion of Cases with Biopsy Recommendations by Diagnostic Method

	DIAGNOSTIC METHOD					
	IN-PERSON CONSENSUS	IN-PERSON ATTENDING	IN-PERSON RESIDENT	STORE-AND-FORWARD	UNCOMPRESSED VIDEO	COMPRESSED VIDEO
Proportion	0.04*	0.02	0.01	0.11	0.08	0.12
N	207	129	128	209	112	98
Standard deviation	0.19	0.15	0.08	0.31	0.28	0.33
Total biopsies	8	3	2	23	9	12

*Significantly lower than store-and-forward overall ($p=0.001$) and significantly lower than compressed video ($p=0.04$), McNemar tests.

with each other and the consensus, with the attending’s agreement with the consensus higher. The mean proportions of agreement with the in-person consensus for the in-person attending, in-person residents, store-and-forward residents, and the uncompressed and compressed video residents are shown in Table 3. The mean agreements for the remote methods were significantly ($p<0.05$) lower than the in-person method and similar to each other.

The number and mean proportion of cases with biopsy recommendation stratified by diagnostic method appear in Table 4. There were few suspected cancers and only eight in-person consensus biopsy recommendations. The in-person proportion of biopsy recommendations was significantly lower than for store-and-forward ($p=0.001$) and uncompressed video ($p=0.04$). The Kappa coefficient (0.43) of agreement between uncompressed video and the in-person consensus for biopsy recommendation was significantly greater than zero ($p=0.001$) as was the Kappa coefficient (0.35) for

Table 5. Mean Confidence Rating for Diagnoses and Biopsy Recommendations by Diagnostic Method

TYPE OF DIAGNOSTIC ASSESSMENT	DIAGNOSTIC METHOD			
	IN PERSON CONSENSUS ^a	STORE-AND-FORWARD IMAGE	UNCOMPRESSED VIDEO ^b	COMPRESSED VIDEO
Diagnosis ($n=201$)	1.42 (0.66) $n=201$	1.74 (1.05) $n=201$	1.72 (0.99) $n=95$	2.51 (1.23) $n=106$
Differential ($n=201$)	1.39 (0.64) $n=201$	1.68 (0.97) $n=201$	1.72 (1.00) $n=95$	2.31 (1.17) $n=106$
Biopsy ($n=199$)	1.19 (0.50) $n=199$	1.51 (0.78) $n=199$	1.57 (0.93) $n=93$	2.16 (1.18) $n=106$

Lower values indicate greater confidence and standard deviations are in parentheses.
^aAll mean ratings for in-person were significantly ($p<0.001$) lower than any remote method, Wilcoxon Signed Rank test.
^bThe mean ratings for uncompressed video were significantly ($p<0.001$) lower than for compressed video for each assessment type, Mann-Whitney test.

store-and-forward ($p=0.001$), while the Kappa coefficient between compressed video and the consensus was low and not significant ($p=0.23$).

Confidence ratings are presented in Table 5. There was significantly less confidence in diagnosis, differential diagnoses, and biopsy decisions for remote methods than for in-person ($p<0.001$) and there were no significant differences in confidence between store-and-forward and un-

compressed live interactive methods. Mean confidence in diagnosis, differential, and biopsy recommendation was significantly lower ($p<0.001$) for compressed video versus the uncompressed video and store-and-forward methods.

Discussion

High levels of diagnostic agreement, diagnostic confidence, and decisions to biopsy for in-person exams significantly contrasted with those for remote methods. The in-person residents’ independent primary and secondary diagnoses agreed with the attending’s in 87% and 96% of the cases and matched the top primary and secondary consensus diagnoses in 91% and 98% of the cases. When the entire differential is considered, there was partial agreement with the attending’s diagnosis and consensus diagnosis 100% of the time (Table 2).

On average, primary diagnoses using remote methods matched the in-person consensus diagnosis about 75% of the time. Agreement for remote methods improved when secondary diagnoses were considered and improved even more if the consensus diagnosis appeared anywhere in differentials (Table 3).

Store-and-forward and both video methods had similar agreement and decisions to biopsy, but store-and-forward and uncompressed confidence levels were significantly higher than those for compressed video. The finding that these variables were significantly different between in-person and all remote exams conforms to two other previous studies having two in-person agreement baselines.^{11,12}

The uniformly lower confidence for compressed video conflicts somewhat with the results of some earlier live interactive studies, many likely conducted with standard definition video at transmission rates well below those for compressed video in this study.

The confidence ratings for uncompressed video and store-and-forward methods in this study were similar and higher than those for compressed video. This parity indicates uncompressed video can close the resolution gap between live interactive and store-and-forward methods, preserving the benefit of immediately collecting additional information. One limitation of this study is the store-and-forward photographs were very high resolution, following a strict protocol, and were taken by highly knowledgeable dermatology residents, which may have inflated the method's concordance and confidence levels.

Another limitation is varied expertise of the attending and the residents and the attending always evaluating patients in person. The residents, however, were all second and third year, the cases patients presented were very typical, and there was still very high agreement between the residents and attending in the in-person method. Since residents rotated between methods, any variance in expertise would have likely been distributed equally among methods. If attending dermatologists were used across all methods, the agreement levels for all methods might be higher, but whether they would be so much higher for remote methods as to produce different results is uncertain since in-person agreement might increase as well. Finally, the decision to biopsy results are significant but inconclusive given the small number of cases.

Conclusion

Diagnoses, decisions to biopsy, and diagnostic confidence for teledermatology consultations differ from those done in-person. Of the remote methods tested, uncompressed live interactive and store-and-forward methods had similar results and, although significantly worse than in-person, were significantly better than compressed video. Compressed video performed poorly on most measures and is not recommended unless used in conjunction with high resolution photography as other studies suggest.^{14,24} Uncompressed video is not a turnkey technology and adopting it versus store-and-forward depends on network infrastructure and technical support and whether protocols and training are sufficient to ensure high quality still image capture.

This study, like most, found some level of agreement for remote methods (higher than chance) and, like others, offers evidence of teledermatology's reliability, since teledermatology may be the only option for many patients²⁵ and always less risky than no assessment. When malignancies and other conditions

having considerable consequences are suspect, however, additional measures are needed.²⁶ The higher propensity to biopsy and overall lower confidence for remote methods found in this study not only reinforce earlier research suggesting biopsy an indicator of uncertainty,^{27,28} but also suggests these biopsies are probably clinically justified as a precaution.

Acknowledgments

The authors thank participating residents Vivian Beyer, Kathryn Dempsey, Brad Greenhaw, Francesa Lewis, Nick Papa-john, Adam Perry, Adam Sperduto, Roger Sullivan, Julie Swick, and Brent Taylor. This study was supported by NIH Research Contracts HHSN276201100424P, HHSN276201100588P, and the NIH/NLM Intramural Research Programs.

Disclosure Statement

No competing financial interests exist.

REFERENCES

1. Hersch W, Wallace J, Patterson P, Shapiro S, Kraemer D, Eilers G, Chan B, Greenlick M, Helfand M. Telemedicine for the medicare population. Evidence Reports/Technology Assessments, No. 24. Rockville, MD: Agency for Healthcare Research and Quality, 2001. Report No. 01-E012.
2. Hersch W, Hickam D, Severance S, Dana T, Krages K, Helfand M. Telemedicine for the medicare population: Update. Evidence Reports/Technology Assessments, No. 131, Rockville, MD: Agency for Healthcare Research and Quality, 2006. Report No. 06-E007.
3. Eminovic N, de Keizer N, Bindels P, Hasman A. Maturity of teledermatology evaluation research: A systematic review. *Br J Derm* 2007;156:412-419.
4. Johnson M, Armstrong A. Technologies in dermatology: Teledermatology review. *G Ital Dermatol Venereol* 2011;146:143-153.
5. Romero G, Garrido J, Garcia-Arpa M. Telemedicine and teledermatology (I): Concepts and applications. *Actas Dermosifiliogr* 2008;99:506-522.
6. Romero G, Cortina P, Vera E. Telemedicine and teledermatology (II): Current state of research on dermatology teleconsultations. *Actas Dermosifiliogr* 2008;99:586-597.
7. Whited J. Teledermatology research review. *Int J Derm* 2006;45:220-229.
8. Levin Y, Warshaw E. Teledermatology: A review of reliability and accuracy of diagnosis and management. *Dermatol Clin* 2009;27:163-176.
9. Warshaw E, Hillman Y, Greer N, Hage E, MacDonald R, Rutks I, Wilt T. Teledermatology for diagnosis and management of skin conditions: A systematic review. *J Am Acad Dermatol* 2011;64:759-772.
10. Whited J, Hall R, Simel D, Foy M, Stechuchak K, Drugge R, Grichnik J, Myers S, Horner R. Reliability and accuracy of dermatologists' clinic-based and digital image consultations. *J Am Acad Dermatol* 1999;41:693-702.
11. Leshner J, Davis L, Gourdin F, English D, Thompson W. Telemedicine evaluation of cutaneous diseases: A blinded comparison study. *J Am Acad Dermatol* 1998;38:27-31.
12. Browns I, Collins K, Walters S, McDonagh A. Telemedicine in dermatology: A randomized controlled trial. *Health Technol Assess* 2006;10:1-39.
13. Whited J, Mill B, Hall R, Drugge R, Grichnik, Simel D. A pilot trial of digital imaging in skin cancer. *J Telemed Telecare* 1998;4:108-112.
14. Baba M, Seckin D, Kapdagli S. A comparison of teledermatology using store-and-forward methodology alone, and in combination with Web camera videoconferencing. *J Telemed Telecare* 2005;11:354-360.

15. Massone C, Hofmann-Wellenhof R, Ahlgrimm-Siess V, Gabler G, Ebner C, Soyer HP. Melanoma screening with cellular phones. *PLoS One* **2007**;2:e483.
16. Moreno-Ramirez D, Ferrandiz L, Nieto-Garcia A, Carrasco R, Moreno-Alvarez P, Galdeano R, Bidegain E, Rios-Martin J, Camacho F. Store and forward teledermatology in skin cancer triage. *Arch Dermatol* **2007**;143:479-484.
17. Moreno-Ramirez D, Ferrandiz L, Bernal AP, Duran RC, Martin JJ, Camacho F. Teledermatology as a filtering system in pigmented lesion clinics. *J Telemed Telecare* **2005**;11:298-303.
18. Oztas MO, Calikoglu E, Baz K, Birol A, Onder M, Calikoglu T, Kitapci M. Reliability of Web-based teledermatology consultations. *J Telemed Telecare* **2004**;10:25-28.
19. Lim A, Egerton I, See A, Shumack S. Accuracy and reliability of store-and-forward teledermatology: Preliminary results from the St George teledermatology project. *Australas J Dermatol* **2001**;42:247-251.
20. Krupinski E, LeSuer B, Ellsworth L, Levine N, Hansen R, Silvis N, Sarantopoulos P, Hite P, Wurzel J, Weinstein R, Lopez AM. Diagnostic accuracy and image quality using a digital camera for teledermatology. *Telemed J* **1999**;5:257-263.
21. Piccolo D, Soyer HP, Chimenti S, Argenziano G, Bartenjev I, Hofmann-Wellenhof R, Marchetti R, Oguchi S, Pagnanelli G, Pizzichetta MA, Saida T, Salvemini I, Tanaka M, Wolf IH, Zgavec B, Peris K. Diagnosis and categorization of acral melanocytic lesions using teledermoscopy. *J Telemed Telecare* **2004**;10:346-350.
22. Ribas J, da Graca Souza Cunha M, Mendes Schettini AP, da Rocha Ribas CB. Agreement between dermatological diagnoses made by live examination compared to analysis of digital images. *An Bras Dermatol* **2010**;85:441-447.
23. Loane M, Bloomer S, Corbett R, Eedy D, Hicks N, Lotery H, Mathews C, Paisley J, Steele K, Wootton R. A comparison of real-time and store-and-forward teledermatology: A cost-benefit study. *Br J Dermatol* **2000**;143:1241-1247.
24. Edison K, Ward D, Dyer J, Lane W, Chance L, Hicks L. Diagnosis, diagnostic confidence, and management concordance in live-interactive and store-and-forward teledermatology compared to in-person examination. *Telemed J E Health* **2008**;14:889-895.
25. Coates S, Kvedar J, Granstein R. Teledermatology: From historical perspective to emerging techniques of the modern era. *J Am Acad Dermatol* **2015**;72:563-574.
26. Tandjung R, Badertscher N, Kleiner N, Wensing M, Rosemann T, Braun R, Senn O. Feasibility and diagnostic accuracy of teledermatology in Swiss primary care: Process analysis of a randomized control trial. *J Eval Clin Pract* **2015**;21:326-331.
27. Pak H, Harden D, Cruess D, Welch M, Poropatich R; National Capital Area Teledermatology Consortium. Teledermatology: An intraobserver diagnostic correlation study, Part I. *Cutis* **2003**;71:399-403.
28. Pak H, Harden D, Cruess D, Welch M, Poropatich R; National Capital Area Teledermatology Consortium. Teledermatology: An intraobserver diagnostic correlation study, Part II. *Cutis* **2003**;71:476-480.

Address correspondence to:

Richard Marchell, MD
 Department of Dermatology and Dermatologic Surgery
 Medical University of South Carolina
 Charleston, SC 29425

E-mail: locatis@nlm.nih.gov

Received: April 19, 2016

Revised: May 26, 2016

Accepted: June 3, 2016