Review article

A systematic review of the efficacy of telemedicine for making diagnostic and management decisions

William Hersh*†, Mark Helfand*†, James Wallace*†, Dale Kraemer*†, Patricia Patterson*†‡, Susan Shapiro†‡ and Merwyn Greenlick†§

*Division of Medical Informatics and Outcomes Research, Oregon Health and Science University; †Evidence-Based Practice Center, Oregon Health and Science University; ‡School of Nursing, Oregon Health and Science University; §Department of Public Health and Preventive Medicine, Oregon Health and Science University, Portland, Oregon, USA

Summary

We conducted a systematic review of the literature to evaluate the efficacy of telemedicine for making diagnostic and management decisions in three classes of application: office/hospital-based, store-and-forward, and home-based telemedicine. We searched the MEDLINE, EMBASE, CINAHL and HealthSTAR databases and printed resources, and interviewed investigators in the field. We excluded studies where the service did not historically require face-to-face encounters (e.g. radiology or pathology diagnosis). A total of 58 articles met the inclusion criteria. The articles were summarized and graded for the quality and direction of the evidence. There were very few high-quality studies. The strongest evidence for the efficacy of telemedicine for diagnostic and management decisions came from the specialties of psychiatry and dermatology. There was also reasonable evidence that general medical history and physical examinations performed via telemedicine had relatively good sensitivity and specificity. Other specialties in which some evidence for efficacy existed were cardiology and certain areas of ophthalmology. Despite the widespread use of telemedicine in most major medical specialties, there is strong evidence in only a few of them that the diagnostic and management decisions provided by telemedicine are comparable to face-to-face care.

Introduction

Telemedicine can be defined as the use of telecommunication for diagnostic, monitoring and therapeutic purposes when distance separates the participants[1]. It has been advocated for improving the provision of health-care to individuals in rural areas, in the home and in other places where medical personnel are not readily available. There are over 450 telemedicine programmes worldwide, and over 360 in the United States[2]. There are programmes in almost every medical specialty, and the populations commonly served include people who live in rural areas, who are elderly, who reside in remote military settings, or who are veterans.

Despite this telemedicine activity, however, research to assess its efficacy has not been performed in all the clinical specialties in which it has been implemented[2]. Previous systematic reviews of telemedicine have examined clinical outcomes[3–6], patient satisfaction[7] and cost[8]. However, there appears to have been no systematic review of the efficacy of telemedicine for diagnostic or management decisions.

The present study had its origins in a report commissioned by the US Health Care Financing Agency (HCFA) and the Agency for Healthcare Research and Quality (AHRQ) to assess the efficacy of telemedicine in the adult, non-pregnant population[2]. A subsequent supplementary report extended the analysis to the paediatric and obstetric population[9]. This paper focuses on one aspect of those reports, namely the efficacy of telemedicine for diagnostic and management decisions.

Accepted 18 February 2002
Correspondence: Dr William Hersh, Division of Medical Informatics and Outcomes Research, Oregon Health and Science University, BICC, 3181 SW Sam Jackson Park Rd, Portland, OR 97201, USA (*Fax: +1 503 494 4551; Email: hersh@ohsu.edu*)

Since the usual goal of telemedicine is to provide equivalent health-care via telecommunication to locations where such care is not readily available, the key question is whether it can provide at least equally good diagnostic and management decisions to face-to-face care. This would then qualify the technology for evaluation at the next levels, of clinical outcomes and cost-effectiveness.

Methods

We divided the studies into three general areas. The first two relate to office/hospital-based telemedicine, in which both patient and clinician are in a health-care setting, such as a practitioner's office or a hospital. The two categories of office/hospital-based telemedicine are *interactive*, where the interaction between the parties occurs in realtime, and *store-and-forward*, where clinical data are collected and transmitted for review later. In the third general area, *home-based telemedicine*, the patient is at his or her residence and there is direct communication between the clinician and the patient or carer. (In principle, there are two categories of home-based telemedicine, but in practice there have been no reports to date of store-and-forward home telemedicine.)

The criteria for inclusion in this systematic review were:

(1) that the study was relevant to at least one of the above areas of telemedicine;
(2) that the report compared data on the diagnostic capability or management decision performance for telemedicine with those for a control condition (i.e. in-person care).

We excluded services:

(1) that would not normally require face-to-face encounters between the clinician and patient (e.g. radiology and pathology diagnosis);
(2) that used only telephone care, traditional photography or email;
(3) that provided medical advice directly to the public.

We developed a search strategy designed to find any publications about telemedicine and used it to search the MEDLINE, EMBASE, CINAHL and HealthSTAR databases through to February 2001 (see Appendix). In addition, we searched telemedicine reports and compilations such as those from the Conseil d'Evaluation des Technologies de la Sante du Quebec[10], the Telemedicine Strategic Healthcare Group[11] and the Association of Telemedicine Service Providers' database[12], and the *Telemedicine Sourcebook 1998*[13]. We also assessed three systematic reviews (different in scope from this study) from the International Network of Agencies for Health Technology (INHATA)[4], the Cochrane Database of Systematic Reviews[6], and the Catalan Agency for Health Technology Assessment and Research[5]. We identified additional articles from the reference lists of included reports and articles, and contacted known telemedicine experts to find more still. Finally, we hand-searched all issues of the two peer-reviewed telemedicine journals, *Journal of Telemedicine and Telecare* and *Telemedicine Journal* (recently renamed *Telemedicine Journal and e-Health*), through to the end of 2000.

Each abstract retrieved was read by two reviewers. If they disagreed over whether the full article deserved examination, a third reviewer made the final decision. We then retrieved the full-text articles. When an article met the criteria for inclusion, the summary statistics (e.g. agreement with a gold standard or agreement with other clinicians) were extracted. To evaluate the quality of the studies, we adapted a scale used to assess the quality of diagnostic test evaluations, previously developed by our evidence-based practice centre (Table 1)[14]. We used a second scale to indicate the direction of the evidence (Table 2).

Results

The literature search of the computerized databases yielded 4628 possibly relevant references. An additional 81 studies were identified by hand-searching, review of references of other reviews and other means. After application of the inclusion criteria, there were 33

Table 1 Classification of evidence for diagnostic studies

Study class	Characteristic
I	Case series of consecutive patients from a relevant population of individuals who would use telemedicine. Use of an objective gold standard with blinded interpretation of results. Inter-observer analysis
II	Case series of patients from a relevant population of individuals who would use telemedicine. Use of an objective gold standard
III	Case series not from a relevant population or not using appropriate methodology for diagnostic test evaluation

Table 2 Direction of evidence for effect

Study class	Characteristic
A	Strong positive effect
B	Weak positive effect
C	Conflicting evidence for effect
D	Negative effect (i.e. evidence that the technology is inferior or ineffective)

articles in interactive office/hospital-based telemedicine (Table 3), 22 in store-and-forward office/hospital-based telemedicine (Table 4) and four in home-based telemedicine (Table 5).

The overall methodological quality of most studies was low. Only four studies of interactive office/hospital-based telemedicine, two studies of store-and-forward office/hospital-based telemedicine and two studies of home-based telemedicine met the criteria for class I studies. Typical problems included small sample sizes, the use of the same clinician for telemedicine and in-person interaction with the patient, and lack of measurement of inter-observer agreement within modalities. For this reason, as well as the general heterogeneous nature of the studies, we did not attempt to aggregate them (i.e. in a meta-analysis). Because there were so few high-quality studies, we did not exclude any from the analysis on the grounds of methodology, as long as the study had a control group comparing the accuracy or agreement of telemedicine with face-to-face care.

Office/hospital-based telemedicine — interactive

The most common specialties in interactive office/hospital-based telemedicine were psychiatry (seven studies), dermatology (seven studies), cardiology (five studies) and emergency medicine (three studies).

Psychiatry

Four of the psychiatry studies assessed various forms of psychometric testing, whereas the other three assessed diagnosis by psychiatrists. Two of the studies assessing diagnosis by psychiatrists were of high quality and showed high kappa values (>0.70) for inter-observer agreement for various psychiatric diagnoses[15,16]; the other study found 96% agreement among telemedicine and in-person psychiatrists in diagnosis and management plans in child psychiatry[17]. All but one of the psychometric test studies had very small sample sizes (4–11 patients)[18–20]. The largest of them found that neuropsychiatry tests administered in-person and via telemedicine yielded very similar results[21].

Dermatology

Seven teledermatology studies that assessed diagnostic and/or management decisions used interactive videoconferencing. In the study of the best methodological quality[22], diagnostic agreement (with an in-person gold standard) for interactive teledermatology was significantly lower than that achieved in an in-person consultation. This was in contrast to the store-and-forward study with the best methodology (see below)[23], which showed comparable agreement for teledermatology and in-person assessment (again, both rates of agreement being with an in-person gold standard). The remaining studies had serious methodological limitations, such as using the same dermatologist for the teledermatology and in-person consultation, assessing the agreement for detecting malignancies with insufficient sample size to estimate the false-negative rate of the teledermatology examination, and not employing a gold standard (e.g. biopsy) to determine whether cancer was present[24–29].

Cardiology

Three of the cardiology studies focused on children and two on adults. Studies in paediatric cardiology demonstrated reasonable agreement for auscultation[30,31] and examination via a telemedicine link[32]. In an adult cardiology study, an attending physician overcame errors in interpretation of electrocardiograms and improved management decisions by senior house officers[33]. Another adult study found a high rate of agreement between in-person and remote assessment of dobutamine stress echocardiography tests, although there was a very high rate of normal tests, which would make it difficult to detect a difference[34].

Emergency medicine

One study of emergency medicine assessed the value of teleradiology consultation in assisting emergency room (ER) clinicians[35], and it showed that the teleconsultation changed the ER diagnosis in 30% of cases and the initial treatment plan in 26%. Another compared physical examination findings by both an in-person ER physician (gold standard) and a paediatric critical care specialist via telemedicine; the sensitivity and specificity of the telemedicine examinations in detecting abnormal findings were 87.5% and 93.0%, respectively[36]. A final study showed improvement in the sensitivity and specificity of diagnoses made by nurse practitioners in a minor accident and trauma service when aided by telemedicine[37].

Table 3 Studies of efficacy for diagnostic and management decisions in interactive office/hospital-based telemedicine

Study (first author)	Quality–evidence score	Clinical specialty	Sample	Intervention	Effects — diagnosis	Effects — management
Casey[32]	II–B	Cardiology	9 patients	Agreement between examination guided by paediatric cardiologist vs in-person paediatric cardiologist	88% of diagnoses correct	Not assessed
Belmont[30]	I–B	Cardiology	116 patients	Diagnostic agreement between paediatric cardiologists in acoustic, electronic and tele-electronic stethoscopy	Kappa for agreement between acoustic and electronic was 0.65, and between acoustic and tele-electronic was 0.64	Sensitivity and specificity of need for follow-up consultation or echocardiology were 88% and 97%, respectively
McConnell[31]	II–B	Cardiology	21 patients	Diagnostic agreement in in-person and telemedicine evaluation of paediatric heart murmurs	Agreement in 90% of cases	Not assessed
Srikanthan[33]	II–B	Cardiology	112 patients	Fax of electrocardiogram to cardiologist by resident to make diagnosis of myocardial infarction and to decide on use of thrombolytics	Electrocardiogram misinterpretation in 9% of cases	Consultant agreed with decision to use thrombolytics in 87.5% of cases
Trippi[34]	III–C	Cardiology	26 patients	Agreement in remote and in-person diagnosis from dobutamine stress echocardiograms in emergency room for acute myocardial infarction	Agreement in 96% of patients but 88% of tests were normal	Not assessed
Baur[38]	III–B	Dentistry	13 patients	Agreement in interactive video and in-person orthognathic examination	Agreement on examination classifications varied from 76.9% to 100%	Not assessed
Lowitt[27]	III–C	Dermatology	139 patients with 318 diagnoses	Diagnostic agreement in interactive video vs in-person examination	For paired observations, agreement in 80%; for 11 patients with biopsies done, interpretation confirmed biopsy in seven, favoured video diagnosis in one and revealed different diagnosis in three	Not assessed
Loane[24]	III–C	Dermatology	351 patients with 427 diagnoses	Diagnostic and management plan agreement in interactive video vs in-person examination	67% of diagnoses in agreement	64% had same management plan
Lesher[22]	II–C	Dermatology	60 patients with 115 problems	Diagnostic agreement in interactive video vs in-person examination	78% agreement between telemedicine and local diagnosis vs 94% agreement between local and local diagnosis	Not assessed
Oakley[25]	III–C	Dermatology	104 patients with 135 conditions	Diagnostic agreement in interactive video vs in-person examination	75% of conditions correctly diagnosed	Not assessed
Phillips[28]	II–C	Dermatology	107 lesions in 51 patients	Agreement on diagnostic and management plan in interactive video vs in-person examination for skin tumours	59% agreement on diagnosis (kappa=0.32)	86% agreement on recommending biopsy (kappa=0.47)
Gilmour[26]	III–C	Dermatology	126 patients	Diagnostic agreement in interactive video vs in-person examination for skin tumours	57% agreement on diagnosis	Not assessed

Continued

Table 3 *Continued*

Study (first author)	Quality–evidence score	Clinical specialty	Sample	Intervention	Effects—diagnosis	Effects—management
Phillips[29]	III–C	Dermatology	79 diagnoses in 60 patients	Diagnostic agreement in interactive video vs in-person examination	77.2% agreement on diagnosis	Not assessed
Lee[35]	II–B	Emergency medicine	123 studies with 460 radiographs for 90 patients	Effect of realtime teleradiology on diagnosis and treatment plans in emergency room	Changed initial diagnosis in 30% of cases	Changed initial treatment plan in 26% of cases
Kofos[36]	II–B	Emergency medicine	15 patients	Agreement between in-person emergency room physician (gold standard) and paediatric critical care specialist via telemedicine for physical examination findings	Sensitivity of abnormal findings was 87.5% and specificity was 93.0%	Not assessed
Tachakra[37]	II–B	Emergency medicine	150 patients	Improvement in diagnosis via telemedicine consultation	Telemedicine improved sensitivity and specificity of working diagnosis from 90% to 97% and 96% to 99%, respectively	Not assessed
Hubble[39]	II–B	Neurology	9 patients	Assessment of Parkinson's disease by interactive videoconferencing vs in-person	High correlation for subjective (0.97), objective (0.91) and staging (0.88) scores	Not assessed
Craig[40]	I–B	Neurology	23 patients	Agreement of junior physician performing examination via telemedicine with in-person examination by consultant neurologist	Kappa varied across findings from 0.42 to 1.0	Not assessed
Marcus[41]	II–C	Ophthalmology	34 eyes of 17 HIV-positive patients, 39 eyes of 20 patients with diabetes mellitus	Interactive video diagnosis of eye disease in HIV and diabetes mellitus	For HIV retinopathy, sensitivity was 83% and specificity was 95%; for diabetes, unable to classify 46% of glaucoma and 36% of retinopathy due to cataract, and poor sensitivity (29–50%) for those who could be assessed	Not assessed
Nitzkin[42]	III–B	Ophthalmology, physical therapy	58 patients with 1826 findings	Interactive video-ophthalmology and physical therapy examination findings by physicians and physical therapists	Ophthalmology agreement was 85.9% (kappa=0.65), physical therapy agreement was 87.7% (kappa=0.61)	Not assessed
Pedersen[43]	III–B	Otolaryngology	17 patients	General practitioners were telementored using otolaryngolical endoscopy (phase 2)	Agreement on diagnosis in all cases	Not assessed
Sclafani[44]	II–B	Otolaryngology	45 patients	Agreement of otolaryngologists in remote non-interactive video vs store-and-forward examination	Diagnostic agreement 85% for local vs non-interactive examination and 64% for local vs store-and-forward examination	Not assessed
Elford[17]	II–B	Psychiatry	23 patients	Diagnostic and management plan agreement as determined by an independent psychiatrist	Diagnostic agreement was 96%	Agreement with management plan was 96%

Continued

Table 3 *Continued*

Study (first author)	Quality–evidence score	Clinical specialty	Sample	Intervention	Effects—diagnosis	Effects—management
Montani[20]	II–B	Psychiatry	9 patients	Reliability of MMSE and CFT diagnostic rating scales administered by interactive video to elderly patients	A small but significant decrease in scores for remote patients	Not assessed
Ball[18]	II–B	Psychiatry	11 patients	Agreement of MMSE examination given remotely and in person	Pearson correlation coefficient was 0.89 between in-person and telemedicine examination	Not assessed
Ball[19]	II–B	Psychiatry	4 patients	Agreement of BPRS diagnostic rating scale given remotely and in person	Pearson correlation coefficient was 0.82 between in-person and telemedicine administration	Not assessed
Baigent[15]	I–B	Psychiatry	63 patients	Agreement of diagnosis with observer in same room or over video-link	Kappa was 0.85 when observer was in the same room, 0.70 when over the video-link	Not assessed
Ruskin[16]	I–B	Psychiatry	30 patients	Agreement on psychiatric diagnoses; half the patients had two in-person interviews and half had one in-person and one remote interview	Kappa was nearly identical for in-person vs in-person and remote interviews, varying from 0.70 to 1.0	Not assessed
Kirkwood[21]	II–B	Psychiatry	26 patients	Agreement in results of neuropsychiatry tests administered in person and via telemedicine	All test results within 95% confidence interval of difference	Not assessed
Pacht[45]	II–B	Pulmonary	40 patients	History and physical examination via interactive videoconferencing by pulmonary specialists	Kappa for agreement varied from 0.29 (history) to 0.66 (physical examination)	Not assessed
Graham[47]	II–D	Rheumatology	20 patients	Agreement between diagnosis by a junior physician presenting over video-phone and in-person diagnosis	Only 40% of diagnoses made correctly via video-phone—little better than 35% rate using only the telephone	Not assessed
Tachakra[48]	II–B	Trauma management	200 patients	Accuracy of telemedicine vs in-person examination for minor injuries	Sensitivity for various findings was 94.1–97.5%; specificity was 81.3% to 98.5%	Not assessed
Hayes[46]	III–C	Urology	14 real and 18 simulated patients	How often real and simulated teleconsultations for urolithiasis changed plan of care	Not assessed	Telemedicine altered treatment recommendations in 50% of real and 17% of simulated patients

MMSE, Mini-Mental State Examination; CFT, Complex Figure Test; BPRS, Brief Psychiatric Rating Scale.

Table 4 Studies of efficacy for diagnostic and management decisions in store-and-forward office/hospital-based telemedicine

Study (first author)	Quality–evidence score	Clinical specialty	Sample	Intervention	Effects — diagnosis	Effects — management
Houston[64]	II–C	Ambulatory care	20 patients	Comparison of diagnosis and treatment plans in store-and-forward telemedicine vs in person	Complete agreement on all diagnoses	93% of in-person consultations had additional treatment recommendations; 79% of store-and-forward consultations had additional treatment recommendations
Patterson[65]	II–C	Dentistry	27 patients	Diagnostic and management agreement among dentists in oral screening	Kappa for agreement on decay was 0.50–0.58	Kappa for agreement on extraction and filling was 0.93–1.0
Zelickson[50]	II–B	Dermatology	30 conditions in 29 patients	Diagnostic and management plan agreement in store-and-forward vs in-person examination	67% correct diagnosis with history alone, 85% with image alone and 88% with both	70% correct treatment plan with history alone, 87% with image alone and 90% with both
Whited[53]	III–C	Dermatology	13 lesions in 12 patients	Diagnostic agreement of store-and-forward vs in-person examination for skin tumours	74% correct diagnosis for in-person examiner vs 58% for store-and-forward examiner	Not assessed
Krupinski[23]	I–B	Dermatology	308 patients	Diagnostic agreement in store-and-forward vs in-person examination	83% concordance between in-person and teledermatology examination, 84% intra-dermatologist concordance, 81% inter-dermatologist concordance	Not assessed
Whited[49]	I–B	Dermatology	168 lesions in 129 patients	Diagnostic and management plan agreement in store-and-forward vs in-person examination	54% agreement among clinic examiners vs 41–55% between clinic and teledermatology examiners and 49–55% between teledermatology examiners	77% agreement among clinic examiners vs 56–77% between clinic and teledermatology examiners and 64–83% between teledermatology examiners
Kvedar[51]	II–B	Dermatology	121 images from 116 patients	Diagnostic agreement in store-and-forward vs in-person examination	61–64% agreement between remote and in-person examiners	Not assessed
Harrison[54]	III–B	Dermatology	210 patients	Diagnostic agreement for dermatologists vs general practitioners for skin tumours based on histological gold standard	Diagnosis for all tumours and malignant tumours was greater for teledermatology (71%, 94%) than for general practitioners (49%, 70%)	Not assessed
Braun[55]	II–B	Dermatology	55 lesions in 51 patients	Diagnostic agreement of dermatologists for pigmented skin lesions using teledermoscopy based on histological gold standard	Accuracy of diagnosis was 64% for dermoscopy and 75% for teledermoscopy	Not assessed
High[56]	II–B	Dermatology	106 lesions in 92 patients	Diagnostic agreement of dermatologists compared with in-person gold standard	Agreement of 82–98% for images of adequate quality, 60–75% for images of poor quality	Not assessed
Lewis[57]	II–B	Dermatology	141 patients	Diagnostic agreement of dermatologists assessing skin tumours in person and via telemedicine	Sensitivity was 88% and specificity was 80%	Not assessed

Continued

Table 4 *Continued*

Study (first author)	Quality–evidence score	Clinical specialty	Sample	Intervention	Effects — diagnosis	Effects — management
Lyon[58]	III–B	Dermatology	100 patients	Agreement between in-person and remote examination with trainee and consultant viewing photographs	For 38 patients with rashes, disagreement in 10%. For 62 patients with tumours, 4% disagreement	Not assessed
Taylor[52]	II–B	Dermatology	190 patients	Agreement between in-person and telemedicine dermatologists 13 months later	Agreement was 77%	31% of cases could have been managed by consultation with general practitioner
Yamamoto[66]	III–B	Neonatology	31 patients	Accuracy of chest radiograph interpretation between neonatologists using telemedicine and general paediatrician, and agreement of level of care assignment between neonatologists and general paediatrician	Neonatologists identified 91–98% of findings vs 82% for general paediatrician; neonatologists assigned level of care 73% of time with written chest radiograph report and 67% of time with image	Not assessed
Yogesan[60]	II–D	Ophthalmology	49 eyes in 25 patients	Agreement of digital camera with standard fundus camera gold standard for patients screened for diabetic retinopathy	Kappa for agreement on images was <0.3	Not assessed
Liesenfeld[59]	II–B	Ophthalmology	129 patients	Agreement of digital photographs with 35 mm photograph gold standard for patients screened for diabetic retinopathy	Sensitivity was 85% and specificity was 90%	Not assessed
Yogesan[61]	II–C	Ophthalmology	51 eyes in 27 patients	Diagnostic agreement among ophthalmologists using portable digital retinal camera for glaucoma	Sensitivity of portable camera was 0.68–0.79 and specificity was 0.67–0.87	Not assessed
Li[62]	II–B	Ophthalmology	26 eyes in 32 patients	Agreement of optometrists in diagnosing glaucoma in digital vs 35 mm images	Cup-to-disc ratio was in agreement for 80.8% of eyes	Not assessed
Schwartz[63]	III–B	Ophthalmology	19 eyes in 10 patients	Diagnostic and management agreement for patients who had retinopathy of prematurity	Agreement was 89–95% for diagnosis	Agreement on management plans in 42% of cases
Sclafani[44]	II–B	Otolaryngology	45 patients	Agreement of otolaryngologists in remote non-interactive video vs store-and-forward examination	Diagnostic agreement 85% for local vs non-interactive examination and 64% for local vs store-and-forward examination	Not assessed
Furukawa[67]	III–B	Otolaryngology	29 patients	Agreement of remote and local otolaryngologists in laryngoscopy	Three otolaryngologists made all identical diagnoses with still images and text of history	Not assessed
Wirthlin[68]	II–B	Wound care	38 wounds in 24 patients	Agreement in diagnosis and management of photographed vs directly observed wound healing	For wound diagnosis items, sensitivity was 0.86 and specificity was 0.72	For wound management, sensitivity was 0.87 and specificity was 0.78

Table 5 Studies of efficacy for diagnostic and management decisions in home-based telemedicine

Study (first author)	Quality–evidence score	Clinical specialty	Sample	Intervention	Effects — diagnosis	Effects — management
de Lusignan[72]	II–C	Home monitoring	20 patients	Accuracy of measurement of heart rate, respiratory rate, temperature and blood pressure vs standard portable unit	Within confidence interval for heart rate, outside and lower for respiratory rate, and outside and higher for temperature and systolic and diastolic blood pressure	Not assessed
Finkelstein[69]	I–A	Pulmonary	18 patients	Comparison between supervised and unsupervised home spirometry	Differences averaged $<3\%$	Not assessed
Bruderman[71]	II–C	Pulmonary	39 patients	Telespirometry to detect early signs of asthma deterioration	Not assessed	Spirometric data correlated with decision for despatch of mobile intensive care unit on 56% of occasions
Finkelstein[70]	I–A	Pulmonary	31 patients	Comparison between supervised and unsupervised home spirometry	No difference between supervised and unsupervised testing	Not assessed

Other clinical specialties

Eleven other studies assessed diagnostic accuracy in a variety of clinical specialties. All of these studies either lacked a definitive gold standard or did not assess inter-observer agreement in addition to telemedicine versus in-person agreement. Most had relatively small numbers of patients and, especially, clinicians[38–48].

Office/hospital-based telemedicine — store-and-forward

The two most common specialties in store-and-forward office/hospital-based telemedicine were dermatology (11 studies) and ophthalmology (five studies).

Dermatology

Two high-quality store-and-forward studies compared diagnostic agreement between teleconsultation and in-person consultation with agreement among different in-person consultants. The first was carried out in a university general dermatology clinic[23]. The information transmitted consisted of a brief statement of the patient history and appropriate images. The concordance between store-and-forward and in-person consultation (83%) was comparable to inter-dermatologist (81%) and intra-dermatologist (84% — two months later) rates. A second study that assessed intramodality as well as intermodality agreement was performed in a Department of Veterans Affairs clinic[49]. This study compared diagnostic and management agreement among patients seen by five examiners, two in the clinic and three using digital images along with a standardized history form. Agreement on the exact diagnosis was 54% for the clinic dermatologists, 41–55% between the clinic and teledermatologists, and 49–55% among the teledermatologists. Higher concordance was obtained when partial agreement over a differential diagnosis was assessed. Most of the differences in agreement and accuracy were not significant. Neither of these high-quality studies made direct comparisons with interactive teledermatology. Nonetheless, they demonstrate that store-and-forward teledermatology can be nearly as accurate for diagnosis as in-person consultation. The other studies examined only inter-modality agreement and showed a range of diagnostic agreement[50–58].

Ophthalmology

Two studies assessed the ability of telemedicine to screen for diabetic retinopathy. One showed relatively high sensitivity and specificity, of 85% and 90%, respectively[59], while the other[60] showed that digital images were more difficult to interpret than photographs, with a kappa of <0.3. Two studies evaluated the ability of telemedicine to screen for glaucoma and found moderate rates of agreement[61,62]. Another study showed high diagnostic but less management decision agreement in paediatric ophthalmology[63].

Other clinical specialties

An outpatient study looked at telemedicine encounters performed on 20 patients from four specialties — cardiology, dermatology, endocrinology and orthopaedics[64]. Fifteen of these patients also had a face-to-face encounter. For those having both store-and-forward and in-person examinations, there was complete agreement on diagnoses, although each modality led to some change in management. Studies in other specialties showed varying rates of agreement[44,65–68].

Home-based telemedicine

Since diagnosis is not usually the focus of home-based telemedicine, few studies have assessed its diagnostic capabilities. However, home spirometry is a diagnostic area that has been evaluated in more than one study. In studies of home-based lung transplant and asthma patients, it was found that home pulmonary function results were comparable to those in the laboratory[69,70]. Another small, uncontrolled study of patients with moderate to severe asthma found that home spirometry measurements led to appropriate triage decisions in only half the cases[71]. A study of home monitoring of vital signs showed relatively close rates of agreement in measuring heart rate, respiratory rate, temperature and blood pressure between a telemedicine system, a local reference system and a home nurse[72].

Discussion

Despite the widespread use of telemedicine in most major specialties, there is strong evidence in only a few of them that the diagnosis and management decisions provided by telemedicine are comparable to those made face to face. Well designed studies indicate that the best evidence for the diagnostic efficacy of telemedicine is in the areas of psychiatry and dermatology. One anomalous finding was that store-and-forward teledermatology appears to have better diagnostic capability than interactive teledermatology. The most likely explanation is that the best store-and-forward study used good-quality static digital images[23], whereas the best interactive study used video images of lower spatial resolution[22]. Another possibility is that store-and-forward telemedicine is more effective than interactive telemedicine, a non-intuitive hypothesis that would need further investigation.

There is also reasonable evidence that history and physical examinations performed via telemedicine yield reasonably good sensitivity and specificity, especially in general medicine and cardiology, but perhaps less so in areas where patient abnormalities are visually finer, such as rheumatology. In ophthalmology, telemedicine is more effective in some areas (e.g. retinitis) than others (e.g. screening for glaucoma or diabetic retinopathy). There are many more clinical specialties, however, where telemedicine is used, yet either poor studies or no studies have assessed its efficacy.

Demonstrating diagnostic efficacy is not, of course, the same as showing overall efficacy for telemedicine. Telemedicine interventions must be shown to achieve at least equivalent patient outcomes, ideally at similar or reduced costs. A number of systematic reviews have concluded that there have been few well designed trials assessing clinical outcomes, and not all of those which have been well designed have shown telemedicine to be comparable to in-person care[2–6,9]. Little can be said about the costs of telemedicine from reviewing cost assessments in the peer-reviewed literature[2,9].

Further studies must therefore be done to demonstrate the efficacy of telemedicine for diagnostic and management decisions. In addition, care must be taken to avoid the common pitfalls of diagnostic test research[73,74]. Based on the problems with studies assessed in this systematic review, future studies must include larger numbers of patients and clinicians. They should also measure either diagnostic capability and management decisions relative to an appropriate gold standard or measure inter-observer agreement between and within telemedicine and in-person methodologies. The fact that telemedicine is a new technique is not a reason for failing to perform appropriate evaluation studies. New methodologies such as 'tracker trials' could be used to assess it systematically[75].

The growth of high-speed telecommunications networks and their ability to transmit high-quality images and other clinical information suggests a bright future for telemedicine, especially as the infrastructure is extended to individuals who have historically been denied access to health-care because of their remote location. As the cost of these services is likely to be significant, it is imperative that decisions to use them be based on appropriate evidence that they lead to similar diagnostic and management decisions, as well as patient outcomes.

Acknowledgements: This study was conducted under Agency for Healthcare Research and Quality contract 290-97-0018, task order no. 3, Rockville, MD. No statement in this article should be construed as an official position of the Agency for Healthcare Research and Quality or the US Department of Health and Human Services.

References

1 Perednia DA, Allen A. Telemedicine technology and clinical applications. *Journal of the American Medical Association* 1995;**273**:483–8

2 Hersh WR, Helfand M, Wallace JA, *et al*. *Telemedicine for the Medicare Population. Evidence Report/Technology Assessment No. 24*. AHRQ Publication No. 01-E012. Rockville, MD: Agency for Healthcare Research and Quality, 2001. See http://www.ahcpr.gov/clinic/evrptfiles.htm#telemed. Last checked 8 February 2002

3 Balas EA, Jaffrey F, Kuperman GJ, *et al*. Electronic communication with patients. Evaluation of distance medicine technology. *Journal of the American Medical Association* 1997;**278**:152–9

4 Ohinmaa A, Hailey D, Roine D. *The Assessment of Telemedicine: General Principles and a Systematic Review*. Helsinki: Finnish Office for Health Care Technology Assessment, 1999. See http://www.ahfmr.ab.ca/hta/hta-publications/reports/telemedicine99/. Last checked 18 February 2002

5 Almazan C, Gallo P. *Assessing Clinical Benefit and Economic Evaluation in Telemedicine*. Barcelona: Catalan Agency for Health Technology Assessment and Research, 1999. See http://www.aatm.es/ang/informes/i.html. Last checked 8 February 2002

6 Currell R, Urquhart C, Wainwright P, Lewis R. Telemedicine versus face to face patient care: effects on professional practice and health care outcomes. *Cochrane Database of Systematic Reviews* 2000;**2**:CD002098. Review

7 Mair F, Whitten P. Systematic review of studies of patient satisfaction with telemedicine. *British Medical Journal* 2000;**320**:1517–20

8 Whitten P, Kingsley C, Grigsby J. Results of a meta-analysis of cost–benefit research: is this a question worth asking? *Journal of Telemedicine and Telecare* 2000;**6**:4–6

9 Hersh WR, Wallace JA, Kraemer DK, *et al*. *Telemedicine for the Medicare Population: Pediatric, Obstetric, and Clinician-Indirect Home Interventions in Telemedicine. Evidence Report/Technology Assessment No. 24, Supplement*. AHRQ Publication No. 01-E060. Rockville, MD: Agency for Healthcare Research and Quality 2001. See http://www.ahcpr.gov/clinic/evrptfiles.htm#telemed. Last checked 8 February 2002

10 *Telehealth and Telemedicine in Quebec: Current Issues*. Montreal: Conseil d'Evaluation des Technologies de la Sante du Quebec, 1999

11 Telemedicine Strategic Healthcare Group. *Issues and Opportunities Facing VHA Telemedicine: Reinventing Service to Patients*, 1999. See http://www.va.gov/telemed/policy.htm. Last checked 8 February 2002.

12 Grigsby B, Brown N. *1998 Report on U.S. Telemedicine Activity*. Portland, OR: Association of Telemedicine Service Providers, 1998

13 Rosenthal B, ed. *Telemedicine Sourcebook 1998: A Resource Guide for the Practical Application of Telemedicine in the Managed Care Marketplace*. New York: Faulkner and Gray, 1997

14 Chesnut RM, Carney N, Maynard H, Patterson PK, Mann NC, Helfand M. *Rehabilitation for Traumatic Brain Injury*. Rockville, MD: Agency for Health Care Policy and Research, 1999. See http://www.ahcpr.gov/clinic/epcsums/tbisumm.htm. Last checked 11 February 2002

15 Baigent MF, Lloyd CJ, Kavanagh SJ, *et al*. Telepsychiatry: 'tele' yes, but what about the 'psychiatry?' *Journal of Telemedicine and Telecare* 1997;**3** (suppl. 1):3–5

16 Ruskin PE, Reed S, Kumar R, *et al*. Reliability and acceptability of psychiatric diagnosis via telecommunication and audiovisual technology. *Psychiatric Services* 1998;**49**:1086–8

17 Elford R, White H, Bowering R, *et al*. A randomized, controlled trial of child psychiatric assessments conducted using videoconferencing. *Journal of Telemedicine and Telecare* 2000;**6**:73–82
18 Ball C, Puffett A. The assessment of cognitive function in the elderly using videoconferencing. *Journal of Telemedicine and Telecare* 1998;**4**:36–8
19 Ball CJ, Mclaren PM. Comparability of face-to-face and videolink administration of the Brief Psychiatric Rating Scale. *American Journal of Psychiatry* 1995;**152**:958–9
20 Montani C, Billaud N, Couturier P, *et al*. 'Telepsychometry': a remote psychometry consultation in clinical gerontology: preliminary study. *Telemedicine Journal* 1996;**2**:145–50
21 Kirkwood KT, Peck DF, Bennie L. The consistency of neuropsychological assessments performed via telecommunication and face to face. *Journal of Telemedicine and Telecare* 2000;**6**:147–51
22 Lesher JL Jr, Davis LS, Gourdin FW, English D, Thompson WO. Telemedicine evaluation of cutaneous diseases: a blinded comparative study. *Journal of the American Academy of Dermatology* 1998;**38**:27–31
23 Krupinski EA, LeSueur B, Ellsworth L, *et al*. Diagnostic accuracy and image quality using a digital camera for teledermatology. *Telemedicine Journal* 1999;**5**:257–63
24 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. *Journal of Telemedicine and Telecare* 1998;**4**:95–100
25 Oakley AM, Astwood DR, Loane M, Duffill MB, Rademaker M, Wootton R. Diagnostic accuracy of teledermatology: results of a preliminary study in New Zealand. *New Zealand Medical Journal* 1997;**110**:51–3
26 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. *British Journal of Dermatology* 1998;**139**:81–7
27 Lowitt MH, Kessler II, Kauffman CL, Hooper FJ, Siegel E, Burnett JW. Teledermatology and in-person examinations: a comparison of patient and physician perceptions and diagnostic agreement. *Archives of Dermatology* 1998;**134**:471–6
28 Phillips CM, Burke WA, Allen MH, Stone D, Wilson JL. Reliability of telemedicine in evaluating skin tumors. *Telemedicine Journal* 1998;**4**:5–9
29 Phillips CM, Burke WA, Shechter A, Stone D, Balch D, Gustke S. Reliability of dermatology teleconsultations with the use of teleconferencing technology. *Journal of the American Academy of Dermatology* 1997;**37**:398–402
30 Belmont JM, Mattioli LF, Goertz KK, Ardinger RH Jr, Thomas CM. Evaluation of remote stethoscopy for pediatric telecardiology. *Telemedicine Journal* 1995;**1**:133–49
31 McConnell ME, Steed RD, Tichenor JM, Hannon DW. Interactive telecardiology for the evaluation of heart murmurs in children. *Telemedicine Journal* 1999;**5**:157–61
32 Casey F, Brown D, Craig BG, Rogers J, Mulholland HC. Diagnosis of neonatal congenital heart defects by remote consultation using a low-cost telemedicine link. *Journal of Telemedicine and Telecare* 1996;**2**:165–9
33 Srikanthan VS, Pell AC, Prasad N, *et al*. Use of fax facility improves decision making regarding thrombolysis in acute myocardial infarction. *Heart* 1997;**78**:198–200
34 Trippi JA, Kopp G, Lee KS, *et al*. The feasibility of dobutamine stress echocardiography in the emergency department with telemedicine interpretation. *Journal of the American Society of Echocardiography* 1996;**9**:113–18
35 Lee JK, Renner JB, Saunders BF, *et al*. Effect of real-time teleradiology on the practice of the emergency department physician in a rural setting: initial experience. *Academic Radiology* 1998;**5**:533–8
36 Kofos D, Pitetti R, Orr R, Thompson A. Telemedicine in pediatric transport: a feasibility study. *Pediatrics* 1998;**102**:E58
37 Tachakra S, Wiley C, Dawood M, Sivakumar A, Dutton D, Hayes J. Evaluation of telemedical support to a free-standing minor accident and treatment service. *Journal of Telemedicine and Telecare* 1998;**4**:140–5
38 Baur DA, Pusateri AE, Kudryk VL, *et al*. Accuracy of orthognathic evaluation using telemedicine technology. *Telemedicine Journal* 1998;**4**:153–60
39 Hubble JP, Pahwa R, Michalek DK, Thomas C, Koller WC. Interactive video conferencing: a means of providing interim care to Parkinson's disease patients. *Movement Disorders* 1993;**8**:380–2
40 Craig JJ, McConville JP, Patterson VH, Wootton R. Neurological examination is possible using telemedicine. *Journal of Telemedicine and Telecare* 1999;**5**:177–81
41 Marcus DM, Brooks SE, Ulrich LD, *et al*. Telemedicine diagnosis of eye disorders by direct ophthalmoscopy. A pilot study. *Ophthalmology* 1998;**105**:1907–14
42 Nitzkin JL, Zhu N, Marier RL. Reliability of telemedicine examination. *Telemedicine Journal* 1997;**3**:141–57
43 Pedersen S, Hartviksen G, Haga D. Teleconsultation of patients with otorhinolaryngologic conditions. A telendoscopic pilot study. *Archives of Otolaryngology, Head and Neck Surgery* 1994;**120**:133–6
44 Sclafani AP, Heneghan C, Ginsburg J, Sabini P, Stern J, Dolitsky JN. Teleconsultation in otolaryngology: live versus store and forward consultations. *Otolaryngology—Head and Neck Surgery* 1999;**120**:62–72
45 Pacht ER, Turner JW, Gailiun M, *et al*. Effectiveness of telemedicine in the outpatient pulmonary clinic. *Telemedicine Journal* 1998;**4**:287–92
46 Hayes WS, Tohme WG, Komo D, *et al*. A telemedicine consultative service for the evaluation of patients with urolithiasis. *Urology* 1998;**51**:39–43
47 Graham LE, McGimpsey S, Wright S, *et al*. Could a low-cost audio-visual link be useful in rheumatology? *Journal of Telemedicine and Telecare* 2000;**6** (suppl. 1):35–7
48 Tachakra S, Lynch M, Newson R, *et al*. A comparison of telemedicine with face-to-face consultations for trauma management. *Journal of Telemedicine and Telecare* 2000;**6** (suppl. 1):178–81
49 Whited JD, Hall RP, Simel DL, *et al*. Reliability and accuracy of dermatologists' clinic-based and digital image consultations. *Journal of the American Academy of Dermatology* 1999;**41**:693–702
50 Zelickson BD, Homan L. Teledermatology in the nursing home. *Archives of Dermatology* 1997;**133**:171–4
51 Kvedar JC, Edwards RA, Menn ER, *et al*. The substitution of digital images for dermatologic physical examination. *Archives of Dermatology* 1997;**133**:161–7
52 Taylor P. An assessment of the potential effect of a teledermatology system. *Journal of Telemedicine and Telecare* 2000;**6** (suppl. 1):74–6
53 Whited JD, Mills BJ, Hall RP, Drugge RJ, Grichnik JM, Simel DL. A pilot trial of digital imaging in skin cancer. *Journal of Telemedicine and Telecare* 1998;**4**:108–12
54 Harrison PV, Kirby B, Dickinson Y, Schofield R. Teledermatology—high technology or not? *Journal of Telemedicine and Telecare* 1998;**4** (suppl. 1):31–2
55 Braun RP, Meier M, Pelloni F, *et al*. Teledermatoscopy in Switzerland: a preliminary evaluation. *Journal of the American Academy of Dermatology* 2000;**42**:770–5
56 High WA, Houston MS, Calobrisi SD, Drage LA, McEvoy MT. Assessment of the accuracy of low-cost store-and-forward teledermatology consultation. *Journal of the American Academy of Dermatology* 2000;**42**:776–83
57 Lewis K, Gilmour E, Harrison PV, *et al*. Digital teledermatology for skin tumours: a preliminary assessment using a receiver operating characteristics (ROC) analysis. *Journal of Telemedicine and Telecare* 1999;**5** (suppl. 1):57–8
58 Lyon CC, Harrison PV. A portable digital imaging system in dermatology: diagnostic and educational applications. *Journal of Telemedicine and Telecare* 1997;**3** (suppl. 1):81–3
59 Liesenfeld B, Kohner E, Piehlmeier W, *et al*. A telemedical approach to the screening of diabetic retinopathy: digital fundus photography. *Diabetes Care* 2000;**23**:345–8

60 Yogesan K, Constable IJ, Barry CJ, Eikelboom RH, McAllister IL, Tay-Kearney ML. Telemedicine screening of diabetic retinopathy using a hand-held fundus camera. *Telemedicine Journal* 2000;**6**:219–23
61 Yogesan K, Constable IJ, Barry CJ, *et al*. Evaluation of a portable fundus camera for use in the teleophthalmologic diagnosis of glaucoma. *Journal of Glaucoma* 1999;**8**:297–301
62 Li HK, Tang RA, Oschner K, Koplos C, Grady J, Crump WJ. Telemedicine screening of glaucoma. *Telemedicine Journal* 1999;**5**:283–90
63 Schwartz SD, Harrison SA, Ferrone PJ, Trese MT. Telemedical evaluation and management of retinopathy of prematurity using a fiberoptic digital fundus camera. *Ophthalmology* 2000;**107**:25–8
64 Houston MS, Myers JD, Levens SP, *et al*. Clinical consultations using store-and-forward telemedicine technology. *Mayo Clinic Proceedings* 1999;**74**:764–9
65 Patterson S, Botchway C. Dental screenings using telehealth technology: a pilot study. *Journal of the Canadian Dental Association* 1998;**64**:806–10
66 Yamamoto LG, Ash KM, Boychuk RB, *et al*. Personal computer teleradiology interhospital image transmission of neonatal radiographs to facilitate tertiary neonatology telephone consultation and patient transfer. *Journal of Perinatology* 1996;**16**:292–8
67 Furukawa M, Furukawa MK, Mizojiri G, Matsuda H. Telemedicine in laryngology. *Telemedicine Journal* 1998;**4**:329–33
68 Wirthlin DJ, Buradagunta S, Edwards RA, *et al*. Telemedicine in vascular surgery: feasibility of digital imaging for remote management of wounds. *Journal of Vascular Surgery* 1998;**27**:1089–99
69 Finkelstein SM, Lindgren B, Prasad B, *et al*. Reliability and validity of spirometry measurements in a paperless home monitoring diary program for lung transplantation. *Heart and Lung* 1993;**22**:523–33
70 Finkelstein J, Cabrera MR, Hripcsak G. Internet-based home asthma telemonitoring: can patients handle the technology? *Chest* 2000;**117**:148–55
71 Bruderman I, Abboud S. Telespirometry: novel system for home monitoring of asthmatic patients. *Telemedicine Journal* 1997;**3**:127–33
72 de Lusignan S, Althans A, Wells S, Johnson P, Vandenburg M, Robinson J. A pilot study of radiotelemetry for continuous cardiopulmonary monitoring of patients at home. *Journal of Telemedicine and Telecare* 2000;**6** (suppl. 1):19–22
73 Reid MC, Lachs MS, Feinstein AR. Use of methodological standards in diagnostic test research: getting better but still not good. *Journal of the American Medical Association* 1995;**274**:645–51
74 Lijmer JG, Mol BW, Heisterkamp S, *et al*. Empirical evidence of design-related bias in studies of diagnostic tests. *Journal of the American Medical Association* 1999;**282**:1061–6
75 Lilford RJ, Braunholtz DA, Greenhalgh R, Edwards SJ. Trials and fast changing technologies: the case for tracker studies. *British Medical Journal* 2000;**320**:43–6

Appendix. Search strategy

The strategy outlined below was used for MEDLINE. Comparable strategies were used for EMBASE, CINAHL and HealthSTAR.

Initial search strategy

(1) exp telemedicine/
(2) telemedicine.mp.
(3) telehealth.tw.
(4) remote consultation$.mp.
(5) 1 or 2 or 3 or 4
(6) exp home care services/
(7) home nursing/
(8) 6 or 7
(9) exp therapy, computer-assisted/
(10) exp computers/
(11) exp computer communication networks/
(12) exp medical informatics/
(13) exp telecommunications/
(14) exp monitoring, physiologic/
(15) monitor$.tw.
(16) blood glucose self-monitoring/
(17) self-examination/
(18) self exam$.tw.
(19) self monitor$.tw.
(20) self test$.tw.
(21) 14 or 15 or 16 or 17 or 18 or 19 or 20
(22) tele$.tw.
(23) (remote or offsite or distance).tw. tw=abstract, title
(24) rural population/
(25) rural health services/
(26) hospitals, rural/
(27) rural.tw.
(28) 22 or 23 or 24 or 25 or 26 or 27
(29) 21 and 28
(30) 9 or 10 or 11 or 12 or 13 or 29
(31) 8 and 30
(32) 31 not 5
(33) limit 32 to english language
(34) 32 not 33
(35) limit 34 to abstracts
(36) 33 or 35

Supplementary search strategy (to identify more home-based telemedicine articles)

(1) exp computer communication networks
(2) patient participation
(3) exp consumer satisfaction
(4) delivery of health care
(5) exp home care services
(6) exp home nursing
(7) house calls/ or housecalls.mp
(8) 2 or 3 or 4 or 5 or 6 or 7
(9) 1 and 8
(10) limit 9 to english language