Agreement between dermatological diagnoses made by live examination compared to analysis of digital images

Concordância entre diagnósticos dermatológicos obtidos por consulta presencial e por análise de imagens digitais

Jonas Ribas [1]   Maria da Graça Souza Cunha [2]
Antônio Pedro Mendes Schettini [3]   Carla Barros da Rocha Ribas [4]

Abstract: BACKGROUND: Teledermatology is seldom used in Brazil, although some incipient initiatives have been implemented in the state of Amazonas. Further studies are still required to confirm the feasibility and efficacy of this diagnostic tool.
Objectives: To evaluate the efficacy of an asynchronous method of teledermatology using simple, inexpensive technological resources.
Methods: One hundred and seventy-four patients were examined by four dermatologists, two clinic-based dermatologists, who diagnosed the patients following live examination (A1 and A2), and two consultant specialists in image dermatology, who reached diagnoses by examining images of the lesions and the patients' clinical histories (B1 and B2). The agreement between live examination and examination of images was evaluated.
Results: Agreement between the two live examiners (A1 and A2) with respect to the principal diagnosis was 83.3% compared to 81.0% between the physicians who diagnosed by examining images (B1 and B2). Agreement between the principal diagnosis established by live examination and that obtained from examining images ranged from 78.2% to 83.9%,
Conclusion: Agreement between diagnoses of dermatological diseases reached following examination of digital images and those reached following live examination of the patient was excellent.
Keywords: Clinical diagnosis; Dermatology; Telemedicine

Resumo: FUNDAMENTOS: A teledermatologia, pouco utilizada no Brasil e com iniciativas incipientes no Amazonas, necessita de estudos para demonstrar sua factibilidade e eficácia.
OBJETIVOS: Avaliar a eficácia de um método assíncrono de teledermatologia, utilizando recursos tecnológicos simples e de baixo custo.
MÉTODOS: Cento e setenta e quatro pacientes foram examinados por quatro dermatologistas; dois efetuaram diagnóstico presencial (A1 e A2) e dois por meio de imagens das lesões e história clínica (B1 e B2). Foi investigada a concordância entre as avaliações presenciais e por imagens.
RESULTADOS: A concordância do diagnóstico principal entre os examinadores presenciais (A1 e A2) foi de 83,3% e entre os de imagens (B1 e B2), de 81%. A concordância entre o diagnóstico principal estabelecido pelo método presencial e o obtido por meio de imagens variou de 78,2% a 83,9%.
CONCLUSÃO: O diagnóstico de doenças dermatológicas realizado por imagens digitais demonstrou concordância ótima quando comparado àquele realizado com a presença física do paciente.
Palavras-chave: Dermatologia; Diagnóstico clínico; Telemedicina

INTRODUCTION

The first experiences in the field of video, audio and image transmission for medical purposes dates back to the US and Russian space missions of the 1960s. The idea of using this technology to evaluate patients at a distance was first conceived in the 1970s due to the need to improve specialized medical care in rural areas of the United States, where medical specialists were scarce. This was the cornerstone of telemedicine. According to the World Health Organization, telemedicine is a means of providing service resources and extending healthcare to geographical locations where they are inexistent or where certain types of procedures are lacking, with the broad objective of permitting equal access to medical services in cases in which distance represents a critical factor. In addition to enabling medical care to be provided at a distance, telemedicine also enables information to be exchanged in research and education, using the same information and communication technology. [1,2]

Two methods of exchanging information were developed in telemedicine: real-time consultation or synchrony and the store-and-forward or asynchronous method. In the synchronous method, two or more individuals communicate in real time, enabling remote diagnosis to be made immediately. The patient, the physician requesting the consultation and the consultant physician all participate in the session simultaneously, generally via a video conference call. In the asynchronous method, text, image, video and audio files are transmitted to data storage devices in which they are kept for later consultation. This method is simpler to perform, since it does not require the simultaneous presence of the parties involved and uses simpler and less expensive material resources. [3]

The diagnostic efficacy of telemedicine has been confirmed in various studies, with agreement rates ranging from 55% to 99%. No statistical correlation has been found with age, gender, race or with the nature of the dermatosis. [4-14] Perednia and Brown were the first to use the resources of telemedicine in dermatology [15] and Zelickson and Homan studied the reproducibility of the method, comparing diagnoses reached by analyzing stored digital images with diagnoses reached by the conventional method (live examination). These authors reported an agreement rate of 88%. [16]

Another benefit of the implementation of teledermatology is the possibility of diagnosing dermatoses at an earlier stage. In a study conducted in a rural area of Cambodia, Brandling-Bennett showed a reduction in the interval between the onset of the symptoms of the disease and diagnosis, after 28 months of use of a teledermatology program. [17] The interval found between the onset of symptoms and diagnosis highlights the longer waiting time for consultation with a specialist when access is difficult and specialists are scarce. [18] Some authors have identified obstacles to increasing the use of teledermatology such as guaranteeing the security of the data, issues regarding the remuneration of the participants and the need to sensitize public health managers to the importance of this program. [10,19]

Amazonas is the largest state in Brazil, occupying an area that corresponds to one-fifth of the entire country. Tropical rainforest covers much of this state, which contains the largest hydrographic basin in the world, and access to healthcare services in the region is extremely difficult. Although river transportation exists, it is precarious due to poor infrastructure. Moreover, the few roads that exist in the region are blocked during the floods or are frequently reclaimed by the forest due to lack of maintenance. Therefore, townships far from the capital city become relatively isolated at certain times of the year. [20]

There is a high prevalence of infectious diseases such as malaria, leishmaniasis, hanseniasis, tuberculosis, hepatitis, dengue, mycoses and skin viruses in Amazonas. The birth rates and infant mortality rates are still high in the region, probably as a consequence of poverty, the difficulty in accessing healthcare services, and inadequate infrastructure and basic sanitation. [21]

Another major difficulty in healthcare service coverage derives from the disproportional distribution of physicians in the state, which results in difficulties in providing a wide, hierarchical network of dermatological care for the population. [22]

Since the estimated prevalence of dermatoses in the Brazilian population ranges from 26% to 37%, a considerable amount of human and material resources would be required to meet this demand. [23] Teledermatology represents a promising alternative that would improve the coverage of dermatological care, principally in areas where distance from the major centers represents a critical factor and in those cases in which there is a need for specialist opinion. Teledermatology is still uncommon in Brazil, and in the state of Amazonas initiatives remain incipient and infrequent. Further studies are required to confirm the feasibility and efficacy of this tool.

The objective of the present study was to evaluate the efficacy of an asynchronous method of teledermatology using simple and inexpensive technology.

MATERIAL AND METHODS

The protocol of this descriptive, cross sectional diagnostic test study was approved by the Internal

Review Board of the *Alfredo da Matta* Foundation (FUAM), Manaus, Amazonas, Brazil. The study was designed to investigate the agreement between the diagnosis of dermatological diseases reached by evaluating digital images (at a distance) and conventional diagnosis reached in the physical presence of the patient. The overall agreement rate between the two types of diagnosis, distance and live examination, was calculated, considering the principal diagnosis alone as well as this diagnosis in association with a differential diagnosis. The agreement rate was also calculated as a function of whether the diseases were classified as infectious inflammatory diseases, non-infectious inflammatory diseases or neoplasias.

Patients with dermatoses receiving care at the FUAM screening clinic, which receives patients spontaneously seeking care, were included in the study. Calculation of sample size was based on an expected agreement of 80%, significance level of 10% and precision of 10%, resulting in the inclusion of 174 randomly selected patients. Patients who refused to participate in the study, those cases in which the images of the dermatoses were considered to be of insufficient quality to enable evaluation to be made, patients who were already in treatment, and those who had only clinical symptoms with no visible dermatological signs were excluded from the study.

All the enrolled cases were examined by four dermatologists, two of whom carried out the examination and reached diagnosis in the presence of the patient (A1 and A2), while the other two were consultants who received digitalized images of the lesions that were sent by electronic mail together with the clinical history of the patients, and who completed a consultancy form following analysis (B1 and B2). Each patient was given a principal diagnosis and two differential diagnoses. Measures were taken to ensure that no communication occurred between the examiners and the consultants during the study with respect to the evaluation of cases. All the physicians participating in the study were dermatologists with specialist certificates and more than five years of practice.

Evaluator A1 took three photographs of each patient at three different focal distances using a Nikon D80 camera with a 60-mm f/2.8D AF Micro-Nikkor lens in aperture priority, 1 megapixel resolution, in artificial light (built-in flash). The most significant dermatological lesion for the purpose of diagnosis was photographed at the shortest possible focal distance and included the adjacent normal skin (close-up photo). A second, panoramic photograph included an anatomical structure that permitted the site of the lesion and the distribution of the lesions on the patient's body to be identified, but did not permit identification of the patient. A third photograph was taken at a distance halfway between the other two.

Agreement between the two examiners and between the two consultants (intra-evaluator) and between examiners and consultants (inter-evaluator) was assessed in the statistical analysis using overall agreement and the Kappa coefficient of agreement with the respective 95% confidence intervals (95%CI) in accordance with the classification defined by Fleiss: Kappa values > 0.7 are considered indicative of excellent agreement; values between 0.4 and 0.7 moderate agreement; and values < 0.4 poor agreement.

The variables regarding the hypotheses of the principal and differential diagnoses reached by each one of the evaluators (live and distant examiners) were used in the statistical analysis. A graduation scale (Table 1) adapted from that initially developed by Whited et al. (1999) was used to classify the degree of agreement between the diagnoses as complete, partial or discordant. [14]

RESULTS

In the study sample, 53.4% of the group was female and 46.6% male. Mean age was 34.7 ± 18.1 years (range 2-90 years). Patients' skin color was brown in the majority of cases (65.3%), followed by white (34.1%) and black (1%). With respect to the

TABLE 1: Scale of agreement

Complete agreement	Evaluates agreement based only on the principal diagnosis. When the principal diagnoses are identical, agreement is classified as complete.
Partial agreement	Evaluates agreement based on the principal or differential diagnosis. When the principal diagnosis made by an evaluator is identical to the differential diagnosis made by a second evaluator or when the differential diagnoses made by both evaluators are identical, this constitutes partial agreement.
Discordant	When neither the principal nor the differential diagnoses are identical, this is considered a lack of agreement.

Source: Whited, et al. [14]

TABLE 2: Most common clinical hypotheses

Diagnoses	Frequency	Percentage (%)
Dermatophytosis	170	12.2
Pitiriasis versicolor	112	8.0
Allergic contact dermatitis	88	6.3
Psoriasis	55	3.9
Pityriasis alba	47	3.4
Scabies	42	3.0
Herpes zoster	40	2.9
Warts	38	2.7
Vitiligo	33	2.4
Seborrheic dermatitis	29	2.1
Lichen simplex chronicus	28	2.1
Pityriasis rosea	28	2.1

duration of the lesions, most (52%) had been present for more than six months, followed by 1-6 months (19.1%), 1-4 weeks (16.8%) and 1-7 days (12.1%).

Of the 1,392 hypotheses made (8 hypotheses for each one of the 174 patients), the three most frequent, both in the live examinations and in those based on images, were dermatophytosis (12.2%), pityriasis versicolor (8.0%) and allergic contact dermatitis (6.3%). Table 2 shows details of the 12 most common clinical hypotheses.

The agreement found between the live examiners (A1 and A2) with respect to the principal diagnosis was 83.3%, whereas agreement between the consultants who reached diagnoses by examining images (B1 and B2) was 81% (Table 3).

When a differential diagnosis was added to the principal diagnosis, agreement between the live examiners (A1 and A2) was 94.3% compared to 96.6% between the consultants who diagnosed from images (B1 and B2) (Table 4).

With respect to agreement between the principal diagnosis established by live examination and that obtained by viewing images, the following results were found: between evaluators A1 and B1 agreement rate was 80.5%; between A1 and B2 agreement was 83.9%; between A2 and B1, 78.2% and between A2 and B2, 83.3%. Table 5 shows the percentages of agreement between the principal diagnoses.

When a differential diagnosis was added to the principal diagnosis, the following agreement rates were found: between A1 and B1, 94.3%; between A2 and B2, 96%; between A2 and B1, 92%; and between A1 and B2, 98.3%, this being the highest agreement rate found in the study (Table 6).

Table 7 shows the agreement (Kappa) between the principal diagnoses reached by the two live examiners (A1 and A2) and the two distance consultants (B1 and B2) when the skin conditions were classified into infectious inflammatory diseases, noninfectious inflammatory diseases or neoplasias.

DISCUSSION

Although the study was not designed to permit findings to be extrapolated to the rest of the population, the demographic characteristics of the study sample are similar to those reported in other studies carried out in the same region. The majority of the participants were brown, which is the predominant skin color in the region. The mean age of the patients in the study and the predominance of females are similar to the characteristics of the population seen in the clinic during the same period in which the study was conducted. [24,25] With respect to the duration of the disease prior to diagnosis, the most common finding was that the condition had been present for periods greater than six months, which is comparable with results published in the literature. [26] In addition, the prevalence of the various dermatoses was not significantly different from results found in other studies or in the area where the study was conducted. [25,27]

The efficacy of diagnosis in asynchronous telemedicine systems has already been referred to in various reports and was confirmed in the present study. The agreement rates found between the live examiners and the consultants who diagnosed by examining images is within a range of values considered excellent. When a differential diagnosis was added to the analysis, the overall agreement rate was even higher.

Other similar studies conducted to evaluate the variability of interobserver agreement showed very similar results: Heffner *et al.* 82%, Krupinski *et al.* 84%, Lim *et al* 95%, Whited *et al.* 92% and Pak *et al.* 91%. [5,7,9,14,28]

With respect to agreement between examiners and consultants (live examination versus examination of images), which was the principal objective of the

TABLE 3: Overall agreement on the principal diagnosis hypotheses between the live examiners and between the consultants who reviewed digitalized images

Evaluators	Modality of Diagnosis	Overall Agreement Rate (%)	95% Confidence Interval (%)
A1 x A2	Live examination x Live examination	83,3	76,9 - 88,5
B1 x B2	Images x Images	81	74,4 - 86,6

TABLE 4: Overall agreement on the principal diagnosis hypotheses between the live evaluators and the consultants who reviewed digitalized images

Evaluators	Modality of Diagnosis	Overall Agreement Rate(%)	95% Confidence Interval(%)
A1 x A2	Live examination x Live examination	94,3	89,7 - 97,2
B1 x B2	Images x Images	96,6	92,6 - 98,7

TABELA 5: Overall agreement on the principal diagnosis hypotheses between the live evaluators and the consultants who reviewed digitalized images

Evaluators	Modality of Diagnosis	Overall Agreement Rate(%)	95% Confidence Interval(%)
A1 x B1	Live examination x Images	80,5	73,8 - 86,1
A1 x B2	Live examination x Images	83,9	77,6 - 89,0
A2 x B1	Live examination x Images	78,2	71,3 - 84,1
A2 x B2	Live examination x Images	83,3	76,9 - 88,5

TABLE 6: Overall agreement on the hypotheses of the principal and differential diagnoses between the live evaluators and the consultants who reviewed digitalized images

Evaluators	Modality of Diagnosis	Overall Agreement Rate(%)	95% Confidence Interval(%)
A1 x B1	Live examination x Images	94,3	89,7 - 97,2
A1 x B2	Live examination x Images	98,3	95,0 - 99,6
A2 x B1	Live examination x Images	92	86,9 - 95,5
A2 x B2	Live examination x Images	96	91,9 - 98,4

TABLE 7: Agreement between evaluators with respect to principal diagnoses, when cases were classified as infectious inflammatory diseases, noninfectious inflammatory diseases or neoplasias

Evaluators	Agreement Rate %	Kappa Index
A2 x A1	95,4	0,92 - Very good
B2 x B1	91,4	0,85 - Very good
B1 x A1	90,2	0,80 - Very good
B2 x A1	90,2	0,80 - Good
B1 x A2	92,5	0,87 - Very good
B2 x A2	92	0,86 - Very good

present study, the values found were considered excellent and agreement increased even further when a differential diagnosis was added to the principal diagnosis. This finding is also in agreement with results published by other authors, who evaluated the agreement rate between the two modalities of diagnosis and found the following rates of agreement for the principal diagnosis and differential diagnoses: 54% and 80% (Gilmour *et al*.)[3] and 81% and 89% (High *et al*.)[6], respectively. This reinforces the need to make at least one differential diagnosis in addition to the principal diagnosis during telemedicine consultations.

Analysis of the Kappa index of agreement between diagnoses when conditions were classified as infectious inflammatory diseases, noninfectious inflammatory diseases or neoplasias showed agreement rates that were considered very good. This highlights the versatility of the method, since it means that telemedicine consultation can be used to evaluate various groups of dermatoses with no differences in accuracy.

CONCLUSION

Agreement between diagnoses of dermatological diseases performed at a distance by analyzing digital images compared to live diagnosis made in the physical presence of the patient was very good.

When a differential diagnosis was added to the

principal diagnosis, the agreement rate increased further, both between the live examiners and between consultants at a distance.

When the diseases were grouped into categories of infectious inflammatory diseases, noninfectious inflammatory diseases and neoplasias, the Kappa index of agreement was still classified as very good between the observers in both forms of diagnosis.

In view of the excellent agreement rates found between live diagnosis and diagnosis using images, the use of the asynchronous method of teledermatology may be recommended. Implementation of teledermatology programs in the Brazilian state of Amazonas may contribute towards improving the efficacy and coverage of dermatological care for the population.   ❑

REFERENCES

1. Chen K, Lim A, Shumack S. Teledermatology: influence of zoning and education on a clinician's abilities observe peripheral lesions. Australas J Dermatol. 2002;43:171-4.
2. who.int/en [homepage]. World Health Organization. Telemedicine [cited 2005 Jan 18]. Avaiable from: http://whqlibdoc who.int/hq/1998/WHO_DGO_98.1.pdf
3. Gilmour E, Campbell SM, Loane MA, Esmail A, Griffiths CE, Roland MO, et al. Comparison of teleconsultations and face-to-face consultations: preliminary results of a United Kingdom multicentre teledermatology study. Br J Dermatol. 1988;139:81-7.
4. Edison KE, Wards DS, Dyer JA, Lane W, Chance L, Hicks LL. Diagnosis, diagnostic confidence, and management concordance in live-interactive and store-and-forward teledermatology compared to in-person examination. Telemed J E Health. 2008;14:889-95.
5. Heffner VA, Lyon VB, Brousseau DC, Holland KE, Yen K. Store-and-forward teledermatology versus in-person visits: A comparison in pediatrio teledermatology clinic. J Am Acad Dermatol. 2009;60:956-61.
6. High WA, Houston MS, Calobrisi SD, Drage LA, McEvoy MT. Assessment of the accuracy of low-cost store-and-forward teledermatology consultation. J Am Acad Dermatol. 2000;42:776-883.
7. Krupinski EA, LeSueur B, Ellsworth L, Levine N, Hansen R, Silvis N, et al. Diagnostic accuracy and image quality using a digital camera for teledermatology. Telemed J E Health. 1999;5:257-63.
8. Kvedar JC, Edwards RA, Menn ER, Mofid M, Gonzalez E, Dover J, et al. The substitution of digital images for dermatologic physical examination. Arch Dermatol. 1997;133:161-7.
9. Lim AC, Egerton IB, Shumack SP. Accuracy and reliability of store-and-forward teledermatology: preliminary results from the St George Teledermatology Project. Australas J Dermatol. 2001;42:247-51.
10. Lowitt MH, Kessler II, Kauffman CL, Hooper FJ, Siegel E, Burnet JW. Teledermatology and in-person examinations: a comparison of patient and physician perception and diagnostic agreement. Arch Dermatol. 1998;134:471-6.
11. Pak HS. Teledermatology and teledermatopathology. Semin Cutan Med Surg. 2002;21:179-89.

12. Scheinfeld N, Kurz J, Teplitz E. A comparison of the concordance of digital images, live examinations, and skin biopsies for the diagnosis of hospitalized dermatology consultation patients. Skinmed. 2003;2:14-9.
13. Scheinfeld N. The use of teledermatology to supervise dermatology residents. J Am Acad Dermatol. 2005;52:378-80.
14. Whited JD, Hall RP, Simel DL, Foy ME, Stechuchak KM, Drugge RJ, et al. Reliability and accuracy of dermatologists' clinic-based and digital image consultations. J Am Acad Dermatol. 1999;41:693-702.
15. Perednia DA, Brown NA. Teledermatology: one application of telemedicine. Bull Med Libr Assoc. 1995;83:42-7.
16. Zelickson BD, Homan L. Teledermatology in the nursing home. Arch Dermatol.1997;133:171-4.
17. Brandling-Bennett HA, Kedar I, Pallin DJ, Jacques G, Gumley GJ, Kvedar JC. Delivering health care in rural Cambodia via store-and-forward telemedicine: a pilot study. Telemed J E Health. 2005;11:56-62.
18. D´Elia PB, Fisher PD, Bordin R, Harzheim E, Ramos MC. Concordância entre diagnósticos dermatológicos feitos presencialmente e por imagens digitais. An Bras Dermatol. 2007;82:521-7.
19. Vidmar DA, Cruess D, Hsieh P, Dolecek Q, Pak H, Gwynn M, et al. The effect of decreasing digital image resolution on teledermatology diagnosis. Telemed J E Health. 1999;5:375-83.
20. Moreo A. Um raio X do Amazonas. Conasems. 2004;8:37-45.
21. datasus.gov.br [homepage]. Ministério da Saúde. DataSus [acesso 15 Maio 2006]. Disponível em: http://tabnet.datasus.gov.br/cgi/ tabcgi. exe?idb2005 /d0114.def.
22. Miranda AG, Pennini S, Pinheiro AG. O cenário geossociopolítico da região norte. In: O Médico e o seu trabalho: resultados da região norte e seus estados. Brasília: Conselho Federal de Medicina; 2005. p.33- 40.
23. Williams HC. Epidemiology of skin disease. In: Burns T, Breathnach S, Cox N, Griffths C. Rook's Textbook of Dermatology. 7th ed. Oxford: Blackwell Science; 2004. p.6.1-6.21.
24. ibge.gov.br [home page]. Censos Demográficos, 2000 [acesso 10 Mar 2006]. Disponível em:

http://www.sidra.ibge.gov.br/bda/tabela/protabl.asp.

25. Fundação Alfredo da Mata. Boletim Epidemiológico. (Folheto de circulação interna). Manaus: Fundação Alfredo da Mata; jan/dez. 2005.
26. Weinstock NA, Bingham SF, Cole GW, Eilers D, Naylor MF, Kalivas J, et al. Reliability of counting actinic keratoses before and after brief consensus discussion: the VA topical tretinoin chemoprevention (VATTC) trial. Arch Dermatol. 2001;137:1055-8.
27. Krupinski E, Barker G, Rodriguez G, Engstrom M, Levine N, Lopez AM, et al. Telemedicine versus in-person dermatology referrals: an analysis of case complexity. Telemed J E Health. 2002;8:143-7.
28. Pak HS, Harden D, Cruess D, Welch ML, Poropatich R. National Capital Area Teledermatology Consortium. Teledermatology: an intraobserver diagnostic correlation study, part I. Cutis. 2003;71:399-403.

MAILING ADDRESS / ENDEREÇO PARA CORRESPONDÊNCIA:
Jonas Ribas
Rua 24 de maio, 220 - Centro
69010 080 - Manaus - AM, Brazil
Phone/fax: +55 92 3234 5019
E-mail: ribas@internext.com.br