
S. Vañó-Galván, a,∗ Á. Hidalgo, b I. Aguayo-Leiva, a M. Gil-Mosquera, c L. Ríos-Buceta, a M.N. Plana, d J. Zamora, d A Martorell-Calatayud, e P. Jaén a

Servicio de Dermatología, Hospital Ramón y Cajal, Universidad de Alcalá, Madrid, Spain
Associate Professor in Fundamentals of Economic Analysis, Universidad de Castilla-La Mancha, Toledo, Spain
Medicina de Familia, Hospital Ramón y Cajal, Madrid, Spain
Unidad de Bioestadística Clínica, Hospital Universitario Ramón y Cajal. CIBER en Epidemiología y Salud Pública (CIBERESP), Instituto de Investigación Sanitaria (IRYCS). Madrid, Spain
Servicio de Dermatología, Hospital de Requena, Valencia, Spain

Manuscript received May 30, 2010; accepted for publication November 3, 2010

Keywords
Teledermatology; Telemedicine; Store-and-forward; Emergency services; Validity; Reliability

Abstract

Background and objectives: The aim of this study was to assess the validity of store-and-forward teledermatology as a tool to support physicians in primary care and hospital emergency services and reduce the requirement for face-to-face appointments. Diagnostic validity and the approach chosen for patient management (face-to-face vs teledermatology) were compared according to patient origin and diagnostic group. Material and methods: Digital images from 100 patients were assessed by 20 different dermatologists and the diagnoses offered were compared with those provided in face-to-face appointments (gold standard). The proposed management of the different groups of patients was also compared. Results: The percentage complete agreement was 69.05% (95% confidence interval [CI], 66.9%-71.0%). The aggregate agreement was 87.80% (95% CI, 86.1%-89.0%). When questioned about appropriate management of the patients, observers elected face-to-face consultation in 60% of patients (95% CI, 58%-61%) and teledermatology in 40% (95% CI, 38%-41%). Diagnostic validity was higher in patients from primary care (76.1% complete agreement and 91.8% aggregate agreement) than those from hospital emergency services (61.8% complete agreement, 83.4% aggregate agreement) (P < .001) and teledermatology was also chosen more often in patients from primary care compared with those from emergency services (42% vs 38%; P = .003). In terms of diagnostic group, higher validity was observed for patients with infectious diseases (73.3% complete agreement and 91.3% aggregate agreement) compared to those with inflammatory disease (70.8% complete agreement and 86.4% aggregate agreement) or tumors (63.0% complete agreement and

∗Corresponding author.
E-mail address: sergiovano@yahoo.es (S. Vañó-Galván).
Introduction

Teledermatology is the clinical evaluation of skin lesions by dermatologists using telemedicine techniques; it allows patients to be diagnosed and, possibly, treated at a distance. There are 2 types of teledermatology: real-time (synchronous) teledermatology and store-and-forward (asynchronous) teledermatology. In the former, the dermatologist analyzes digital images of the patients as soon as these are taken and provides an immediate diagnosis. In the latter, the dermatologist analyzes stored images obtained at an earlier point in time and then provides a diagnosis and management plan.

Several studies have analyzed the reliability and validity of teledermatology. The level of diagnostic agreement reported for real-time teledermatology ranges from 0.69 (complete agreement) to 0.84 (aggregate agreement). The corresponding figures for store-and-forward teledermatology are 0.60 and 0.80, respectively. Few studies that have compared teledermatology and face-to-face consultations have analyzed diagnostic agreement by pathology and those that have have produced conflicting results. It should be noted, however, that agreement has generally been found to be higher for tumors than for inflammatory disease. In our review of the literature, we found no studies that had compared diagnostic agreement in teledermatology according to the origin of the patient (primary care vs hospital emergency services).

Study Objectives

1. To evaluate the validity of store-and-forward teledermatology. The concept of validity refers to the ability of a tool to measure what it is intended to measure. In other words, it reflects how a particular test performs compared to a gold standard. Validity is measured by level of diagnostic agreement.

2. To evaluate the possible application of store-and-forward teledermatology as a support tool for primary care and as a triage tool for hospital emergency services. The usefulness of the application is determined by its capacity to reduce the number of face-to-face consultations.

3. To compare diagnostic validity and patient management approaches (face-to-face consultations vs teledermatology) according to patient origin (primary care vs emergency departments) and diagnostic group (inflammatory disease, infectious diseases, and tumors).
Material and Methods

Informed consent was obtained from all the patients who participated in the study. Digital images of 100 patients, together with a brief clinical history taken according to a standardized protocol, were forwarded to 20 dermatologists (observers) for diagnostic/therapeutic evaluation. The level of diagnostic agreement between store-and-forward teledermatology and face-to-face consultation (the gold standard) was calculated for each observer and for the group of observers as a whole (validity study). The management approaches chosen by each observer were also compared to those recommended by the dermatologists who performed the face-to-face consultations to evaluate the usefulness of store-and-forward teledermatology as a support tool for primary care and hospital emergency services. For the second analysis, we calculated the most common management approach suggested for a given patient by the 20 observers and then calculated the overall frequencies for the approaches used in the 100 patients.

1. Type of study. We performed a descriptive repeated measures cross-sectional study.

2. General information. The study was performed between January and April 2009. Patients presenting at primary care or emergency department facilities were seen by a physician (not a dermatologist) who produced a brief clinical history following a standardized protocol and took 1 or more digital photographs of the patient's lesions. Minutes later, the patient was seen by a dermatologist who conducted a routine evaluation and established a clinical diagnosis based on the results of this evaluation and additional tests such as biopsy when considered necessary. This diagnosis was established as the gold standard diagnosis.

3. Patient selection. Two groups of patients were formed: one consisting of 50 patients seen in a primary care dermatology outpatient clinic and another consisting of 50 patients from the emergency department at Hospital Ramón y Cajal in Madrid, Spain. The patients were recruited consecutively and they all gave their informed consent for the digital photographs to be taken. The nonspecialist physicians who saw the patients during the initial visit were primary care physicians or first-year residents.

4. Digital images and clinical history. The physicians/residents were briefly instructed on how to take the digital images prior to the study. All the lesions were photographed using a digital Olympus camera. The clinical history was taken following a brief protocol covering 5 items (Appendix 1).

5. Presentation of data. The images and clinical history were shown, by projector, to 20 dermatologists in a viewing room. The dermatologists filled in a standardized form that included sections on diagnosis, management decisions, and reasons for referral for face-to-face consultation (Appendix 2). All of the dermatologists completed a questionnaire for each of the 100 patients.

6. Evaluation of diagnostic agreement. Each dermatologist provided a maximum of 3 diagnoses. Complete agreement was achieved when there was just 1 diagnosis and this coincided with the gold standard diagnosis. Partial agreement was achieved when 1 of the diagnoses provided by the dermatologist coincided with the gold standard. Aggregate agreement was calculated as the sum of the complete and partial agreement. It was considered that there was no agreement when none of the diagnoses coincided with the gold standard diagnosis.

7. Statistical methods. The data were analyzed using the software programs SPSS (version 13.0, Chicago, Illinois, USA) and STATA (version 10.0, StataCorp, College Station, Texas, USA). A descriptive analysis was made of patient characteristics. Diagnostic validity was calculated by analyzing complete and aggregate agreement for each observer and for the group of observers as a whole. Diagnostic agreement and management approaches were also analyzed by diagnostic group (inflammatory disease, infectious diseases, or tumors) and patient origin (primary care vs emergency department), with the calculation of 95% confidence intervals (CIs). The \( \chi^2 \) test was used and statistical significance was set at \( P < .05 \).

Results

Descriptive Analysis

The time spent on collecting data for each patient in the face-to-face consultations with the nonspecialist physician was 4 to 6 minutes, and that spent by the dermatologists who remotely analyzed the stored and forwarded data to produce a diagnosis and management strategy was 2 to 3 minutes per patient. The descriptive data (specific diagnoses, diagnostic groups, and patient origin) are summarized in Tables 1 and 2.

Level of Diagnostic Agreement

Complete diagnostic agreement between store-and-forward teledermatology and face-to-face consultation was found in 69.05% (95% CI, 66.9%-71.0%) of cases. The corresponding percentage for aggregate agreement was 87.80% (95% CI, 86.1%-89.0%). The results for patients analyzed by origin and diagnostic group are shown in Table 3. Specifically, diagnostic validity was significantly higher for primary care patients than for emergency department patients (76.1% vs 61.8% for complete agreement and 91.8% vs 83.4% for aggregate agreement; \( P < .001 \) in both cases). and for patients with infectious diseases than for those with inflammatory disease (73.3% vs 70.8% for complete agreement and 91.3% vs 86.4% for aggregate agreement; \( P < .001 \) in both cases) and tumors (73.3% vs 63.0% for complete agreement and 91.3% vs 87.2% for aggregate agreement; \( P < .001 \) in both cases).

Management Approaches Recommended by Observers

Forty percent of the observers opted for a remote management strategy (discharge, referral to a primary care physician, or a second teledermatology evaluation) while 60% referred the patients to a dermatologist for a
priority or nonpriority face-to-face visit. Discharge was the majority option in 21 patients (21%). The corresponding percentages were 15% for a visit with a primary care physician, 4% for a second teledermatology evaluation, 37% for nonpriority referral to a dermatologist, and 23% for priority referral. Among the main reasons mentioned for recommending face-to-face consultation were poor-quality images (5.1%), need for palpation (1.7%) or dermatoscopy (4.9%), in-person treatment (eg, cryotherapy) (15.3%), and other reasons such as patch tests or culture (11.9%).

The majority diagnosis by the observers did not coincide with the gold standard diagnosis (diagnostic errors) for 6 patients, all of whom were referred for face-to-face consultation with a dermatologist. The reasons given for referral were skin biopsy (4 cases) and poor-quality image (2 cases) (Table 4).

The results for the comparison of recommended management approaches by patient origin and diagnostic group are shown in Table 5. Teledermatology was chosen for a greater proportion of primary care patients (42%) than emergency department patients (38%) (P = .003), and for a greater proportion of patients with infectious diseases (52%) than either inflammatory disease (40%) or tumors (28%) (P = .001).

Discussion

We performed a validity analysis of store-and-forward teledermatology in which we compared the level of diagnostic agreement between 20 remotely located observers and onsite dermatologists who provided a gold standard diagnosis following a face-to-face consultation, including a histology study where necessary. Validity studies of store-and-forward teledermatology have yielded varying results, with agreement levels ranging between 0.63 and 1.00. Histologic diagnosis is the usual gold standard for skin lesions. The power of studies in this area is directly related to the total number of observations, which is calculated by multiplying the number of patients by the number of observers. In this respect, the most powerful study published to date is that by Whited et al,23 which comprised 168 patients and 6 observers (total of 1008 observations). One particular strength of our study is the fact that we analyzed 2000 observations (100 patients each analyzed by 20 observers). Our diagnostic agreement results are similar to those reported by High et al,20 and slightly higher than those reported by Whited et al23 for complete agreement (due to differences in methodology) (Table 6).

Studies that have analyzed level of agreement by diagnostic group have conducted reliability analyses and offer varying results, but the majority have reported higher agreement levels in inflammatory disease compared with infectious disease (Table 6). We found a similar trend, with complete agreement ranging from 61.8% (primary care) to 73.3% (emergency department), and aggregate agreement ranging from 76.1% (primary care) to 91.8% (emergency department) (Table 6). The reasons for this difference may be related to the greater complexity and variability of skin infections compared with inflammatory skin conditions.

Table 1 Number of Patients in Study By Diagnosis

<table>
<thead>
<tr>
<th>No. of</th>
<th>Diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Eczema, acne</td>
</tr>
<tr>
<td>6</td>
<td>Viral warts</td>
</tr>
<tr>
<td>5</td>
<td>Seborrheic keratosis</td>
</tr>
<tr>
<td>4</td>
<td>Soft fibromas, melanocytic nevi, seborrhoeic dermatitis</td>
</tr>
<tr>
<td>3</td>
<td>Infectious cellulitis, basal cell carcinoma, molluscum contagiosum, psoriasis, herpes zoster, toxic dermatitis</td>
</tr>
<tr>
<td>2</td>
<td>Insect and spider bites, urticaria, actinic keratosis, squamous cell carcinoma, epidermal cyst, panniculitis, vitiligo, pityriasis rosea</td>
</tr>
<tr>
<td>1</td>
<td>Telogen effluvium, androgenetic alopecia, labial herpes, pityriasis versicolor, tinea, xanthoma, keloid scarring, keratosis pilar, pyoderma gangrenosum, vasculitis, melanoma, alopecia areata, venous lake, Bowen disease, lichen planus, impetigo, black hairy tongue, dermatofibroma, myxoid cyst, lichen sclerosus et atrophicus, rosacea, cutaneous lupus erythematosus, perioral dermatitis, erythema multiforme exudativum, nail dystrophy, granuloma annulare, sebaceous hyperplasia, scabies</td>
</tr>
</tbody>
</table>

Table 2 Diagnostic Groups and Origin of Patients

<table>
<thead>
<tr>
<th>Diagnostic Group</th>
<th>Primary Care</th>
<th>Emergency Department</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inflammatory disease</td>
<td>22</td>
<td>28</td>
<td>50</td>
</tr>
<tr>
<td>Tumors</td>
<td>20</td>
<td>10</td>
<td>30</td>
</tr>
<tr>
<td>Infectious disease</td>
<td>8</td>
<td>12</td>
<td>20</td>
</tr>
<tr>
<td>All diseases</td>
<td>50</td>
<td>50</td>
<td>100</td>
</tr>
</tbody>
</table>

Figures refer to percentage of patients.

Table 3 Level of Complete and Aggregate Agreement and Between-Group Comparisons (2000 Observations)

<table>
<thead>
<tr>
<th>Agreement Level</th>
<th>Overall</th>
<th>Primary Care</th>
<th>Emergency</th>
<th>Inflammatory Disease</th>
<th>Infectious Disease</th>
<th>Tumor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete agreement</td>
<td>69.0 (66.9-71.0)</td>
<td>76.1 (73.3-78.7)</td>
<td>61.8 (58.7-64.8)</td>
<td>70.8 (67.9-73.6)</td>
<td>73.3 (68.6-77.5)</td>
<td>63.0 (59.0-66.9)</td>
</tr>
<tr>
<td>Aggregate agreement</td>
<td>87.8 (86.1-89.0)</td>
<td>91.8 (89.9-93.4)</td>
<td>83.4 (80.9-85.7)</td>
<td>86.4 (84.1-88.5)</td>
<td>91.3 (88.0-93.8)</td>
<td>87.2 (84.2-89.7)</td>
</tr>
</tbody>
</table>

Agreement levels are expressed as percentages and 95% confidence intervals. Complete agreement, primary care group vs emergency group: P < .001; aggregate agreement, primary care group vs emergency group: P < .001; complete agreement, comparison by diagnostic group: P < .001; complete agreement, comparison by diagnostic group: P < .001.
to the fact that these patients tend to have more complex conditions than those who visit primary care facilities. It is noteworthy that all the patients for whom the majority diagnosis provided by the observers did not coincide with the gold standard diagnosis (n=6) were referred for face-to-face consultation as this would have reduced the risk of a true diagnostic error.

We calculated the potential usefulness of store-and-forward teledermatology as a support tool for primary care and a triage tool for emergency services by analyzing the reduction in the requirement for face-to-face consultation. Our results show that such a consultation would have been avoided in 40% of patients; 21% were discharged, 15% were referred to their primary care physician, and 4% were scheduled for a second teledermatology evaluation. It is unlikely that this last option would have been widely considered in a real-life situation as teledermatology is not yet available in the Spanish public health care system.

Recent studies by Moreno-Ramírez et al.\textsuperscript{27} and Moreno et al.\textsuperscript{29} yielded interesting results on the use of store-and-forward teledermatology in skin cancer and pigmented lesions. Our results show a higher level of agreement for infectious diseases, partly because this diagnostic group contains entities that are easy to diagnose visually, such as verruca vulgaris, molluscum contagiosum, herpes simplex, and herpes zoster. One reason why our results for tumors are slightly lower than those reported by other studies may be because we did not include dermatoscopic images.

In our review of the literature, we did not find any teledermatology studies that had analyzed differences in agreement levels according to patient origin (primary care vs emergency services). The lower values observed for patients from the emergency department in our study may be related to the fact that these patients tend to have more complex conditions than those who visit primary care facilities.

It is noteworthy that all the patients for whom the majority diagnosis provided by the observers did not coincide with the gold standard diagnosis (n=6) were referred for face-to-face consultation as this would have reduced the risk of a true diagnostic error.

We calculated the potential usefulness of store-and-forward teledermatology as a support tool for primary care and a triage tool for emergency services by analyzing the reduction in the requirement for face-to-face consultation. Our results show that such a consultation would have been avoided in 40% of patients; 21% were discharged, 15% were referred to their primary care physician, and 4% were scheduled for a second teledermatology evaluation. It is unlikely that this last option would have been widely considered in a real-life situation as teledermatology is not yet available in the Spanish public health care system. On analyzing the usefulness of store-and-forward..
teledermatology by diagnostic group, it was seen that the highest number of face-to-face consultations would have been avoided in patients with infectious disease, possibly because of the higher diagnostic accuracy associated with such diseases (this was also the group with the highest level of diagnostic agreement), but also because confirmation of diagnosis by skin biopsy is necessary in a smaller percentage of patients with infectious diseases than of those with inflammatory disease or tumors, in particular.

The main reason given for referral for face-to-face consultations was the need for skin biopsy (61% of cases). While this figure is high, it should be borne in mind that half of the patients were seen in the emergency department, where lesions tend to be more complex than in primary care. Indeed, patients who visit emergency departments with skin complaints tend to have hyperacute diseases which often require a histologic diagnosis. This, added to the possibility that there is a greater tendency to seek confirmation of remote diagnoses, would explain the large proportion of patients in our series who were referred for a histology study.

One limitation of our study is that the data were obtained under artificial conditions. Indeed, the remote evaluation was purely experimental and patient care was based exclusively on face-to-face examination by a dermatologist. Our results thus do not reflect the variations and unanticipated events that can affect the real-life application of teledermatology. It would therefore be interesting to collect data and perform retrospective analyses of data from dermatology departments that already offer teledermatology services as part of routine clinical practice.

One of the fundamental benefits and justifications for using teledermatology is that it would result in savings for the public health care system, as it is less costly than face-to-face consultation in certain cases. In this study, we did not set out to perform a health economic analysis of teledermatology, but such studies would be of value for investigating whether this technique is more cost-effective than face-to-face consultations with a dermatologist and, if so, to establish in exactly which situations it is appropriate and to establish guidelines on its use. Future studies in this area are likely to analyze the experiences of teledermatology units that are already operating in the public health system to evaluate the economic and health care impact of these units in real-life situations.

Conclusions

1. Store-and-forward teledermatology had a high level of diagnostic validity, particularly in the case of primary care patients and patients with infectious diseases.
2. The technique is useful for the remote diagnosis and management of patients as, based on our results, it would reduce the need for face-to-face consultation by 40%.

Conflict of Interest

The authors declare that they have no conflict of interest.

Acknowledgments

We thank all the members of the Department of Dermatology at Hospital Ramón y Cajal for their help with the conduct of this study.

Appendix 1.

Standardized Clinical History Taken for Each Patient

- REASON FOR CONSULTATION
- SEX, AGE, AND RELEVANT DISEASES
- CURRENT DISEASE (time since onset, clinical manifestations, extent of disease)
- SPECIFIC INFORMATION (recent change of drugs, sexual relationships, etc)
- TREATMENTS RECEIVED and response to these

Appendix 2.

Questionnaire completed by the observer dermatologists

1. DIAGNOSIS OR DIFFERENTIAL DIAGNOSES (maximum 3):
2. PATIENT MANAGEMENT APPROACH
   1. MANAGEMENT BY TELEDERMATOLOGY:
      a) DISCHARGE (with or without treatment)
      b) EVALUATION BY PRIMARY CARE PHYSICIAN
      c) SECOND TELEDERMATOLOGY EVALUATION
   2. FACE-TO-FACE MANAGEMENT:
      d) NONPRIORITY REFERRAL FOR FACE-TO-FACE CONSULTATION WITH A DERMATOLOGIST
      e) DIRECT, PRIORITY REFERRAL FOR FACE-TO-FACE CONSULTATION WITH A DERMATOLOGIST
3. REASON FOR DIRECT REFERRAL FOR FACE-TO-FACE CONSULTATION WITH A DERMATOLOGIST
   a) Incomplete clinical history
   b) Poor-quality photograph, does not permit a diagnosis to be made
   c) Lesions require palpation
   d) Dermatoscopy required
   e) Skin biopsy required
   f) Requires in-person treatment (cryotherapy)
   g) Other (specify)
References


