

Evaluating telemedicine: lessons and challenges

Michelle Brear

Abstract

Telemedicine applications have been rapidly implemented since the early 1990s and are now in use in a wide range of healthcare settings. There is, however, limited evidence of clinical benefits resulting from their use. This paper describes the design and implementation of a multi-method evaluation of an emergency telemedicine application's clinical impact. In particular, the challenges faced and lessons learnt regarding the development of a suitable methodology and collection of health information from a variety of sources are discussed. In order to understand the application's clinical impact it was necessary to use different types of data from multiple sources, and to interpret the results from each enquiry in relation to results from the others.

Keywords (MeSH):

Telemedicine; Evaluation Studies, Data Collection; Evaluation Methodology; Information Management.

Introduction

The use of telemedicine applications (i.e. applications which enable the practice of medicine at a distance) has increased remarkably since the early 1990s (Hersh et al. 2001). Rapid technological advances, most notably the emergence of the Internet and high bandwidth connections to it, have created new and significant opportunities for telemedicine use (Benger 2000). The need to improve the efficiency and equity of health service delivery, in response to the increasing specialisation and shortage of medical professionals and centralisation of health facilities, has contributed to the rapid implementation of telemedicine.

Technologically advanced telemedicine applications are now used in a wide range of healthcare settings (e.g. hospital and general practices) and medical specialties (e.g. radiology and neurology) (Ohinmaa, Hailey & Roine 2001). They enable access to specialist expertise at previously unavailable times, for example by transmitting images to mobile phones or wireless pocket devices (Yamamoto & Williams 2000); and from previously inaccessible locations for example

submarines and space-ships (Bland 2000). Telemedicine applications have the potential to reduce patient travel requirements, for example by facilitating the home monitoring or treatment of patients undergoing rehabilitation (Phillips et al. 2001), or those who are chronically ill (McKay et al. 2002). Emergency telemedicine applications can enhance the timeliness of care delivery (Schwamm et al. 2004). They are also being implemented in order to provide previously unavailable health services, for example telephone triage services (O'Connell et al. 2001).

Telemedicine in Australia

In Australia, where the geographically dispersed population complicates the provision of equitable and efficient health care, telemedicine has been practiced for decades (Ellis 2004). In many remote areas where doctors are not physically located, nurses continue to utilise the telephone to assist their clinical practice, as they have done so in the past (Ellis 2004). The facsimile has been used to transmit images such as echocardiograms and diagrammatic instructions about treatment (Ellis & Riley 2005). These telemedi-

cine applications have facilitated the provision of timely care and reduced unnecessary transport of patients. Similar applications are also widely and frequently used in non-remote areas, however technological advances have led to the implementation of many new and innovative telemedicine applications based on video conferencing, email and the Internet in Australia in recent years.

Telemedicine has been most frequently applied in Australia to enable the provision of psychiatry, ophthalmology and dermatology services to geographically remote locations (Ellis 2004). These applications are now expanding into almost all areas of health service provision being used in remote and non-remote locations. A 2003 survey revealed that approximately 50% of Australian hospitals are using telemedicine applications (Wootton, Blignault & Cignoli 2003). Telemedicine applications are now used to provide specialist advice about paediatric care in areas as diverse as cardiography, sexual assault and palliative care (Wootton & Batch 2004). In remote areas, for example, they are being piloted to assist the treatment of lower limb ulcers (Santamaria & Carville 2004), diagnosis of radiographs (Dillon & Loermoans 2003) and post-operative burns care (Johansen et al. 2004). There are also telemedicine links between urban clinicians (Olver & Selva-Nayagam 2000).

These examples demonstrate the enormous potential for the employment of telemedicine applications to enhance the timeliness, effectiveness and efficiency of health care delivery, which in turn could produce cost savings to both the health system and individual patients. Costs can potentially be minimised by reducing the need to transport patients to central locations (Ohinmaa, Hailey & Roine 2001) and preventing costly health complications resulting from delayed treatment (Schwamm et al. 2004). Telemedicine applications can potentially increase the equity of health service provision between remote and urban locations (Ellis 2004). They also have the potential to improve working conditions for health service providers in remote locations, by providing additional support and educational opportunities (Olver & Selva-Nayagam 2000).

Methodological issues in evaluation of telemedicine applications

Current evidence

Despite the rapid implementation of and enormous potential for telemedicine applications, systematic reviews of literature published internationally have revealed that evidence of clinical or other benefits resulting from their use is scarce. One review of telemedicine randomised controlled trials (RCTs) concluded that 'despite the widespread use of telemedicine in virtually all major areas of health care, evidence concerning the benefits and its use exists in only a small number of them' (Hersh et al. 2001). Another, which also assessed non-randomised controlled trials, indicated that '...most of the available literature refers only to pilot projects and short terms outcomes...[T]elemedicine applications still lack scientific evidence regarding their effectiveness' (Ohinmaa, Hailey & Roine 2001).

The lack of available evidence severely limits the capacity of health service providers to make informed decisions about the appropriate use of telemedicine applications. Without knowing whether or not, to what degree, and for which medical specialties telemedicine can produce clinical benefits, it is not possible to assess the potential cost effectiveness in comparison to other models of care (e.g. transporting patients), nor is it possible to determine the most suitable applications. As technologically advanced telemedicine applications have been used for such a short period of time, the lack of evidence is not entirely surprising. However, the considerable resources now being directed to the implementation of telemedicine applications need to be justified by the production of evidence regarding its clinical impact. To utilise health information effectively from existing sources is a key challenge in the production of such evidence.

Challenges in telemedicine evaluation

There are many challenges to producing reliable evidence about the effects of telemedicine applications. Evaluation is rarely seen as an integral part of implementation, thus resources are not usually dedicated to evaluation. Sometimes resources are dedicated after implementation, so prospective collection of baseline data is not

possible. Retrospective collection of health information presents many challenges in terms of identifying appropriate sources and collection techniques. These issues complicate the evaluation of many interventions, while the specific nature of telemedicine evaluation presents further complications.

Most telemedicine interventions have been pilot projects, limited to a single site and influencing the care of a small number of patients. This limits the extent to which traditionally employed methods of enquiry and statistical analysis can produce meaningful results about the clinical and economic impact of an application (e.g. randomised controlled trials are designed to evaluate large, not small, populations). Telemedicine applications depend on communication between two sites which may have different methods of recording and storing data. The context of each site, the perspectives of the individuals working within them and the patterns by which they work, are also likely to differ. These factors will influence the effective use of, and clinical benefit resulting from, a telemedicine application and the extent to which similar results can be expected in other organisations (Lorenzi 2004). The technical features and performance of a telemedicine application will also influence the clinical results, for example if an application does not allow clinicians to follow their normal work routines, it may not be used and thus will not produce clinical benefits (Kaplan 2001).

Thus, a telemedicine evaluation must examine not only to what extent an application was effective, but also why it was or was not effective. An evaluation methodology must utilise health information data not only from traditionally utilised, mainly quantitative health information sources (e.g. medical records data),

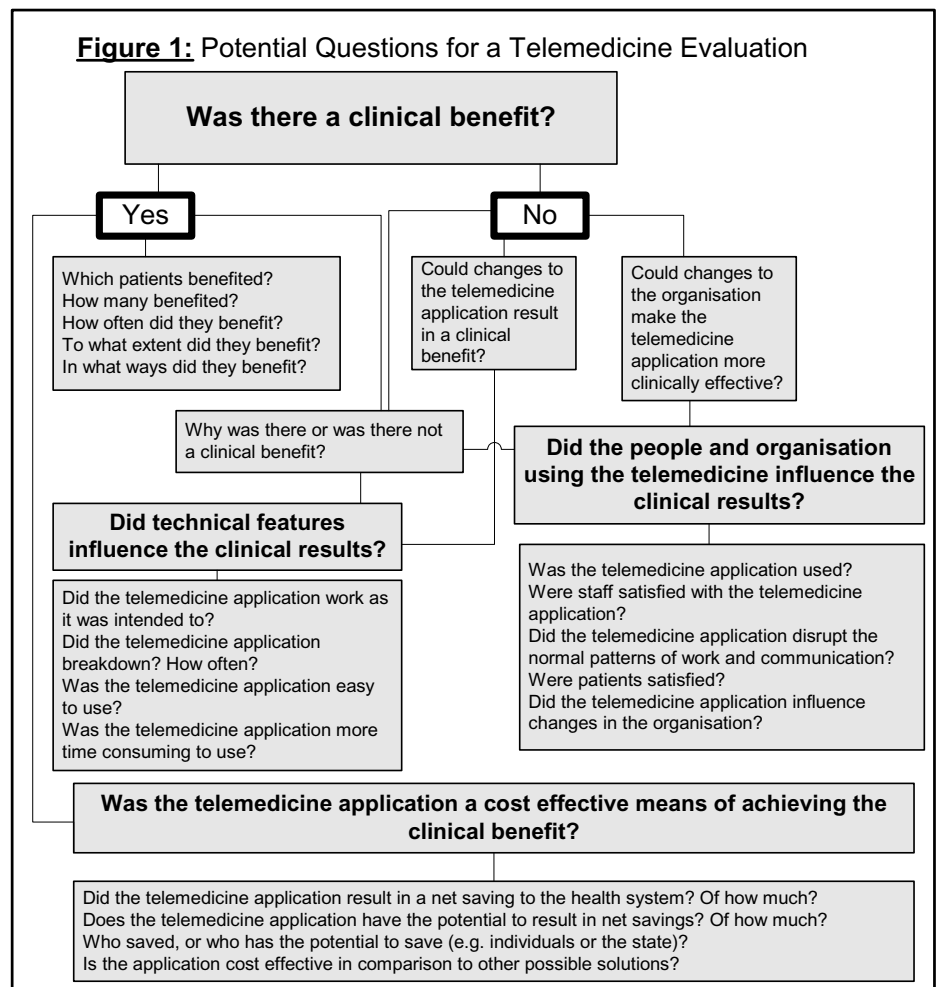
but also from qualitative sources (e.g. interview data with doctors and nurses). Due to the infancy of telemedicine projects, reliable evaluation methodologies, which consider both if and why an application was effective, have not been tested.

Questions for telemedicine evaluation

The various factors influencing the effective use of telemedicine applications lead to multiple evaluation questions¹ (Figure 1). As most telemedicine projects aim to enhance the provision of clinical care, the pertinent question for most evaluation projects is ‘was there a clinical benefit?’ There are however many sub-questions which must be answered in order to fully understand the nature and extent of such clinical benefit, if the telemedicine is found to be beneficial, or the reasons why a clinical benefit did not ensue if this was the case (Figure 1).

The available resources and aims of telemedicine implementation will determine which and

¹ The questions shown in Figure 1 were developed by researchers in conjunction with key informants from the participating hospitals.



how many questions are answered. Economic resources will place limits on the extent of the study and human resources (e.g. the availability of people and the skills they possess) may impose further limits. However determining whether or not the telemedicine application produces a clinical benefit is essential, in order to assess, meaningfully, the cost effectiveness or influence of technical features and the organisation into which an application was implemented.

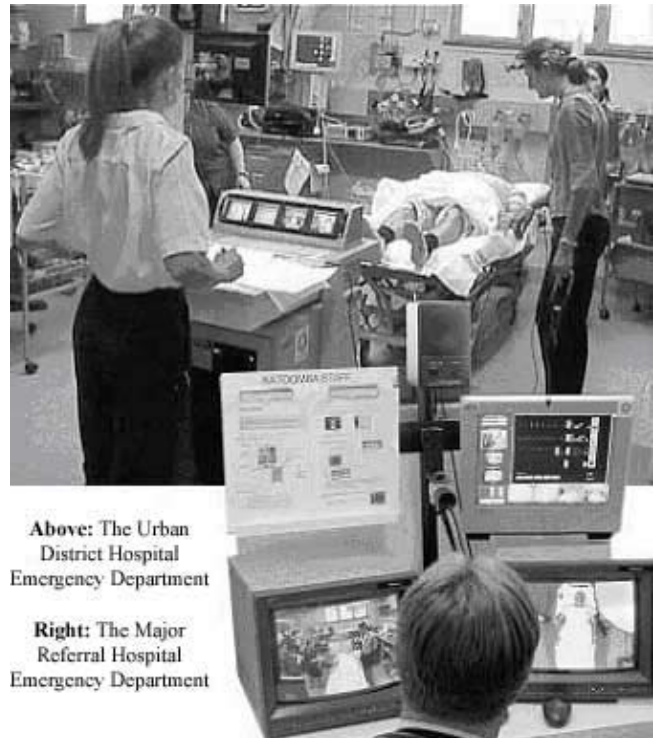
Research methods and techniques for telemedicine evaluation

Once evaluation questions have been determined there is a range of methods and techniques which can be applied to answer them. The available human and economic resources will also influence this choice, as will the philosophical perspectives of the researchers and availability of data. There is no ultimately right or wrong method; each has strengths and weaknesses which must be considered in relation to the specific nature of telemedicine evaluation.

It has traditionally been considered that randomised controlled trials (RCTs) produce the highest quality evidence of the effect of an intervention because the process of randomisation controls for factors which may confound the intervention's effect (e.g. staff changes) (Friedman & Wyatt 1997). However telemedicine evaluations do not always use RCTs. Assumptions that telemedicine applications will result in improved clinical care can create ethical concerns about randomising patients to non-telemedicine models of care. The small number of patients treated at rural and remote health facilities where telemedicine applications are usually implemented, make obtaining a sample sufficient to conduct a RCT impractical.

Due to these difficulties many telemedicine evaluations have employed non-randomised controlled trials. Like RCTs, they compare the results for a telemedicine treated and a non-telemedicine treated group of patients, and telemedicine evaluations often take the form of experiments or 'before and after' studies. In experiments clinicians use a telemedicine application to enact scenarios or provide diagnoses based on retrospective cases. Such tests can show an application's potential to perform in ideal condi-

Figure 2: The Virtual Critical Care Unit (Wilson, 2003)



Above: The Urban District Hospital Emergency Department

Right: The Major Referral Hospital Emergency Department

tions; however they cannot take account of how the telemedicine application will actually perform in a clinical environment where stress, work routines and other factors will influence results. Thus one cannot assume that the effects seen in a laboratory will be replicated in the real world. 'Before' and 'after' studies examine telemedicine applications in real work contexts, taking into consideration factors related to the context; however they are often not able to control for changes after the implementation, which may simultaneously influence clinical results. These changes must be considered when attempting to establish a cause and effect relationship between the telemedicine application and clinical results.

Studies which describe the clinical situation following a telemedicine intervention (e.g. number and types of patients treated or the perspectives of clinicians) are not a reliable way of determining a telemedicine application's clinical impact. They can, however, be a useful way of uncovering unexpected effects of a telemedicine application and understanding the specific ways in which it was used and influenced patient care (Friedman & Wyatt, 1997).

Due to the strengths and weaknesses of each method, telemedicine interventions can be

evaluated most effectively by employing a combination of descriptive and comparative techniques. Such an approach allows the identification of unexpected and specific effects of the telemedicine application, as well as the comparison of clinical indicators between telemedicine and non-telemedicine patients. These approaches are known as multi-method or triangulated study designs.

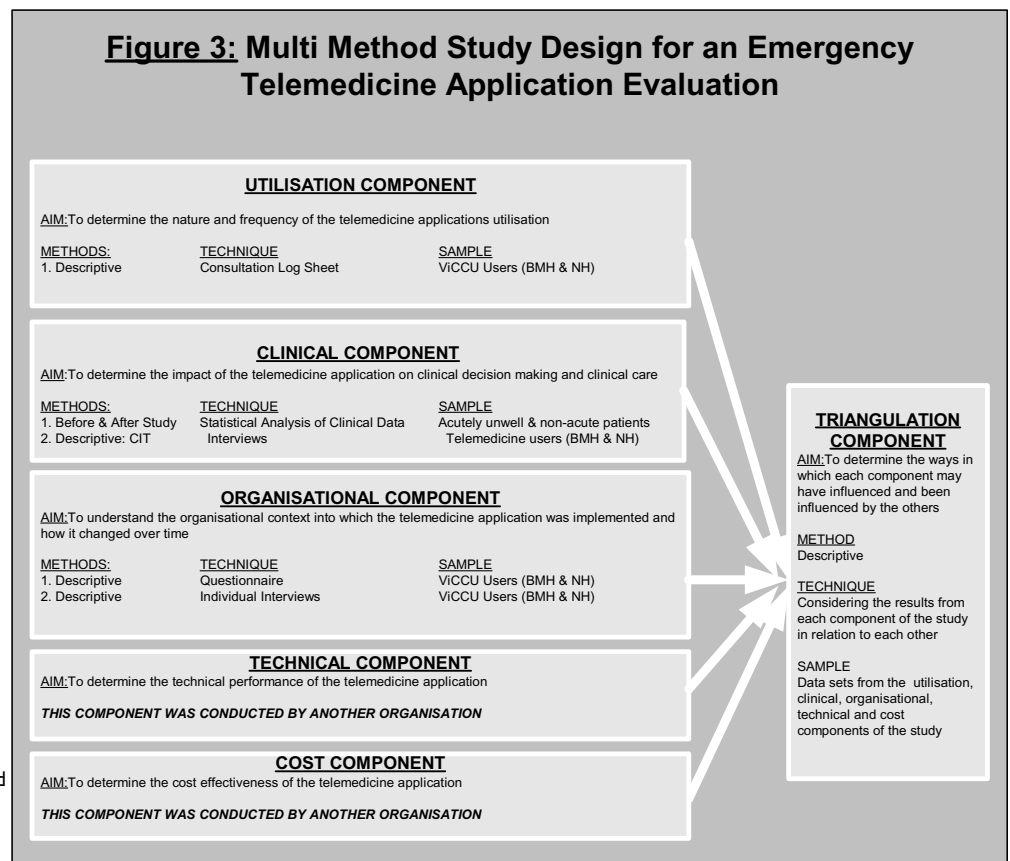
Method: evaluation of an emergency telemedicine application

The Virtual Critical Care Unit

This paper discusses the challenges faced and lessons learnt in the process of designing a methodology to evaluate the clinical impact of an emergency telemedicine application, the Virtual Critical Care Unit (Wilson 2003). The application connects the emergency departments (EDs) of an urban district hospital and a metropolitan referral hospital located approximately 50km apart. The telemedicine application simultaneously transmits up to four real-time video feeds between a team of clinicians working in the resuscitation bay of the urban district hospital ED, and an emergency staff specialist in the metropolitan referral hospital ED, via a dedicated GB/sec connection (Figure 2)² (Wilson 2003). The staff specialist consultant simultaneously views three screen-size, real-time feeds, choosing between feeds from fixed video cameras located above and at the end of the patient’s bed, a portable video camera, a document camera with X-ray

viewing capacity and the patient’s vital signs monitor. A real time video-feed of the consulting doctor in the metropolitan referral hospital is transmitted back to the urban district hospital.

The telemedicine application was implemented to overcome difficulties associated with the itinerant nature and varying expertise of the medical staff in the urban district hospital. The urban district hospital ED provides a full range of emergency services for acutely unwell patients but receives a relatively small number of such presentations (e.g. approximately 100 patients are triaged to the resuscitation category each year). Due to the unpredictable nature of emergency medicine, resuscitation cases may arrive at any time; however the ED is not consistently staffed by doctors with critical care expertise, nor are there sufficiently experienced doctors in other wards of the urban district hospital (Brear et al. 2005). The aim of implementing the telemedicine application was to leverage expertise from the metropolitan referral hospital ED, which is consistently staffed by doctors skilled in emergency medicine, to assist in the clinical decision making for, and care of,



2 The telemedicine application described was developed by the Commonwealth Scientific and Industrial Research Organisation

acutely unwell patients presenting to the urban district hospital.

Study design

A multi-method evaluation, including components for assessing the telemedicine application’s utilisation, clinical and organisational impact, technical functioning and cost effectiveness, was designed (Figure 3). Each component employed different methods of evaluation, while the clinical and organisational components employed multiple methods of enquiry, the results from which were analysed in relation to each other. There was also a component of triangulation, in which the results from each component of the study were analysed in relation to each other (Figure 3). In making a meaningful assessment of the cost effectiveness of the application, the impact of its technical performance or its influence on the organisation was dependent on determining its clinical impact; we thus focused our resources on this component.

Clinical questions and data sources

We asked a number of questions regarding the application’s clinical impact, based on anti-

pated changes (Figure 1). Each of the questions was answered using one or more clinical indicators. The data were derived from electronic and physical patient records and telemedicine users (Table 1). We developed criteria based on triage category and diagnostic codes to select a group of acutely unwell patients for intensive study, as these were the patients most expected to benefit from the telemedicine intervention. We also examined, for all patients presenting to the urban district hospital ED, indicators obtained from the Emergency Department Information System.

Discussion

Limitations of the data collection process

Collection of data from each source posed a number of challenges. Electronic data were extracted from hospital databases relatively easily; however inconsistent coding of data within fields (due to software changes mid-project), and the fact that the sites each used a different medical record number to identify the same patient records, meant that data needed to be recoded and matched manually. Collecting data from physical patient records was time consuming, limiting the number of patients who could be studied intensively. Accessing large volumes of medical records was a strain on the Medical Records Department. Collection of interview data was complicated by difficulties in relieving clinicians from their work in the EDs, although their willingness to participate assisted greatly. None of these issues ultimately prevented us from collecting data; however dealing with each was time consuming.

Reliability and nature of the evidence produced

In the preliminary analysis of data which was performed six months after the evaluation’s commencement, each indicator allowed us to make assessments of the telemedicine’s clinical impact in different ways; however each had

Box 1: Research questions posed to determine the impact of the telemedicine application on clinical decision-making and care

Clinical Decision Making

Does the telemedicine application enable:

- the provision of clinical decision support to UDH clinicians about the treatment and management of patients?
- the nature and gravity of patients’ illnesses/injuries to be diagnosed more accurately in the UDH?
- early recognition of treatment failures and provision of additional treatment or escalated therapy?

Clinical Care

Does the telemedicine application:

- allow the remote supervision of clinical procedures in the UDH ED?
- result in additional clinical procedures being performed in the UDH ED?
- result in an increased number of particular clinical procedures (e.g. intubation) being performed in in the UDH ED?
- result in the improved stabilisation of acutely unwell patients in the UDH ED?
- expedite the transfer of patients from the UDH ED?
- increase the proportion of patients appropriately retained in the UDH ED?
- reduce the proportion of unnecessary transfers from the UDH ED?
- reduce the number of inappropriate admission to the UDH?
- result in reduced hospital stays?

Table 1: Clinical Indicators and Grouping Variable used and their data sources

INDICATOR	COLLECTION TECHNIQUE AND DATA SOURCE
Telemedicine User perceptions	Interviews with Telemedicine Users
Telemedicine User Experiences	Interviews with Telemedicine Users
Clinical Procedures performed	Extracted from Physical Medical Records
Rapid Acute Physiology Score (Rhee et al. 1987)	Extracted from Physical Medical Records
Treatment Time in UDH ED	Extracted from Emergency Department Information System (EDIS)
Time to contact Medical Retrieval Unit	Extracted from Medical Retrieval Unit Electronic Database
Rate of Admission	Extracted from EDIS
Rate of Discharge	Extracted from EDIS
Rate of Transfer	Extracted from EDIS
Rate of transfer for patients admitted to the UDH from ED	Extracted from HOSPAS
Rate of death for patients admitted to BMH from ED	Extracted from HOSPAS
Length of Stay in the UDH	Extracted from HOSPAS
Length of Stay in the MRH (for those patients transferred there)	Extracted from HOSPAS
GROUPING VARIABLES	COLLECTION TECHNIQUE & DATA SOURCE
Case Type	Extracted from Physical Medical Records
Triage Category	Extracted from EDIS

limitations. The perceptions and experiences of telemedicine users identified unanticipated effects of the application (e.g. its use for the treatment of non-acute patients) and important instances of telemedicine use. In the final analysis this data will contribute to our understanding of why the telemedicine application resulted in the particular clinical effects it did. These data were, however, anecdotal and their integrity dependent upon the ability of individuals, some of whom had vested interests in the project, to accurately recall and assess their experiences.

The clinical indicators extracted from medical records will allow a comparative assessment of the telemedicine application's clinical impact. These were, however, collected for the purpose of statistical analysis, a process designed to highlight changing trends in clinical indicators, not to explain the reasons for their presence or absence or specific nature. Preliminary analysis indicates that the statistical analysis of individual indicators, examined in isolation from the others, will be insufficient to assess whether or not changing trends represent clinical benefits.

Examining the qualitative and quantitative indicators together

As many of the questions for this evaluation were posed in terms of appropriateness and necessity, individual indicators were often insufficient to answer the questions posed. In the preliminary analysis, therefore, it was necessary to examine the data from a number of qualitative and quantitative indicators simultaneously, as well as to look for the presence or absence of trends within various patient groups (defined by grouping variables [Table 1]). For example, to determine whether or not the proportion of *unnecessary* transfers had been reduced, it was necessary to examine changes not only to the overall rate of transfer, but to examine the rates of transfer, admission and discharge amongst different patient groups. In addition, the perspectives and experiences of users were valuable in indicating why these changes may have occurred and to what extent they may or may not be appropriate. The perceptions and experiences of staff also guided analysis of the quantitative data; for example they informed the analysis of data collected to determine whether there was an increase to clinical procedures performed, by indicating particular types of procedures worthy

of statistical examination. Examining indicators from a range of sources simultaneously allowed us to make assessments and draw conclusions about the telemedicine's clinical impact which would otherwise not have been possible.

Conclusions

There is an urgent need for evidence about the clinical impact of telemedicine applications. Production of such evidence is complicated by:

- the limited resources dedicated to many telemedicine projects
- the context specific nature of the evidence produced
- the extensive range of potential questions to ask and research methods to apply
- the need to identify and determine the nature and extent of both the intended and unintended consequences of an application.

Using a multi-method approach, which obtains qualitative and quantitative data and analyses it using various techniques while taking into consideration the evidence from other indicators, can overcome some of these complications.

Applying a multi-method approach to the evaluation of an emergency telemedicine application, we faced many challenges and learnt valuable lessons regarding the collection of reliable health information data from multiple sources and development of an analysis strategy which would integrate the various data. Preliminary collection and analysis in this study indicate that:

- while electronic patient data are fast and simple to collect, time needs to be taken to ensure they are accurately coded and matched
- these types of data are appropriate for indicating broad trends, not their specific nature or reasons; more specific clinical data can be obtained from physical patient records; however, it is time consuming to collect and statistical analysis of such data only produces evidence of trends and not the reasons for them
- data from interviews are useful for examining the specific nature of trends and their reasons; they are also time consuming to collect and their integrity is dependent upon the memory and objectivity of the interviewees.

Due to the nature and limitations of evidence which can be derived from each type of data, it

is necessary to use multiple methods of collection and analysis. In this way it is possible to determine the presence or absence of changes resulting from a telemedicine application and explain the nature of and reasons for these changes.

Acknowledgements

This research was funded by NSW Health.

References

- Benger, J. (2000). A review of telemedicine in accident and emergency: the story so far. *Journal of Accident and Emergency Medicine* 17(3): 157-64.
- Bland, S. A. (2000). The submarine service of the future? *Journal of Naval and Medical Service* 86(3): 150-55.
- Brear, M., Westbrook, J., Newman, C., Salter, C., Staplton, S., Murphy, M., Cregan, P. and Coiera, E. (2005). Organisational barriers to telemedicine implementation: results from a qualitative study. In *Proceedings of the Health Informatics Conference 2005*. H. Grain and M. Wise (Eds). Melbourne, Health Informatics Society of Australia.
- Dillon, E. and Loeremoans, J. (2003). Telehealth in Western Australia: the challenge of evaluation. *Journal of Telemedicine and Telecare* 9(Supp 2): S15-9.
- Ellis, I. (2004). Is telehealth the right tool for remote communities? Improving health status in rural Australia. *Contemporary Nurse* 16(3): 163-8.
- Ellis, I. and Riley, D. (2005). The fax as an effective telehealth tool. *Australian Journal of Rural Health* 13(56): 56.
- Friedman, C. P. and Wyatt, J. C. (1997). Subjective approaches to evaluations. In *Evaluation methods in medical informatics*. C.P. Friedman and J. C. Wyatt (Eds). New York, Springer, New York: 205-222.
- Hersh, W. R., Hefland, M., Wallace, J., Kraemer, D., Patterson, P., Shapiro, S. and Greenlick, M. (2001). Clinical outcomes resulting from telemedicine interventions: a systematic review. *BMC Medical Informatics and Decision Making* 1(5).
- Johansen, M., Mill, J., Smith, A., Wooten, R., Kimble, R. and Hockey, A. (2004). A feasibility study of email communication between the patient's family and specialist burns team. *Journal of Telemedicine and Telecare* 10(Supp 1): 53-6.
- Kaplan, B. (2001). Evaluating informatics applications-some alternative approaches: theory, social interactionism, and call for methodological pluralism. *International Journal of Medical Informatics* 64(1): 39-56.
- Lorenzi, N. M. (2004). Beyond the gadgets. *British Medical Journal* 328(7449): 1146-7.
- McKay, H., Glasgow, R., Feil E., Boles, S. and Barrera, M. (2002). Internet-based diabetes self management and support: initial outcomes from the Diabetes Network Project. *Rehabilitation Psychology* 47: 31-48.

- O'Connell, J., Johnson, D., Stallmeyer, J. and Cokington, D. (2001). A satisfaction and return-on-investment study of nurse triage service. *American Journal of Managed Care* 7: 159-69.
- Ohinmaa, A., Hailey, D. and Roine, R. (2001). Assessing telemedicine: a systematic review of the literature. *Canadian Medical Association Journal* 165(6): 765-771.
- Olver, I. and Selva-Nayagam, S. (2000). Evaluation of a telemedicine link between Darwin and Adelaide to facilitate cancer management. *Telemedicine Journal* 6(2): 213-8.
- Phillips, V., Vesmarovich, S., Hauber, R., Wiggers, E. and Egner, A. (2001). Telehealth: reaching out to newly injured spinal cord patients. *Public Health Reports* 116(supp 1): 94-102.
- Rhee, K. J., Fisher, C. J. Jr. and Willitis, N. H. (1987). The rapid acute physiology score. *American Journal of Emergency Medicine*. 5: 278-82.
- Santamaria, M. and Carville, K. (2004). The effectiveness of digital imaging and remote expert wound consultation and healing rates in chronic lower leg ulcers in the Kimberley region of Western Australia. *Primary Intention* 12(2): 66-8.
- Schwamm, L. H., Rosenthal, E.S., Hirshberg, A., Schaefer, P.W., Little, E.A., Kvedar, J.C., Petkovska, I., Koroshetz, W.J. and Levine, S. (2004). Virtual telestroke support for the emergency department evaluation of acute stroke. *Academic Emergency Medicine* 11(11): 1193-7.
- Wootton, R. and Batch, J. Eds (2004). *Telepediatrics: telemedicine and child health*, London, Royal Society of Medicine Press.
- Wootton, R., Blignault, I. and Cignoli J. (2003). A national survey of telehealth in Australian hospitals. *Journal of Telemedicine and Telecare* 9(Supp 2): S73-5.
- Yamamoto, L. G. and Williams, D. R. (2000). A demonstration of instant pocket wireless CT teleradiology to facilitate stat neurosurgical consultation and future telemedicine implications. *American Journal of Emergency Medicine* 18(4): 423-6.
- Wilson, L. S. (2003). Broad band telehealth for complex and critical healthcare delivery. *Proceedings of World Congress on Medical Physics and Biomedical Engineering*, August 24-29, Sydney.

Michelle Brear BAppSci(EnvHealth)

Centre for Health Informatics

Cliffbrook House

University of New South Wales

Coogee, NSW 2052

AUSTRALIA

email: m.brear@hotmail.com

