Data Quality, Health Care Planning and Delivery on Data management, Administrative and Clinical Sources

Karolyn Kerr
The University of Auckland

Abstract
Health care planning and delivery rely heavily on data from management, administrative and clinical sources; nearly all health care activities involve gathering, analysing, or using data. Quality data can lead to quality and cost-effective health care delivery by improving patient outcomes through better decision making. Data quality is inextricably linked to the use of information systems and the health sector is increasingly an information-driven service (Hovenga et al., 1996), where information held in databases and other electronic repositories delivered in a reliable and timely manner, is critical to the health and well-being of patients, the wider population, and to the management of health care organisations (Long & Seko, 2002). Along with an increase in information complexity, there has been a parallel increase in the complex nature of organisations and organisational relationships within the health sector (Gendron & D'Onofrio, 2001).

Keywords: Data Quality, Health Care Planning, Health Delivery

1.0 INTRODUCTION

Health data deal with a continuum that ranges from the clinical records of individual patients, detailing their interactions with health services, through aggregated data and data warehouses, knowledge-based data for planning and decision support, to comparative and community data that can be used at a policy development level (Al-Shorbaji, 2001). Health care data include the administrative data required to manage the complex business of health care, but also data captured from the patient's medical record. Data are often collected at the individual patient level and referred to as 'unit level data'. Aggregated data are the sum of unit level data. Data are used to support the proper care of the patient from whom the data were obtained, and they may also contribute to the good of society through the aggregation and analysis of data concerning populations of individuals. In an aggregated form, health data within the health care organisation provides information of the prevalence and distribution of disease in a population, and at the government level informs policy makers.

Diagnosis relies on data gathered through observations. Gathering data and interpreting their meaning is central to the health care process because these steps are central to the process of decision-making. Increasingly wide ranges of data are collected from the multidisciplinary team to eliminate or confirm the possibility of the presence of a specific disease. Data in health care are found in many different formats and are used in many different ways. One item of datum may be used several times over:

- in the patient's electronic health record;
- in expert systems for artificial intelligence;
- by a decision support system;
- by the hospital administration system;
- for claiming payment from central government or insurance companies for services provided;
- for the purpose of reporting to a national health collection.

Health data can also be made up of varying types of information. Clinical information is generally based on patient related data and often found in patient records. The data describes the characteristics of the patient, of the illness, and of the health care process. Medical knowledge abstracts from the individual patient and describes general insights, for example, about a disease, diagnosis or procedure (Liener, Gaus, Haux, & Knaup-Gregori, 2003) and administrative data provide information on the cost or services required for delivering care.

Within these types the data format also varies. Data may be narrative, textual, numerical measurements, recorded signals, still pictures and videos. In some fields of medicine, data in the form of continuous signals are particularly important, such as the ECG, a tracing of the electrical activity from a patient's heart (Shortliffe & Barnett, 2000). Data collection in health care provides the ability to:

- anticipate future health problems;
- record standard preventative measures;
- identify deviations from expected trends;
- provide a legal record;
- support clinical research (Shortliffe & Barnett, 2000).

Table 5 describes the various types of health care users for the types of data found in health care, and the scope of the data.

Users	Data/Information	Scope
World health officials Policy makers Researchers Lawmakers	General health status and health-related needs of individual nations	World wide data
	Abstracted, Summarised, aggregated	
Policy makers Researchers Lawmakers Insurers	Trend in incidence, prevalence, outcomes, and costs by region, by diagnosis, by type of provider	Nationwide data
	Abstracted, Summarised, aggregated	
Analysts, Researchers Quality management Public health officials	Comparison of treatments, outcomes, and costs by locality and by provider. Incidence and prevalence of diagnosis by region	Community/ region wide data
	Abstracted, Summarised, aggregated	
Administrators Researchers Accreditors Quality managers	Costs of care by category of patient. Number of patients admitted with specific diagnosis, volume of tests, procedures and interventions, outcomes for patients grouped by diagnosis	Funder/ Provider organisation wide data
	Abstracted, Summarised, aggregated	
Care givers Provider organisation departments Insurers QA personnel	Unit level patient specific data e.g. assessments, diagnosis, interventions, diagnostic test results, procedures, treatments, outcomes. Used to provide most appropriate care	Individual patient data

Table 5: Examples of Types and Uses of Health Care Data. Adapted from Saba and McCormick (Saba & McCormick, 2001).

Health data are collected by physicians, nurses, office staff, admissions personnel, physical and respiratory therapists, laboratory personnel, radiology technicians, and pharmacists. Therefore, all those who work in health care collect data through the nature of their occupations, as these data are required in the performance of their jobs. Recording medical data electronically can be difficult when the data can be viewed and perceived in many different ways. A blood pressure is a good example – a reading of 120/80 can be read as a single datum, but also provides decomposed information as a dual view of a systolic blood pressure of 120 and a diastolic pressure of 80. Whilst humans are easily able to make this pluralistic distinction, a computer would need this allowed for in the design process for data storage and processing. The contextual requirements of such information in health means that computers can have a limited use – any subjective or intuitive knowledge of the clinician is difficult to measure and record on a one-level, non-relational database.

2.0 LITERATURE REVIEW
2.1 Medical Registries and Repositories
Progress in information technology and increasing demands for accountability have led to an increase in the number of medical registries over recent years. A medical registry is defined by Arts, De Keizer, and Scheffer as; *a systematic collection of a clearly defined set of health and demographic data for patients with specific health characteristics, held in a central database for a predefined purpose. (Arts, De Keizer, & Scheffer, 2002).* These registries can serve different purposes, such as a tool to monitor and improve quality of care or as a resource for epidemiological research. To be useful data in a medical registry must be of sufficient quality to meet its purpose. However, several studies note the often poor quality of data held in medical registries (Arts et al., 2002; Davidson, Melinkovich, Beatty, Chandramouli, Hambridge, Phibbs, Braun, LeBaron, & Steiner, 2003). Arts defines data quality

within a medical registry as 'the totality of features and characteristics of a data set, that bear on its ability to satisfy the needs that result from the intended use of the data' (Arts et al., 2002). A case study by Broder (Broder, 2004) found that the implementation of a medical registry to monitor quality of care improved patient mortality rates by 28% and morbidity by up to 55%. Using data from the repository, treatment algorithms were developed for patients who are at risk for certain conditions. In some cases, this has involved doing something as simple as putting important information about patient risk factors on laminated pocket cards that doctors can easily reference. Such a case study indicates the impact of access and use of health care data on the delivery of care.

2.2 The Patient Record

The patient record is composed of all data and documents generated or received during the care of a patient at a health care institution. Every patient record contains information on demographics, General Practitioner details, the patients' health history, complaints and reason for seeking treatment, information and results about diagnostic examinations, outcomes to treatment and interventions, adverse events, and discharge summary information. These records are archived for future reference, such as the patients' next hospital visit and are increasingly being computerised. It has been estimated that the average patient generates up to 50,000 data items during their life (Weed, 1989) and computers are able to assist with the management of this information. Individual health care organisations generally have electronic patient management systems that capture demographic and sometimes clinical data and provide for appointment scheduling. These systems may link to other information systems, such as those holding laboratory and radiology data. In primary care, patient management systems are often advanced in their clinical applications and provide the general practice with many of the tools required to manage a busy practice and significant patient workloads. In New Zealand at least 85% of GPs now use computers, some for administration purposes only but many use them to record all clinical notes, prescribing and laboratory test ordering and diagnosis coding (Gribben, Coster, Pringle, & Simon, 2001). However, most of these systems do not integrate in any way to systems outside the individual general practice. A clinical data repository is an aggregate of all computer based clinical data for patients. Electronic records support the integration of care across different health care providers by:

- supporting the capture of and access to a patient's outpatient clinical data as he/she moves through the system;
- capturing data for quality measurement and research;
- supporting best care practices by reminding physicians to perform certain tests and complete documentation (clinical decision support);
- structuring data collection to enhance the quality of care by guiding, for example, regular testing of certain parameters for patients with diabetes (Glaser, 2002).

2.3 Electronic Health Records

The 'electronic health record' (EHR) is increasingly being used to deliver health information to the point of care. Considerable work is now underway in many countries to develop the components of effective and comprehensive EHRs, and EHR development is central to many national health information strategies (Orfanidis, Bamidis, & Eaglestone, 2004). EHRs consist of components that enable health care providers to access a patient's health information regardless of geographical location. Other components refer patients to various providers and enable providers to access clinical decision support and electronic prescribing, physician order entry, and integrated communication with laboratories, imaging centres, colleagues and patients; and would include population health management with the overall goal of assisting providers to give better quality health care (Kerr, 2004). EHRs could simply be software applications that provide integrated, longitudinal views of patient data; providing several software applications in a networked environment. The Australian Health Online project provides a definition of an EHR that includes all the components of a comprehensive EHR model: *An electronic health record is an electronic longitudinal collection of personal health information, usually based on the individual, entered or accepted by health care providers, which can be distributed over a number of sites or aggregated at a particular source. The information is organised primarily to support continuing, efficient and quality health care. The record is under the control of the consumer and is to be stored and transmitted securely.* (Australian Department of Health and Aging, 2003)

Orfanidis (Orfanidis et al., 2004) found evidence that gaining the acceptance of EHR systems by physicians, nurses and other health care professionals is often problematic. A primary reason cited is a general lack of trust in EHR systems with respect to the critical issue of data quality, since possible errors may cause serious problems to patient health. A further barrier to effective understanding and use is the poor computer literacy of some health care personnel. Also, EHR systems can be perceived as a distraction, which impedes the physician's focus on the patient.

It may be that health care is missing out on the potential benefits of the EHR through these perceived data quality issues. There is a building body of evidence that suggests capturing and providing health care data in an electronic form improves the safety and quality of care provided and reduces the clinician's time spent on administrative tasks (Menke, Broner, Campbell, McKissick, & Edwards, 2001), (Patel, Kushniruk, Yang, & Yale, 2000), (Bates, 2002), (Ball & Douglas, 2002), (Institute of Medicine, 2000). In particular, components of the EHR, such as electronic prescribing, reduce likelihood of human error with as much as a 55% reduction in serious medication error rates (Bates, 2002). This is achieved through changing the 'systems of care' rather than merely computerising the existing workflow.

2.4 Clinical Coding

In clinical coding a short and formal code is recorded against clinical data to enable more accurate data analysis, making clinical data shorter and easier to record, but restricting the variability of expression. A coding system is defined as a documentary language that is based on a concept system (Shortliffe & Barnett, 2000). A concept system is a system of labels that are defined in an ordered way. This means, therefore, operations can be divided into operations on the nervous system, chest or blood vessels, then further divided into operations on the chest that can be broken down into operations on the heart, lung etc, ending up with further distinctions such as valvotomy of the aortic valve (Liener et al., 2003), (Shortliffe & Barnett, 2000). The health status and medical condition of a patient is difficult to describe using comprehensive, coded data. Factors that contribute to this problem include the following:

- accepted methods for formally decomposing many key components of the patient record, such as admission and physical history, into coded concepts have not yet been developed;
- where available, vocabularies to represent the terms within the model in a standard way are difficult to develop. The condition of the patient is often complex, probabilistic, and highly nuanced. Multi-factorial and temporal relationships can exist between pieces of data;
- there is no single way to organise automated medical data. There are often idiosyncratic ways developed to code data within and across organisations;
- data definitions vary across settings. These variations exist in problem lists, laboratory tests and procedures, for example;
- care processes vary, such as referrals, medication ordering, care documentation;
- definitions of optimal care vary, such as frequency of visits or appropriate tests;
- differing technologies, vendors, and levels of IT sophistication are found across settings, as well as different levels of computerisation (Glaser, 2002).

Therefore, establishing clinical meaning, measuring care, determining the health status of a patient or population, and developing clinical information systems that interoperate can be difficult. The complexity of boundaries compounds the complexities of the medical process and health data (Glaser, 2002), (Sanderson et al., 2004).

2.5 Privacy of Patient Data

Patients have a right to have personal health data kept private. The outcome of indiscreet record keeping could be that patients no longer disclose intimate details that are necessary for their proper medical care, as they do not trust the clinician to keep this information private. The patient owns his or her data and no one has the right to access this without the patient's authorisation. In the era of electronic communication, this is somewhat more complicated to ensure. To address the changing environment and the changing attitudes of health consumers, most OECD countries now have privacy regulations, with New Zealand being one of the founding developers of such regulations. The Privacy Act of 1993 has provided guidelines for the collection, use and storage of health information in New Zealand.

In the United States, HIPPA, or the Health Insurance Portability and Accountability Act of 1996, has further developed the privacy regulations for health data in which the patient is identifiable. The regulations refute the view of computer experts that personal control of private information is an illusion in the computer age and require that all patients receive a notice of privacy practices around their health care record. Privacy regulation needs to take into account legitimate users of the data for public good. These include:

- health care providers;
- patients;
- health care organisations;
- research institutions;
- public health departments;

- insurance companies;
- law enforcement;
- accrediting bodies (such as for health care organisations) (Saba & McCormick, 2001).

Adequate protection of health information depends on both technological and organisational policies.

2.6 Stewardship of Health Care Data

Stewardship involves taking responsibility for the good care of a set of resources on behalf of others (Ministry of Health, 2004). This involves:

- protection and development of the resource;
- ensuring management of the resource in a sustainable manner; and
- representing the interests of stakeholders in the management of that resource.

When multiple agencies, organisations or individuals share common needs and interests in a set of resources, there is a need to have input and representation from these group(s) in the management of the data. Stewardship arrangements should provide processes and mechanisms for these groups to act collectively in the interest of the effective and sustainable management of data (Ministry of Health, 2004).

Without appropriate and effective stewardship arrangements it is not clear how stakeholder input is gained, or who is responsible and accountable for defining requirements. Both duplication of activities and gaps in activities can occur where no-one assumes responsibility for decisions (Seiner, 1997). The result is poor definition and understanding of stakeholder needs and the development and use of resources is less than ideal (Loshin, 2004). Health strategies are key drivers of national collection establishment and change. Stewardship guides the alignment of national collections within the health strategy, but has no role in defining those strategies. Stewardship also recognises consumer interest in data, and provides opportunity for consumer representation in the stewardship processes (Ministry of Health, 2004). The key design principles to achieve effective stewardship are:

- stakeholder representation – stewardship relies on the involvement of stakeholder groups including both the providers of data and the users of that data;
- link to organisation directions and priorities – the underlying driver of all stewardship actions is support of the organisation's strategies and priorities;
- structure alignment with organisation culture – in order to be effective the stewardship roles and structures must align with existing organisational culture and accountability arrangements;
- management commitment – stewardship must be sponsored and supported by senior management in all the organisations affected;
- sustainability – any implementation of stewardship needs to be as practical and streamlined as possible, so that the role is sustainable (Ministry of Health, 2004).

2.7 Health Informatics

Health informatics has emerged as a discipline to develop appropriate computer systems to assist with the management of health care information. The World Health Organisation defines health informatics as: *the combination of technology and methodology which makes possible the computer-assisted collection, storage, processing, retrieval, distribution and management of health information* (World Health Organisation, 1988). Ranson (in (Hovenga et al., 1996)) notes that $28 billion could be saved in health care costs through the efficient use of telecommunications and the movement of patient management information. Health care costs are increasing; in western countries this ranges from about 6% to more than 12% of gross domestic product (GDP) (van Bemmel & Musen, 1997). Now that technology has progressed to the point where all types of data, with the exception of smell, may be produced in digital form it has become feasible to develop fully integrated health information systems (Hovenga et al., 1996). However, most of the evidence needed to make more informed decisions from the clinician at the bedside up to the formulation of national strategy remains embedded in fragmented, irretrievable, and often illegible paper-based patient records (Bates, 2002). Consumers are now coming to expect good quality, co-ordinated care and interoperable information systems provide the communication means required to deliver this kind of care (Rigby, 1998).

However, as Effken (Effken, 2002) notes, health care systems are complex sociotechnical systems in which many information system innovations fail because of problems in planning or design. Designers with an understanding of this complexity are required to provide appropriate analytical approaches to development. Ribiere, LaSalles, Khorramshahgol and Gousty (Ribiere, LaSalle, Khorramshahgol, & Gousty, 1999) found, when researching hospital information systems quality, that there is a significant lag behind other industries in the health care sector's ability to

deliver good quality information. Further, at present there are no current accreditation or regulatory requirements for health care software or hardware to provide quality assurance around such things as the privacy of health information through the appropriateness of design and integrity of systems (Rigby, 1998).

Health informatics research, in its position across the disciplines of health care and information systems, provides for the significant developments needed to manage health care in the future. Health informatics research may enable the development of systems that increase flow of information from the health care recipient and the provider to various organisations for fund and resource allocation and for management.

3.0 WHY IMPROVE DATA QUALITY IN HEALTH CARE?

Examples of the implications of poor quality data in health care provide considerable impetus for professionals and managers to ensure data quality management is a priority. Single errors can have significant effects on patients and are highlighted by the examples below:

- errors in a reference database calculation of Down's Syndrome screening, giving false negatives (Rigby, 1998);
- an age cohort of women omitted from call up for cervical screening (Rigby, 1998);
- duplicate unique identifiers leading to patient clinical data being sent to the wrong patient.

Improving data quality in the health care industry is linked to better overall management of health plans (Henderson in (Leitheiser, 2001)). Further, health care delivery has been under scrutiny recently following an extensive review of the quality of care provided in the United States (Institute of Medicine, 2000). Error in care provision has been found to cause significant 'adverse events', where the patient suffers an injury resulting from medical intervention. The report estimates that between 44,000 and 98,000 people die each year in the United States as a result of medical errors. This caused 'shock waves' around the western world as other nations calculated their possible death rates. There is now considerable interest in health care to improve the safety of patients through the improvement of the system of delivering care (Hagland, 2003). Whilst human error is considered a factor, often this human error is a direct result of a system or process failure (Institute of Medicine, 2000).

The practice of evidence-based medicine requires access to significant research data, collated and presented in a way that the clinician can use at the time of decision making. There are now overwhelming amounts of health research results available and continually being published, so much so that no clinician could ever remain completely up to date. In order to provide clinicians with this information, organisations such as the Cochrane Collaboration[1], an international non-profit and independent organisation, provide up-to-date information about the effects of health care readily available worldwide, produce and disseminate systematic reviews of health care interventions and promote the search for evidence in the form of clinical trials and other studies of interventions. Guidelines Groups throughout the world develop evidence-based guidelines that assess policy, purchasing and practise options, and make recommendations for best practice. If appropriately implemented, an evidence-based approach will improve quality and outcomes by introducing effective care and services, while reducing unnecessary costs and ineffective care. Effectiveness takes into account:

- international and national clinical research-based evidence;
- improved patient/consumer outcomes (i.e. patient safety) and satisfaction;
- service delivery improvements;
- rational resource use.

Many of these guidelines are integrated into electronic clinical decision support tools (Institute of Medicine, 2000) and linked to the patient's EHR. By providing access to interactive clinical decision support tools, the clinician, for example, can be alerted to potential adverse drug reactions when prescribing for a particular patient. These sophisticated tools provide information on the latest research, applicable to the patient in question i.e. what co-morbidities and current drugs is the patient on that would make the drug being prescribed unsafe? Data of high quality are required for such tools to be effective.

3.1 Improving Data Quality in Health Care

The health care sector's widely decentralised and largely autonomous data collection efforts make data quality a significant challenge (Berndt, Fisher, Hevner, & Studnicki, 2001). Attempts have been made for some time to manage and control the quality of health data including data within health-based information systems. For example, Article 6 of the European Union data protection directive (Council of the European Union, 1995) relates specifically to data quality and observes that data must be:

---

[1] www.cochrane.org.nz

- processed fairly and lawfully;
- collected for specified, explicit and legitimate purposes and not further processed in a way incompatible with those purposes;
- adequate, relevant and not excessive;
- accurate and, where necessary, kept up to date;
- kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the data were collected or for which they are further processed.

Given the complexity of health care and the data required, as outlined above, it is not surprising that improving data quality in health care is a complex issue. The multiple uses of the same data provide considerable difficulty in ensuring and measuring for adequate data quality in health care. Olson (Olson, 2003) talks about the 'principles if unintended uses', whereby you cannot separate data from its uses. To assess data quality you must undertake a thorough requirements analysis and design an assessment tool that meets these requirements. But often data collections remain in place and the environment changes. Considerable change to the management of health care has taken place over the last 15 years, particularly in New Zealand, as well as changes to legislation and the move towards evidence-based medicine. Many data collections and/or registries continue to collect data in the same format they did when developed many years previously and the use of the data has expanded into ways not thought of at the time of the database's conception. Without good metadata, the understanding of the context of data collection is likely to be unclear (Olson, 2003).

All medical records and registries are necessarily incomplete because they reflect the selective collection and recording of data, known as the hypothetico-deductive approach (Shortliffe & Barnett, 2000). This approach is one of sequential, staged data collection, followed by data interpretation and the generation of hypotheses around possible diagnoses (differential diagnoses), leading to hypothesis directed selection of the next most appropriate data to be collected. Therefore, any data quality improvement in health care needs to consider all dimensions of data quality, for example, the relevance of missing data, or the timeliness of data depending on the context of its use. Data required by health care providers may not be required to be complete to be relevant to the delivery of care, but may be insufficient to meet the needs of government reporting or hospital management requirements. Clinician involvement in the validation and use of data has been found to be significant in ensuring and improving data quality (Croft & Williams, 2005). Whilst some data collections are managed with teams that include doctors, many do not involve nurses or allied health professionals (Black, Barker, & Payne, 2004).

A review of data quality management practices in US health care organisations (Lorence & Jameson, 2001) found that there was a technology lag and regional variation in practices, making it difficult to make comparisons with broader health sector data such as for evidence based health care. Where data quality assessment did occur within organisations, traditional audits were still preferred and system based audits were uncommon. Lorence and Jameson (Lorence & Jameson, 2001) feel the national comparison of health care data is difficult when there are no common standardised practices for data quality assessment.

3.2 Total Quality Management in Health Care

In response to the recent highlighting of poor quality patient safety (Institute of Medicine, 2000) and the drive for improved efficiency, many health care providers are implementing TQM practices to improve the organisational processes for care delivery and administrative management. For example, the National Health Service (NHS) in the United Kingdom, through the 'NHS Modernisation Programme[2]', has implemented considerable changes at the local, regional and national level to the culture of health care provision and management. Similar efforts are also taking place in Health Management Organisations in the United States. TDQM (Lee, Pipino, Strong, & Wang, 2004) aligns closely with this philosophy of providing a context for data management improvements within the overall management philosophy.

'Clinical care pathways' are an example of quality management practices being implemented in many health care organisations. The pathways are developed through the analysis of evidence to ascertain the best health outcomes for patients following a care pathway devised for a specific diagnosis. The data obtained when documenting the patient's progress or 'pathway' through the devised treatment plan provide health care service planners with clear expected outcomes, for example, average length of stay and cost of treatment. Patient data that vary from the expected outcomes are analysed to assess the reason for the variation. The variation may be unavoidable co-morbidities, related or not to the diagnosis or treatment. Data may also provide valuable information on the quality of care provided or

---

[2] http://www.wise.nhs.uk/cmsWISW/default.htm

clinical decisions made. Aggregated data from several patients on the same clinical pathway may provide information on errors in processes within the organisation that consistently cause variation from the clinical pathway expected outcomes.

Throughout the data collection and analysis process, the quality of data is paramount to providing the clinician and the organisation with a true picture of the quality of care provided. With the movement towards more customer focused health care provision and increasing consumer responsibility for health, data related to patient outcomes is sometimes being made public now (Carey & Lloyd, 2001). This enables health care consumers to make decisions about which health care organisation and even which clinician to choose when requiring treatment, giving clinicians greater impetus to collect good quality data, or at the very least to consider these data more carefully. Given the considerable importance of high quality data to the management of health care, any TQM programme would by necessity include data quality improvement. TDQM provides a structure to prioritise data quality improvement in a complex environment where cost containment can considerably restrict the choices available for improvement projects.

3.3 Improving Data Quality in the EHR

Given the considerable interest in the development and implementation of EHRs, it is not surprising that studies are being undertaken into ensuring and improving data quality in the EHR. De Lusignan, Stephens, Adal and Majeed, and Fine, Keogh, Cretin, Orlando and Guild, and Lorenzoni, Da Cas and Aparo (De Lusignan, Stephens, Adal, & Majeed, 2002; Fine, Keogh, Cretin, Orlando, & Gould, 2003; Lorenzoni, Da Cas, & Aparo, 1999) all found that continuous monitoring and feedback comparing a clinician's performance with other clinicians proved to be successful in improving the quality of data on medical records. More detailed feedback appears to have a greater effect (Gribben et al., 2001). This is supported by the findings of the UK Audit Commission (UK Audit Commission, 2004) who noted data quality in health care was hampered by a lack of understanding of the importance of the role of the data collector to the downstream quality and the usefulness of the data.

Data capture, manipulation and retrieval at the locations where they are generated and/or utilised is also necessary if an EHR system is to be integrated into the environments within which it is used and into the working practices of its users (Haan et al., 2004). Mobile technology (palmtop, pocket PC, mobile phones, etc.) is the key to achieving this (Orfanidis et al., 2004). Preventing the need to write data on paper based forms to be entered later means data are timelier and input errors are less likely to occur. Clear data definitions are also required to ensure all are interpreting fields correctly (Thiru, Hassey, & Sullivan, 2003).

Giving patients access to their records is a key part of EHR management strategies. For example, the NHS plans that every person in the United Kingdom will have his or her own record. Patient access to records is important, as it allows them to take control over their own health and make decisions about their health care, and also improves data quality (NHS Information Authority, 2003). In particular, this provides patients with chronic illnesses the opportunity to get involved with the management and treatment of long-term conditions (Orfanidis et al., 2004). A research project that allowed parents of paediatric patients to view and change their child's records found 'improved sensitivity for detection of historical risk factors for illness can be achieved by augmenting the paediatric EHR with a section for direct parental data input' and significantly improved data quality (Porter & Mandl, 1999).

However, a study in New Zealand revealed that in general the public are unaware that health information is held about them and that they can view and make amendments to that information (Ryan, 2004), indicating considerable public awareness education is required if patients are to have understanding about input into their medical records. Orfandis et al (Orfanidis et al., 2004) and Dugas, Hoffman, Janko, Hahewald, Matis and Uberla (Dugas, Hoffman, Janko, Hahnewald, Matis, & Uberla, 2001) in their studies of data quality in EHRs, found that analysis and design of data quality issues are an integral part of the development of an EHR and should be addressed from the inception of the project, but that there are currently technological limitations to some of the issues, such as storage requirements, which still need to be resolved.

3.4 Clinical Research

Clinical research often utilises the results of data collected from international clinical trials, for example from double-blind drug trials using 'gold standard' research methodologies, usually with many thousands of participants from several countries taking part over 3-5 years. Increasingly, high quality data are expected by authorities and consumers before new drugs or new uses for existing drugs are approved. These trials are now utilising the World Wide Web to enable investigators to directly input clinical trial data at the point of observation. Marks, Conlon and Ruberg (Marks, Conlon, & Ruberg, 2001) found the accuracy and timeliness of the data improved considerably with a simplified data collection process when using the web to collect clinical data. Arts et al (Arts et al., 2002) also note the requirement for good study design, with extensive training for data collectors to ensure the prevention of errors.

In addition detection through routine monitoring and action on errors is required throughout the study. Measures need to be taken at the local data collection source and standard definitions of data quality and data quality attributes are necessary to be able to compare data quality among research registries or within a registry at different points in time (Thiru et al., 2003).

3.5 Health Information Standards

Within health care systems data can be documented in many different ways. A particular level of detail may be chosen; synonymous labels may be selected, spelt differently and structured according to the author's preference. This causes considerable problems in subsequent data analysis. Retrieval is harder and less reliable; the use of homonymous terms may lead to the selection of irrelevant data objects when searching data collections. It may be nearly impossible to count the frequency of certain similar objects (such as diagnostic categories for administrative or scientific purposes), because the terms used do not indicate the degree of similarity between them (Liener et al., 2003).

Standards facilitate the sharing and exchange of information between departments, health agencies and health workers (Hovenga et al., 1996), providing a defined set of rules, conditions or requirements concerning definition of terms, classification of components, specifications of materials, performance of operations, and delineation of procedures. International and national standards developments are providing the benchmarks for data access and storage. Health Level 7 (HL7) is emerging as a significant international health data and messaging standard that includes the development of a standard format for the content of the EHR.

Providing international standards allows for the movement of health data across any geographical boundary. Murray, Lopez and Wibulpolprasert (Murray, Lopez, & Wibulpolprasert, 2004) note the importance of internationally agreed data standards to enable improved global health monitoring. The increasing importance of monitoring epidemics globally was recently highlighted by the SARS virus epidemic and its spread to distant continents within a very short space of time requiring appropriate international responses.

The benefits of standards development in improving data quality are considerable. The ability to transfer data throughout the health sector is greatly increased (Kenney & Macfarlane, 1999), reducing the requirement for subsequent mapping of data to different coding or classification standards when reaching its destination (Shortliffe & Barnett, 2000). All who send and receive data have a consistent understanding of its meaning. Agreed standards can then be included into a 'metadata' repository, providing the format for storage of metadata which may be applicable within the organisation, the region or nationally.

4.0 NATIONAL DATA QUALITY IMPROVEMENT STRATEGIES IN HEALTH CARE

Government departments of health are beginning to recognise the need for national strategies that provide for the consistent improvement of data across the national health sector. A review of international data quality improvement strategies in health care, including the NHS in the United Kingdom (United Kingdom Department of Health, 2004), the Canadian Institute for Health Information (CIHI) (Canadian Institute for Health Information, 2003b), Health*Connect* Australia (Australian Department of Health and Aging, 2003) and the United States Department of Health and Human Services (United States Department of Health and Human Services, 2002) identifies similarities between the strategies. All the reviewed international strategies note the complexity of data quality improvement in their multi-level, multi-dimensional initiatives. Principles are based on the need to manage data proactively to ensure ongoing integrity by preventing data quality problems, using a systematic TDQM approach. There is also commonality of role expectations – the data suppliers are responsible for the quality of the data they provide to central government, while central government is required to provide leadership and assistance to data suppliers by developing sector-wide standards and best practice guidelines. A common theme underlying many of their quality issues has been previous low levels of investment in information technology in health care.

The UK National Health Service Information Authority (NHSIA) views data quality as an intrinsic part of the Information for Health Strategy (NHS Information Authority, 1998). Data quality is seen as critical to ensure the accuracy and timeliness of data that 'supports information used for improving health and more effective management' (NHS Information Authority, 1998). A vital aspect of the UK's strategy revolves around the ability of the health professional to access consistently recorded and coded clinical information. Standards are clearly defined and applied to ensure consistency such that comparisons can be made both across time and between data sources. Despite data quality underpinning the strategic use of information in health, the UK approach to data quality is focused on data entry and consistency of initial coding. There is little consideration of data abstraction, relying instead on the extension of the World Health Organisation's ICD-10 coding classification through the use of internationally agreed metadata standards. ICD-10 has also been adopted as a standard coding model used in Australia (Australian Department of Health and Aging, 2003) and New Zealand.

The NHSIA (NHS Information Authority, 1998) sees high quality data being produced as part of the routine daily activity within a hospital or general practice surgery. The NHSIA data quality framework is then seen as providing confidence in their findings, providing necessary information to make relevant changes and achieve continuous quality improvement. The framework outlines clearly, and in detail, substantial work that is required by health care providers to ensure good data quality. The NHS developed an accreditation scheme that was initially thought to be all that would be required to ensure the supply of good quality data. The scheme is extensive and was found to be very successful but did not sufficiently identify the responsibilities of the data supplier; central government was still monitoring more than it was leading. This led to the more extensive guideline development on the principles and theories of data quality supported within the NHS. Several NHS Trusts have developed and published, on their websites, data quality strategies that align to the central NHS core strategy requirements.

A recent review of the programme by the UK Audit Commission (UK Audit Commission, 2004) found significant improvements to levels of data quality in the NHS. However similar issues are still apparent after five years of targeted improvements and the report recommended:

- developing a more co-ordinated and strategic approach to data quality;
- developing a NHS wide strategy for specifying, obtaining and using both national and local information;
- making more and better use of patient-based information;
- involving NHS Trust board members;
- training and developing staff;
- keeping systems up to date.

The CIHI also briefly discussed accreditation for enhancing collaboration with data suppliers (Canadian Institute for Health Information, 2003b). They have undertaken extensive work on data quality through collaborative work with experienced statisticians from Statistics Canada. They base their theories on research by the Massachusetts Institute of Technology in the United States. The CIHI Data Quality Strategy focuses on the following areas:

- fostering a data quality culture within CIHI and the health sector;
- strengthening data quality infrastructure and capacity;
- cultivating the data supply chain;
- enhancing external collaboration;
- establishing dedicated funds for fast track priority projects;
- communicating strategies and action plans, importance of data quality, opportunities for collaborative action and progress achieved.

The Health *Connect* model (Australian Department of Health and Aging, 2003) is essentially the requirements for the receipt, storage and retrieval of clinical and demographic data. It is recognised that in order to be a success, Health *Connect* will address a number of key data issues such as:

- validity and integrity of the data;
- completeness of data;
- responsibility for data
- controlling access to data;
- errors in data;
- corrections to data;
- data consistency and standardisation;
- data definitions and descriptors;
- ownership of data.

The series of projects related to Health *Connect* are still in the pilot phase. The researcher was unable to find any evidence of the development of a specific strategic data quality plan for ensuring good data quality for the Health *Connect* model. The Department of Health and Human Services in the United States (United States Department of Health and Human Services, 2002) disseminated very high level guidelines to those who manage health care data, no specific data quality strategy was found by the researcher. The guidelines pertained to the quality for information disseminated to the public and all health care providers were requested to develop their own guidelines by the Office of Management and Budget (OMB). The guidelines noted that information should be accurate, reliable, clear,

complete, unbiased, and useful. Guidelines were provided on the policies and practices that support information dissemination. Data quality is to be treated as an integral step in every aspect of the information development process, but organisations also need to weigh the costs and benefits of quality. Organisations should:

- make data management methods transparent by providing documentation;
- ensure quality by reviewing methods, consulting with experts and users, notifying users about corrections and revisions;
- make data and information supported with public funds available to the public consistent with confidentiality concerns and resource availability;
- review all data before dissemination.

5.0 CONCLUSION
Data quality is now becoming an increasingly important issue for health care providers, managers and government departments. The movement towards total quality management in health care to improve patient safety and health care efficiency is demanding high quality information. Further, evidenced-based care requires the assimilation of large amounts of relevant research data to be available at the point of clinical decision making.

The increasing movement of data through electronic innovations is highlighting the contextual nature of data in health care. There is a need for good metadata regarding the context of its collection to ensure applicable use of the data and to prolong the usefulness of the data. The development and implementation of national EHRs is providing clinical decision support and integrated care that demands high levels of data quality. Consumers themselves are demanding greater access to their health data. Whilst this is likely to improve the quality of the data it also places greater demands on the data manager to maintain high quality data to ensure consumer confidence in their health care provider.

Methods engaged outside of health care are proving valuable in improving health care data. Many of the issues found in other industries are found in health care, but the complexity of health care provision means that data quality improvement requires a multi-faceted approach. Strategic prevention, national consistency of improvement practices and data standards, and targeted improvements with increasing consumer involvement are moving health care towards a TDQM approach to data quality management.

Much of the current literature on data quality in health care focuses on data quality improvement for single data collections or one-off projects for multiple collections. Further academic research is required to elicit the theoretical underpinnings on which to understand the needs of health care in the management of data quality on a systemic and strategic level. This will enable increasing data mining and assessment of data across not only multiple organisations, but also nationally and internationally.

It can be said then, that ensuring sufficient data quality in health care can impact over all individual health care management and decision making in the areas of:
- national and regional service delivery decision making, impacting on health economics;
- patient safety whilst undergoing care;
- evidence to support clinical decision making through health care research;
- information provided to patients on their illness and care;
- effectiveness of clinical care pathways;
- movement of data across government departments to provide a holistic view of health care requirements.

References

Abate, M. L., Diegert, K. V., & Allen, H. W. (1998). *A hierarchical approach to improving data quality.* Retrieved 6.08.06, from http://www.dataquality.com/998abate.htm

Adelman, S. (2001). *Data Strategy Introduction*. Retrieved 18.8.06, from http://www.dmreview.com/article_sub.cfm?articleID=4291

Agmon, N., & Ahituv, N. (1987). Assessing data reliability in an information system. *Journal of Management Information Systems, 4*(2), 34-44.

Alshawi, S., Missi, F., & Eldabi, T. (2003). Healthcare information management: the integration of patient's data. *Logistics Information Management, 16*(3/4), 286-295.

Al-Shorbaji, N. (2001). *Health and medical informatics*. Cairo: World Health Organisation.

Argyris, C., & Schön, D. (1978). *Organisational Learning: A Theory of Action Perspective*. Reading: Addison-Wesley.

Arts, D. G. T., De Keizer, N. F., & Scheffer, G. J. (2002). Defining and improving data quality in medical registries: a literature review, case study, and generic framework. *Journal of the American Medical Informatics Association, 9*, 600-611.

Atkinson, C., Eldabi, T., Paul, R. J., & Pouloudi, A. (2001). *Investigating integrated socio-technical approaches to health informatics.* Paper presented at the 34th Hawaii International Conference on System Sciences.

Australian Department of Health and Aging. (2003). *Health Connect*. Retrieved 18.10.05, from http://www.healthconnect.gov.au/index.htm

Avison, D., & Myers, M. (1995). Information systems and anthropology: an anthropological perspective on IT and organisational culture. *Information Technology & People, 8*(3), 43-56.

Ball, M., & Douglas, J. (2002). IT, patient safety and quality care. *Journal of healthcare Information Management, 16*(1), 28-33.

Ballou, D. P., & Pazer, H. (1995). Designing information systems to optimise the accuracy-timeliness tradeoff. *Information Systems Research, 6*(1), 51-72.

Ballou, D. P., & Pazer, H. (2003). Modelling completeness versus consistency tradeoffs in information decision contexts. *Transactions on Knowledge and Data Engineering, 15*(1), 240-243.

Ballou, D. P., & Tayi, G. K. (1989). Methodology for allocating resources for data quality enhancement. *Communications of the ACM, 32*(3), 320-329.

Ballou, D. P., & Tayi, G. K. (1999). Enhancing data quality in data warehouse environments. *Communications of the ACM, 42*(1), 73-78.

Ballou, D. P., Wang, R., Pazer, H., & Tayi, G. K. (1998). Modelling information manufacturing systems to determine information product quality. *Management Science, 44*(4), 462-484.

Baskerville, R. (1999). Tutorial: Investigating information systems with action research. *Communications of the Association for Information Systems, 2*(19).

Baskerville, R., & Myers, M. (2004). Special issue on action research in information systems: making IS research relevant to practice - foreword. *MIS Quarterly, 28*(3), 329-335.

Baskerville, R., & Wood-Harper, A. T. (1996). A critical perspective on action research as a method for Information Systems research. *Journal of Information Technology, 11*(3), 235-246.

Baskerville, R., & Wood-Harper, A. T. (1998). Diversity in information systems action research methods. *European Journal of Information Systems, 7*, 90-107.

Bates, D. W. (2002). The quality case for information technology in healthcare. *BMC Medical Informatics and Decision Making, 2*(7).

Bauer, K. (2004). *The power of metrics: KPIs - the metrics that drive performance management*. Retrieved 18.08.06, from http://www.dmreview.com/article_sub.cfm?articleID=1009207

Berndt, D. J., Fisher, J. W., Hevner, A. R., & Studnicki, J. (2001). Health care data warehousing and quality assurance. *Computer, 34*(12), 56-65.

Black, N., Barker, M., & Payne, M. (2004). Cross sectional survey of multicentre clinical database in the United Kingdom. *British Medical Journal, 328*, 1478-.

Bowen, P., Fuhrer, D. A., & Guess, F. M. (1998). Continuously improving data quality in persistent databases. *Data Quality, 4*(1).

Bowers, B. (1988). Grounded Theory. In B. Sarter (Ed.), *Paths to Knowledge: Innovative Research Methods for Nursing* (pp. 33-59). New York: National League for Nursing.

Broder, C. (2004). *Case Study: Data repository gets to the heart of quality improvement*. Retrieved 18.08.06, from www.ihealthbeat.org/index.cfm?Action=dspItem&itemID=103350

Bryant, A. (2002). *Grounding systems research: re-establishing grounded theory.* Paper presented at the 35th Hawaii International Conference on System Sciences.

Canadian Institute for Health Information. (2003a). *Data Quality Framework.* Retrieved 18.8.06, from http://secure.cihi.ca/cihiweb/dispPage.jsp?cw_page=quality_e

Canadian Institute for Health Information. (2003b). *Earning Trust. Key Findings and Proposed Action Plan from the Data Quality Strategies Study*.

Carey, R. G., & Lloyd, R. C. (2001). *Measuring Quality Improvement in Healthcare. A Guide to Statistical Process Control Applications*. Milwaukee: Quality Press.

Carson, C. S. (2000). *What is data quality? A distillation of experience*. Retrieved 18.08.06, from http://www.thecre.com/pdf/imf.pdf

Carson, C. S. (2001). *Toward a framework for assessing data quality*. Retrieved 18.08.06, from http://www.imf.org/external/pubs/cat/longres.cfm?sk=3991.0

Champagne, F. (2002). *The ability to manage change in health care organisations*: Commission of the Future of Health Care in Canada.

Chengalur-Smith, I., Ballou, D. P., & Pazer, H. (1999). The impact of data quality information on decision making: an exploratory analysis. *IEE Transactions on Knowledge and Data Engineering, 11*(6), 853-864.

Chung, W. Y., Fisher, C., & Wang, R. (2002). *What skills matter in data quality?* Paper presented at the Seventh International Conference on Information Quality, Boston.

Codd, E. F. (1986). *An evaluation scheme for database management systems that are claimed to be relational.* Paper presented at the Second International Conference on Data Engineering, Los Angeles.

Committee on Quality of Health Care in America. (2000). *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington DC: National Academy Press.

Council of the European Union. (1995). *Article 6. European Union Data Protection Directive*. Retrieved 3.11.05, from http://www2.warwick.ac.uk/fac/soc/law/elj/jilt/1996_1/special/directive/#a2.1

Croft, G., & Williams, J. (2005). The RCP Information Laboratory (iLab): breaking the cycle of poor data quality. *Clinical Medicine, 15*(1), 47-49.

Crosby, P. (1980). *Quality is Free*. New York: Penguin Group.

Cunningham, N. (2001). A model for the role of information systems in strategic change within health care organisations. *Organisation Development Journal, 19*(1), 93-108.

Davenport, T. H. (1998). *Working Knowledge: How Organisations Manage What They Know*. Boston: Harvard University.

Davidson, A. J., Melinkovich, P., Beatty, B., Chandramouli, V., Hambridge, S. J., Phibbs, S. L., et al. (2003). Immunisation registry accuracy. Improvement with progressive clinical application. *American Journal of Preventive Medicine, 24*(3), 276-280.

Davidson, B. (2004). Developing data production maps: meeting patient discharge submission requirements. *International Journal of Healthcare Technology and Management, 6*(2), 223-240.

Davidson, C., & Voss, P. (2002). *Knowledge Management*. Auckland: Tandem.

Dravis, F. (2003). *Data Quality Strategy*. Retrieved 14.8.06, from http://www.firstlogic.com/_gics10643/DQ/article.asp?articleID=32

Dravis, F. (2004a). *Data quality strategy: A step by step approach.* Paper presented at the Ninth International Conference on Information Quality, Boston.

Dravis, F. (2004b). *IQ Maturity, it will be smaller than your shoe size*. Retrieved 18.08.06, from http://www.firstlogic.com/resourcelibrary/bestpractices/iqview/iqview0104.asp

De Lusignan, S., Stephens, P. N., Adal, N., & Majeed, A. (2002). Does feedback improve the quality of computerised medical records in primary care? *Journal of the American Medical Informatics Association, 9*, 395-401.

Deming, W. E. (1982). *Out of the Crisis*. Cambridge: Massachusetts Institute of Technology.

Detmer, D. E. (2003). Building the national health information infrastructure for personal health, health care services, public health, and research. *BMC Medical Informatics and Decision Making, 3*(1).

Dey, I. (1993). *Qualitative Data Analysis. A User Friendly Guide for Social Scientists*. London: Routledge.

Dubois, L. (2003). *What Works: Curing the blown budget in data consolidation*. Retrieved 11.07.05, from www.tdwi.org/research/displayaspx?ID=6877

Dugas, M., Hoffman, E., Janko, S., Hahnewald, S., Matis, T., & Uberla, K. (2001). *XML-based visual data mining in medicine.* Paper presented at the MEDINFO 2001, Amsterdam.

Eckerson, W. W. (2002). *Data Quality and the Bottom Line*. Retrieved 26.08.06, from http://www.adtmag.com/article.asp?id=6321

Effken, J. (2002). Different lenses, improved outcomes: a new approach to the analysis and design of health care information systems. *International Journal of Medical Informatics, 65*, 59-74.

Elrod, P. D. (2002). The death valley of change. *Journal of Organisational Change, 15*(3), 273-291.

English, L. (1998, May 1998). *Plain English on data quality: Data quality point 6: Institute training on data quality*. Retrieved 23.08.06, from http://www.dmreview.com/master.cfm?NavID=198&EdID=619

English, L. (1999a). *Improving Data Warehouse and Business Information Quality*. New York: John Wiley & Sons.

English, L. (1999b, October 1999). *Information quality in the knowledge age*. Retrieved 21.08.06, from http://www.dmreview.com/master.cfm?NavID=198&EdID=1452

English, L. (1999c, March 1999). *Plain English on data quality: Data quality point 13: Education and self-improvement*. Retrieved 21.08.06, from http://www.dmreview.com/master.cfm?NavID=198&EdID=117

English, L. (2003, June 9, 2003). *Plain English about information quality: How to save $576 925 000 through IQ management*. Retrieved 21.08.06, from http://www.dmreview.com/article_sub.cfm?articleId=6823

Eppler, M. J., & Wittig, D. (2000). *A Review of Information Quality Frameworks from the Last Ten Years*. Retrieved 14.08.06, from http://www.knowledgemedia.org/modules/pub/view.php?id=knowledgemedia-2&page=1

Fine, L. G., Keogh, B. E., Cretin, S., Orlando, M., & Gould, M. M. (2003). How to evaluate and improve the quality and credibility of an outcomes database: validation and feedback study on the UK cardiac surgery experience. *British Medical Journal, 326*, 25-28.

Fisher, C., Chengalur-Smith, I., & Ballou, D. P. (2003). The impact of experience and time on the use of data quality information in decision making. *Information Systems Research, 14*(2), 170-188.

Fletcher, D. (2004). Achieving data quality: new data from a paediatric health information system earns trust of its users. *Journal of AHIMA, 75*(10), 22-26.

Fox, C. J., & Levitin, A. V. (1996). Data and Data Quality. *Encyclopaedia of Library and Information Science, 57*(Supplement 20), 100-122.

Fry, R. (2002). *Appreciating your past in order to transform the future*. Retrieved 26.9.05, from http://www.clevelandshrm.com/

Fuller, B., & Redman, T. C. (1994). *Data Quality lessons learned at Telecom Australia.* Paper presented at the QUALCON 94 Conference, Melbourne.

Gendron, M. S., & D'Onofrio, M. J. (2001). *Data Quality in the Healthcare Industry*. Retrieved 21.08.06, from http://www.dataquality.com/901GD.htm

Glaser, B. G., & Strauss, A. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Chicago: Aldine Publishing.

Glaser, J. P. (2002). *The Strategic Application of Information Technology in Health Care Organisations* (2nd ed.). San Francisco: Jossey-Bass.

Gribben, B., Coster, G., Pringle, M., & Simon, J. (2001). Non-invasive methods for measuring data quality in general practice. *New Zealand Medical Journal, 114*, 30-32.

Haan, C., Adams, M., & Cook, R. (2004). Improving the quality of data in your database: lessons from a Cardiovascular Centre. *Joint Commission Journal on Quality and Safety, 30*(12), 681-688.

Hagland, M. (2003). Reduced errors ahead. *Healthcare Informatics*(August), 31-40.

Hamel, G., & Prahalad, C. K. (1994). *Competing for the Future. Breakthrough Strategies for Seizing Control of Your Industry and Creating the Markets of Tomorrow*. Boston: Harvard Business School Press.

Hovenga, E. J. S., Kidd, M. R., & Cesnik, B. (1996). *Health Informatics. An Overview*. Melbourne: Churchill Livingstone.

Huang, K.-T., Lee, Y. W., & Wang, R. (1999). *Quality Information and Knowledge*. New Jersey: Prentice-Hall.

Institute of Medicine. (2000). *To Err is Human. Building a Safer Health System*. Washington: National Academy Press.

Juran, J. M., & Godfrey, A. B. (1999). *Juran's Quality Handbook* (5 ed.). New York: McGraw-Hill.

Kamira, R. (2003). *Kaitiakitanga: Introducing useful indigenous concepts of governance.* Paper presented at the Health Informatics New Zealand Conference, Auckland.

Karr, A. F., Sanil, A. P., & Sacks, J. (2000). *Workshop Report, Affiliates Workshop on Data Quality*. Morristown, New Jersey: National Institute of Statistical Sciences.

Kenney, N., & Macfarlane, A. (1999). Identifying problems with data collection at a local level: survey of NHS maternity units in England. *British Medical Journal, 319*, 619-622.

Kerr, K. (2004). *The Electronic Health Record in New Zealand*. Retrieved 12.8.06, from
http://hcro.enigma.co.nz/website/index.cfm?fuseaction=articledisplay&FeatureID=040304

Kerr, K., & Norris, T. (2004). *The development of a healthcare data quality framework and strategy.* Paper presented at the Ninth International Conference on Information Quality, Boston.

Klein, B., & Rossin, D. F. (1999). Data errors in neural network and linear regression models: An experimental comparison. *Data Quality, 5*(1), 25.

Kmietowicz, Z. (2004). Data collection is poor because staff don't see the point. *British Medical Journal, 328*, 786.

Kovac, R., Lee, Y. W., & Pipino, L. L. (1997). *Total Data Quality Management: The case of IRI.* Paper presented at the 1997 Conference on Information Quality, Boston.

Langley, P. (1994). *Doing Social Research: a guide to coursework*. Lancashire: Causeway Press Ltd.

Lee, R. D., Conely, D. A., & Preikschat, A. (2000). eHealth 2000: health care and the internet in the new millennium. *WIT Capital Industry Report*.

Lee, Y. W. (2004). Crafting rules: context reflective data quality problem solving. *Journal of Management Information Systems, 20*(3), 92-119.

Lee, Y. W., Pipino, L. L., Strong, D. M., & Wang, R. Y. (2004). Process-embedded data integrity. *Journal of Database Management, 15*(1), 87-103.

Lee, Y. W., & Strong, D. M. (2003). Knowing why about data processes and data quality. *Journal of Management Information Systems, 20*(3), 13-39.

Lee, Y. W., Strong, D. M., Kahn, B. K., & Wang, R. Y. (2002). AIMQ: A methodology for information quality assessment. *Information and Management, 40*(2), 133-146.

Leitheiser, R. L. (2001, 2001). *Data Quality in Health Care Data Warehouse Environments.* Paper presented at the 34th Hawaii International Conference on System Sciences.

Levitin, A. V. (1998). Data as a resource: properties, implications, and prescriptions. *Sloan Management Review, 40*(1), 89-101.

Levitin, A. V., & Redman, T. C. (1993). A model of the data (life) cycles with application to quality. *Information and Software Technology, 35*(4), 217-223.

Lewin, K. (1951). *Field Theory in Social Science*. New York: Harper and Brothers.

Liener, F., Gaus, W., Haux, R., & Knaup-Gregori, P. (2003). *Medical Data Management*. New York: Springer-Verlag.

Long, J. A. (2004). *Where to start? A preliminary data quality checklist for emergency medical services data.* Paper presented at the 9th International Conference on Information Quality, Boston.

Long, J. A., & Seko, C. E. (2002). *A new method for database data quality evaluation at the Canadian Institute for Health Information.* Paper presented at the 7th International Conference on Information Quality, Boston.

Lorence, D. P., & Jameson, R. (2001). Adoption of information quality practices in US healthcare organisations. A national assessment. *International Journal of Quality and Reliability Management, 19*(6), 737-756.

Lorenzi, N., Riley, R., Blyth, A., Southon, G., & Dixon, B. (1997). Antecedents of the people and organisational aspects of medical informatics: review of the literature. *Journal of the American Medical Informatics Association, 4*(2), 79-93.

Lorenzoni, L., Da Cas, R., & Aparo, U. L. (1999). The quality of abstracting medical information from the medical record: the impact of training programmes. *International Journal of Quality in Health Care, 11*(3), 209-213.

Loshin, D. (2001). *Enterprise Knowledge Management. The Data Quality Approach*. California: Academic Press.

Loshin, D. (2003). *Knowledge Integrity: Assessment in the golden years*. Retrieved 23.9.05, from http://www.dmreview.com/article_sub.cfm?articleId=6803

Loshin, D. (2004). *Issues and opportunities in data quality management coordination*. Retrieved 11/04/05, from http://www.dmreview.com/article_sub.cfm?articleId=1000836

Lynes, D. A. (1998). *Analysis of power dialects occuring between health professionals and clients*. Lancaster: University of Lancaster.

Malcom, A. (1998). *Poor data quality costs 10% of revenues, survey reveals*. Retrieved 18.10.05, from http://computerworld.co.nz/news.nsf/UNID/CC256CED0016AD1ECC25684C000E0278?OpenDocument&Highlight=2,Poor,data,quality,costs

Marks, R. G., Conlon, M., & Ruberg, S. J. (2001). Paradigm shifts in clinical trials enabled by information technology. *Statistics in Medicine, 20*, 2683-2696.

McFadden, F. R., Hoffer, J. A., & Prescott, M. B. (1999). *Modern Database Management* (Fifth ed.). Reading: Addison-Wesley.

McGrath, G. M., & More, E. (2001). *Data integration along the health care supply chain: the Pharmaceutical Extranet Gateway Project.* Paper presented at the 34th Hawaii International Conference on Systems Sciences.

Menke, J., Broner, C., Campbell, D., McKissick, M., & Edwards, -. B., Joy. (2001). Computerised clinical documentation system in the pediatric intensive care unit. *DMC Medical Informatics and Decision Making, 1*(3).

Miles, M., & Huberman, M. (1994). *Qualitative Data Analysis. An Expanded Sourcebook* (2nd ed.). Thousand Oaks, California: Sage.

Miller, H. (1996). The multiple dimensions of information quality. *Information Systems Management, Spring*, 79-82.

Ministry of Health. (1996). *Health Information Strategy for the Year 2000*. Wellington.

Ministry of Health. (2000). *The New Zealand Health Strategy*. Wellington.

Ministry of Health. (2002). *He Korowai Oranga – Maori Health Strategy*. Wellington.

Ministry of Health. (2003a). *DHB CIO Health Sector Survey*. Wellington.

Ministry of Health. (2003b). *Health and Independence Report*. Wellington.

Ministry of Health. (2003c). *Ministry of Health ISSP Project*. Wellington.

Ministry of Health. (2004). *Draft National Collections Stewardship Framework*. Wellington.

Ministry of Health. (2005). *The Health Information Strategy for New Zealand*. Wellington.

Mintzberg, H. (1987). Crafting Strategy. *Harvard Business Review, 65*(4), 66-.

Mitchell, S., & Brown, A. (2002). *Performance Measurement Data: Is it fit for use? Data Quality Initiatives at CIHI*. Retrieved 23.9.05, from http://secure.cihi.ca/cihiweb/en/downloads/quality_perforamancepaper_e.pdf

Moody, D. L., & Shanks, G. G. (1998). What makes a good data model? A framework for evaluating and improving the quality of entity relationship models. *The Australian Computer Journal, 30*(3), 97-110.

Mumford, E. (2001). Action Research: Helping Organisations to Change. In E. M. Trauth (Ed.), *Qualitative Research in IS: Issues and Trends* (pp. 298). London: Idea Publishing Group.

Murray, C., Lopez, A., & Wibulpolprasert, S. (2004). Monitoring global health: time for new solutions. *British Medical Journal, 329*, 1096-1100.

New Zealand Health Information Service. (2003). *Current State Analysis of Data in the National Health Collections*. Wellington: Ministry of Health.

NHS Information Authority. (1998). *An Information Strategy for the Modern NHS. A national strategy for local implementation*. Retrieved 26.9.05, from http://www.nhsia.nhs.uk/def/pages/info4health/1.asp

NHS Information Authority. (2003). *Electronic record development and implementation programme. Lessons learned topic summary: Data quality*. Retrieved 3.11.05, from http://www.nhsia.nhs.uk/erdip/pages/publications/ERDIPDataQualitySummary_3.pdf

Nord, G. D., Nord, J. H., & Xu, H. (2005). An investigation of the impact of organisation size on data quality issues. *Journal of Database Management, 16*(3), 58-71.

Olson, J. E. (2003). *Data Quality. The Accuracy Dimension*. San Francisco: Morgan Kaufmann Publishers.

One Six Sigma Glossary. Retrieved 11.07.05, from http://www.onesixsigma.com/tools_resources/glossary_i.php

Orfanidis, L., Bamidis, P., & Eaglestone, B. (2004). Data quality issues in electronic health records: an adaption framework for the Greek health system. *Health Informatics Journal, 10*(1), 23-36.

Orr, K. (1998). Data quality and systems theory. *Communications of the ACM, 41*(2), 66-71.

Pandit, N. R. (1996). The creation of theory: A recent application of the grounded theory method. *The Qualitative Report, 2*(4), 1-14.

Patel, V., Kushniruk, A., Yang, S., & Yale, J.-F. (2000). Impact of a computer-based patient record system on data collection, knowledge, organisation and reasoning. *Journal of the American Medical Informatics Association, 7*(6), 569-585.

Pauleen, D. J., & Yoong, P. (2004). Studying human-centred IT innovation using a grounded action learning approach. *The Qualitative Report, 9*(1), 137-160.

Pautke, R. W., & Redman, T. C. (2001). The Organisation's Most Important Data Issues. In G. Mario, M. G. Piattini, C. Calero & M. F. Genero (Eds.), *Information and Database Quality*. Dordrecht: Kluwer.

Penchas, S. (2003). Complexity in health care systems. *Complexus, 1*, 149-156.

Pettigrew, A., Ferlie, E., & McKee, L. (1992). *Shaping Strategic Change. Making Change in Large Organisations. The Case of the National Health Service*. London: Sage Publications.

Pierce. (2004a, 25 July). Health care Rx: Information. *The Seattle Times*.

Pierce, E. (2003). *Pursuing a career in information quality: the job of the data quality analyst.* Paper presented at the 8th International Conference on Information Quality, Boston.

Pierce, E. (2004b). *Job wanted: the data quality analyst*. Retrieved 6.08.06, from http://www.tdan.com/i028hy03.htm

Pipino, L. L., Lee, Y. W., & Wang, R. Y. (2002). Data quality assessment. *Communications of the ACM, 45*(4), 211-218.

Plsek, P., & Greenhalgh, T. (2001). The challenge of complexity in health care. *British Medical Journal, 232*, 625-628.

Plsek, P., & Wilson, T. (2001). Complexity, leadership, and management in health care organisations. *British Medical Journal, 323*, 746-749.

Plummer, A. A. (2001). *Information systems methodology for building theory in health informatics: the argument for a structured approach to case study research.* Paper presented at the 34th Hawaii International Conference on System Sciences.

Porcheret, M., Hughes, R., Evans, D., Jordan, K., Whitehurst, T., Ogden, H., et al. (2004). Data quality of general practice electronic health records: the impacts of a program of assessments, feedback and training. *Journal of the American Medical Informatics Association, 11*, 78-86.

Porter, M. (1991a). Towards a dynamic theory of strategy. *Strategic Management Journal, 12*, 95-117.

Porter, M. (1998). What is strategy? In D. Ulrich (Ed.), *Delivering Results: A New Mandate for Human Resource Professionals*: Harvard Business School Press.

Porter, M. E., & Olmestead Teisberg, E. (2004). Redefining competition in health care. *Harvard Business Review, June*, 65-76.

Porter, S. (1991b). A participant observation study of power relationships between nurses and doctors in a general hospital. *Journal of Advanced Nursing, 16*, 728-735.

Porter, S. C., & Mandl, K. D. (1999). *Data quality and the electronic medical record: a role for direct parental data entry.* Paper presented at the American Medical Informatics Association Symposium, Washington DC.

Pretzlik, V. (1994). Observational methods and strategies. *Nursing Researcher, 2*(2), 13-21.

Price, R. J., & Shanks, G. (2005). *Empirical refinements of a semiotic information quality framework.* Paper presented at the 38th Hawaii International Conference on Systems Sciences.

Pringle, M., Wilson, T., & Grol, R. (2002). Measuring "goodness" in individuals and healthcare systems. *British Medical Journal, 325*, 704-707.

Redman, T. C. (1994). Data Quality for telecommunications. *IEEE Journal on Selected Areas in Communications, 12*(2), 306-312.

Redman, T. C. (1996a). *Data Quality in the Information Age*: Artech House.

Redman, T. C. (1996b). Why Care About Data Quality. In *Data Quality in the Information Age* (pp. 303): Artech House.

Redman, T. C. (2001). *Data Quality. The Field Guide.* Boston: Digital Press.

Ribiere, V., LaSalle, A., Khorramshahgol, R., & Gousty, Y. (1999). *Hospital information systems quality: A customer satisfaction assessment tool.* Paper presented at the 32nd Hawaii International Conference of System Sciences.

Rigby, M. (1998). Integrated record keeping as an essential aspect of a primary care led health service. *British Medical Journal, 317*, 579-582.

Robson, W. (1997). *Strategic Management and Information Systems* (Second ed.). London: Pitman Publishing.

Rosenhead, J. (1998). *Complexity theory and management practice*. Retrieved 20.08.06, from http://human-nature.com/science-as-culture/rosenhead.html

Rothenburg, J. (1996). *Meta data to support data quality and longevity.* Paper presented at the First IEEE Metadata Conference, Silver Spring, MD.

Rusin, R. S., & Williams, A. M. (2001). Challenges in incorporating a quality strategy into information systems development efforts. *Topics in Health Information Management, 21*(4), 21-29.

Ryan, K. (2004). Universal electronic health records: a qualitative study of lay perspectives. *NZFP, 31*(3), 149-154.

Saba, V. K., & McCormick, K. A. (2001). *Essentials of Computers for Nurses. Informatics for the New Millennium* (Third Edition ed.). New York: McGraw-Hill.

Sanderson, H., Adams, T., Budden, M., & Hoare, C. (2004). Lessons from the central Hampshire electronic health record pilot project: evaluation of the electronic health record for supporting patient care and secondary analysis. *British Medical Journal, 328*, 875-878.

Schultze, U. (2001). Reflexive Ethnography of Information Systems Research. In E. M. Trauth (Ed.), *Qualitative Research in IS: Issues and Trends* (pp. 78-103). London: Idea Group Publishing.

Segev, A. (1999). Data quality challenges in enabling ebusiness transformation. Paper presented at the Seventh International Conference on Information Quality, Boston.

Seiner, B. (1997). *Improving data quality through information stewardship.* Paper presented at the Seventh International Conference on Information Quality, Boston.

Shankaranarayan, G., & Cai, Y. (2005). Supporting data quality management in decision-making. *Decision Support Systems*.

Shankaranarayan, G., Ziad, M., & Wang, R. Y. (2003). Managing data quality in dynamic decision environments: an information product approach. *Journal of Data Management, 14*(4), 14-32.

Shortliffe, E. H., & Barnett, G. O. (2000). Medical Data: Their Acquisition, Storage and Use. In E. H. Shortliffe, L. E. Perrecult, G. Wiederhold & L. M. Fagan (Eds.), *Computer Applications in Health Care and Biomedicine* (2nd ed., pp. 37-69): Springer.

Sidthorpe, B., Glasgow, N., & Longstaff, D. (2004). *Complex adaptive systems: A different way of thinking about health care systems*. Retrieved 25.8.05, from http://www.anu.edu.au/aphcri/Publications/Background_paper_stream1.pdf

Skinner, B. F. (1963). Behaviourism at fifty. *Science, 140*, 951-958.

Smith, M. K. (2001). *Peter Senge and the learning organisation*. Retrieved 30/8/05, from http://www.infed.org/thinkers/senge.htm

Spradley, J. P. (1980). *Participant Observation*. New York: Holt, Rinehart and Winston.

Stacey, R. D. (1993). *Strategic Management and Organisational Dynamics*. London: Pitman.

Stackpole, B. (2001). *Keeping info clean*. Retrieved 3.11.05, from http://www.eweek.com/article2/0,1759,1250280,00.asp

Stewart, D. W., & Shamdasani, P. N. (1990). *Focus Groups: Theory and Practice*. California: Sage.

Storey, V. C., & Wang, R. Y. (1994). *Modelling quality requirements in conceptual database design.* Paper presented at the International Conference on Information Quality, Massachusetts.

Strauss, A. (1987). *Qualitative Analysis for Social Scientists*. Cambridge: Cambridge University Press.

Strauss, A., & Corbin, J. (1998). *Basics of Qualitative Research*. Thousand Oaks: Sage Publications.

Strong, D. M., Lee, Y. W., & Wang, R. Y. (1997). Data quality in context. *Communications of the ACM, 40*(5), 103-110.

Suh, B., & Han, I. (2003). The impact of customer trust and perception of security control on the acceptance of electronic commerce. *International Journal of Electronic Commerce, 7*(7), 135-162.

Susman, G., & Evered, R. (1978). An assessment of the scientific methods of action research. *Administrative Science Quarterly, 23*(4), 582-603.

Sutcliffe, K., & Weber, K. (2003). The high cost of accurate knowledge. *Harvard Business Review, May 2003*, 74-82.

Tayi, G. K., & Ballou, D. P. (1998). Examining data quality. *Communications of the ACM, 41*(2), 54-57.

TechRepublic. (2005). *Information Management Trends in 2005*. Retrieved 8.8.05, from www.techrepublic.com

Thiru, K., Hassey, A., & Sullivan, F. (2003). Systematic review of scope and quality of electronic patient record data in primary care. *British Medical Journal, 326*, 1070-1072.

UK Audit Commission. (2004). *Information and Data Quality in the NHS*. London: Audit Commission.

United Kingdom Department of Health. (2004). *A Strategy for NHS Information Quality Assurance*.

United States Department of Defence. (2003). *DOD Guidelines on Data Quality Management*: Defence Information Systems Agency.

United States Department of Health and Human Services. (2002). *HHS Guidelines for ensuring and maximising the quality, objectivity, utility, and integrity of information dissemination to the public*. Retrieved 6.11.05, from http://www.hhs.gov/infoquality/part1.html

Urquhart, C. (2001). An encounter with grounded theory: Tackling practical and philosophical issues. In E. M. Trauth (Ed.), *Qualitative Research in IS: Issues and Trends* (pp. 298). London: Idea Group.

van Bemmel, J. H., & Musen, M. A. (1997). *Handbook of Medical Informatics*. Houston: Springer-Verlag.

van Gool, K., Lancsar, E., Viney, R., Hall, J., & Haywood, P. (2002). Diagnosis and prognosis of Australia's health information for evidence-based policy. *Journal of Health Services Research Policy, 7*(Suppl 1), S1:40-S41:45.

van Manen, M. (1997). *Researching Lived Experience: Human Science for an Action Sensitive Pedagogy* (2nd ed.). London, Ontario: Althouse Press.

Wand, Y., & Wang, R. Y. (1996). Anchoring data quality dimensions in ontological foundations. *Communications of the ACM, 39*(11), 86-95.

Wang, R. Y. (1998). A product perspective on total data quality management. *Communications of the ACM, 41*(2), 58-63.

Wang, R. Y., Lee, Y. W., Pipino, L. L., & Strong, D. M. (1998). Manage your information as a product. *Sloan Management Review*(Summer), 95-105.

Wang, R. Y., Storey, V. C., & Firth, C. P. (1995). A framework for analysis of data quality research. *IEE Transactions on Knowledge and Data Engineering, 7*(4), 623-640.

Wang, R. Y.& Strong, D. M. (1996). Beyond accuracy: what data quality means to data consumers. *Journal of Management Information Systems, 12*(4), 5-33.

Wang, R. Y., Ziad, M., & Lee, Y. W. (2001). *Data Quality*. Massachusetts: Kluwer Academic Publishers.

Ward, J., & Peppard, J. (2002). *Strategic Planning for Information Systems* (Third ed.). Chichester: John Wiley & Sons.

WAVE Advisory Board. (2001). *From Strategy to Reality. The WAVE (Working to add Value through E-information) Project. Health Information Management and Technology Plan* (Project Report). Wellington: Ministry of Health.

Weed, L. (1989). *New premises and tools for medical care and medical education.* Paper presented at the International Symposium of Medical Informatics and Education, British Columbia, Canada.

Willshire, M. J., & Meyen, D. (1997). A process for improving data quality. *Data Quality, 3*(1).

World Health Organisation. (1988). *Informatics and Telematics in Health. Present and potential uses*. Geneva.

Wu, J. (2004). *Empowering the information enterprise: Business processes are the foundation of an information enterprise*. Retrieved 3.11.05, from http://www.dmreview.com/article_sub.cfm?articleID=1003984

Zeffane, R., Cheek, B., & Meredith, P. (1998). Does user involvement during information systems development improve data quality? *Human Systems Management, 17*, 115-121.

Zmud, R., Lind, M., & Young, F. (1990). An attribute space for organisational communication channels. *Information Systems Research, 1*(4), 440-457.