

Association for Information Systems  
**AIS Electronic Library (AISeL)**

---

ACIS 2019 Proceedings

Australasian (ACIS)

---

2019

## Perceptions and Challenges of EHR Clinical Data Quality

Shane Downey

*University of Queensland, s.downey@uqconnect.edu.au*

Marta Indulska

*University of Queensland, m.indulska@business.uq.edu.au*

Shazia Sadiq

*University of Queensland, shazia@itee.uq.edu.au*

Follow this and additional works at: <https://aisel.aisnet.org/acis2019>

---

### Recommended Citation

Downey, Shane; Indulska, Marta; and Sadiq, Shazia, "Perceptions and Challenges of EHR Clinical Data Quality" (2019). *ACIS 2019 Proceedings*. 25.

<https://aisel.aisnet.org/acis2019/25>

This material is brought to you by the Australasian (ACIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in ACIS 2019 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact [elibrary@aisnet.org](mailto:elibrary@aisnet.org).

# Perceptions and Challenges of EHR Clinical Data Quality

*Full Paper*

## **Shane Downey**

School of Information Technology and Electrical Engineering  
The University of Queensland  
St Lucia, QLD 4072 Australia  
Email: [s.downey@uqconnect.edu.au](mailto:s.downey@uqconnect.edu.au)

## **Marta Indulska**

UQ Business School  
The University of Queensland  
St Lucia, QLD 4072 Australia  
Email: [m.indulska@business.uq.edu.au](mailto:m.indulska@business.uq.edu.au)

## **Shazia Sadiq**

School of Information Technology and Electrical Engineering  
The University of Queensland  
St Lucia, QLD 4072 Australia  
Email: [shazia@itee.uq.edu.au](mailto:shazia@itee.uq.edu.au)

## **Abstract**

Despite the premise of better data, Electronic Health Record (EHR) data quality remains problematic. Traditional approaches for improving data quality through semantic and syntactic controls have not resolved the problems. To use the medical vernacular – “we have addressed the symptoms but not the cause.” This paper reports on an exploratory study undertaken in a large maternity hospital with an aim to expose detractors from high-quality data in EHRs. The study involved a perceptions survey that was completed by Nursing and Midwifery staff; chosen because of known data quality challenges in their area of practice. The study results indicate social, cultural and environmental aspects of information systems (IS) use are equally as problematic as the IS itself. A lack of agreement amongst healthcare practitioners surrounding what data quality means is also evident, with time, culture and lacking formal education on data quality being contributors to lower data quality outcomes.

**Keywords:** Data quality perceptions, survey, health care, Midwifery and Nursing, perinatal data.

## 1 INTRODUCTION

Electronic Health Records (EHRs) enhance patient care through the centralised collection, storage and management of patient demographic, clinical and diagnostic data. This assists in providing clinicians and healthcare service providers timely and efficient access to all relevant information, inclusive of medical history, for an episode of patient care (Sundararaman and Ramanathan 2016). More advanced solutions use this information for automated decision support, increasing the speed of primary care diagnoses alongside early warning detection of change in a patient's condition. The success of EHRs lie in the level of reliability of the information it provides – high quality data is always required to operate appropriately. EHRs aim to reduce data quality issues through the cross-integration of various components of healthcare management under one umbrella, making data available across key stakeholders in the organisation as it is captured. For example, as soon as the results of diagnostic imaging or pathology results are captured in the EHR, they are immediately accessible by the treating clinicians in the ward. In general, systems with lesser quality data risk facilitating wrong diagnoses, potentially leading to inefficient processes and delays, financial loss and patient harm (ECRI Institute PSO 2015; Haug *et al.* 2011; Keene *et al.* 2018; Shulte and Fortune 2018; Vanderhook and Abraham 2017).

Data quality literature in healthcare tends to focus on semantic and syntactic data quality (i.e. the structural definition of data as it relates to IS) (Batini *et al.* 2009). This is a bottom up approach, retrospectively attempting to manage data quality issues through identification of problems and implementing controls to manage these issues (Almutiry *et al.* 2013; Orr 1998). Although EHRs provide better integration of data to minimise manual data entry duplication, various technology constraints and a lack of standards can restrict the degree of integration possible. This has caused manual data entry to remain a principle method for EHR data acquisition (Mays and Mathias 2019). Regardless of this manual approach, there remains less focus on understanding the problems and challenges faced by the people performing this manual data entry. Studies of human behaviour, and of IS use, theorise that attitudes and behaviours of the people performing the data entry effect data quality outcomes, with some studies suggesting problems of manually acquired data may lie with the data entry operator (Charnock 2019; Haegemans *et al.* 2016). Despite their prevalence in research, there is limited empirical insights of attitudes and behaviours relating to data quality and manually acquired data in practice, especially in the EHR clinical setting.

To analyse the effects of this manual entry, this study was undertaken with the support of a large tertiary maternity hospital in Australia. This hospital delivers close to 10,000 babies each year and caters for the full spectrum of care from straightforward deliveries to the most complex care cases. In 2007, this hospital implemented an obstetric EHR. This solution employs an archetype model to capture and represent data (Beale 2001) and was tailored to the facility to ensure that the key focus was placed on capturing the maternity care path. This was achieved by mirroring the pre-existing paper-based record structure. The implementation of the EHR however, did not replace the existing paper record, thus it remains today as a core part of what is a hybrid patient record (Houben *et al.* 2015).

Since the inception of this EHR, there have been ongoing challenges with data quality – most noticeably in meeting statutory reporting obligations. Data for reporting must be exported from the EHR and manually adjusted to meet the data quality requirements imposed by the State Government health authority. There are gaps in the data available in the EHR, as not all data is transcribed from the paper record. Various attempts have been made to reduce the manual cleansing required – improvements to the structure of the data capture interface (including increasing syntactic controls), simplification of archetypes, enhancements to the EHR rules, and additional resourcing to assist with workload. These attempts have had mixed success in reducing data quality issues and consequently a significant amount of on-going manual work is required to make the data satisfactory for reporting.

The aim of this exploratory study is to expose detractors from high-quality data in the EHR, as identified by the users of the EHR within the clinical care setting. This study expands the body of research about data quality in healthcare, by providing fresh insights from the perspective of the EHR user and in the context of manually acquired data.

## 2 RELATED WORK

Much of the research within the field of data quality centres on the semantic and syntactic aspects of data (i.e. the way data is defined and the rules that surround it) and predefined sets of measurements, commonly referred to as dimensions (Jayawardene *et al.* 2013). These dimensions are generally used to determine the degree to which data is 'fit for purpose' – thus the focus is on data use after it has been

collected, not the collection, or collection method, itself. Orr (1998) hypothesised that “data quality in an information system is a function of its use and not of its collection” (P7). However, the collection method of manual data entry remains a primary mechanism for acquiring data in EHRs, and if the data is incorrect then the impact to patients and patient care could be significant.

Models such as those proposed by Nelson *et al.* (2005;) and Pai and Huang (2011) have examined how the dimensions of data quality may effect IS use and perceived usefulness – but their examinations do not consider how key behavioural (i.e. social, cultural and environmental) aspects effect data quality. These effects on data quality have been explored, to some degree, by examining the role of subjective norms and beliefs on perceived levels of control and personal innovativeness in IS use (Craswell *et al.* 2014a; Hubona and Geitz 1997; Jackson *et al.* 2013; Kortteisto *et al.* 2010). However, while these findings are useful for prediction, they do not explain why data quality problems occur.

In contrast, studies in healthcare have largely focused on the technical aspects of data quality (Craswell *et al.* 2013; Kerr 2006; Strong *et al.* 1997) with a specific focus on data use. Conversely, there are frameworks developed to address data quality in healthcare with a focus on EHRs themselves (Almutiry *et al.* 2013). Yet, there are few studies that attempt to tease out specific causes of data quality problems (particularly in relation to manual data entry) within EHRs. Charnock (2019, p. 94) also noted that there is no definition of data quality specific to health, and proposed the following: “Data which is accurate, complete, used solely for the purpose it was originally collected for and is available when the user expects it to be available.” However, this definition lists only three dimensions of data quality (completeness, accuracy and timeliness), whereas established research indicates that there are eight dimensions and thirty-three patterns that apply in the data quality context (Jayawardene *et al.* 2013). The definition is further limiting as a significant amount of health data is captured for primary and secondary use (including perinatal data reporting). The suggested problem can be theorised as a lack of understanding about how data is used.

The assessment of data quality in healthcare appears to be somewhat constrained to the same three data quality dimensions as identified above - completeness, accuracy, and timeliness – as identified by Chen *et al.* (2014) in a review of literature. The researchers noted that “measuring the perceptions of end users or consumers towards data quality will enrich our understanding of data quality issues” (P5187). In our study we aim to contribute to the body of knowledge by answering this call and studying perceptions of EHR users in a clinical care setting.

### 3 METHODS

This study was conducted with the intention of focusing on the management of perinatal data by Nurses and Midwives. This group of participants was chosen because they are representative of users who work directly with manually acquired data in a clinical care setting, and because there is evidence that the data they deal with (perinatal data) has extensive problems. By way of example, data recorded in the EHR by midwifery staff can be inconsistent with data recorded in the paper record by medical staff. Handwritten notes are entered contemporaneously, however the electronic record is usually completed some hours after the event, resulting in errors in times of critical events, such as time of birth of baby. This group was also chosen because perinatal data extracts are submitted to a health authority for analysis, research and health planning, and is a hospital licencing requirement. The health authority has detailed (documented) data validation rules which they apply and feedback in terms of data problems are returned to the source hospital for correction. There are strict deadlines for correction and resubmission of data, which are often exceeded again due to the human effort required to manually source and correct data. Non-compliance could be seen as a breach and there are financial penalties that can be applied by the health authority.

Given the aim of this study is to expose underlying detractors from high quality, manually acquired EHR data - we adopted a survey-based approach to elicit the perceptions and challenges of the people performing data entry. The following sub-sections provide further methodological details.

#### 3.1 Instrument Design

An online survey was chosen as the tool to obtain perceptions and challenges about data quality from staff so as to minimise intrusion, and give staff the most flexibility in terms of when they could participate (Evans and Mathur 2005). The survey was voluntary and anonymous to protect the identity of respondents. As the hospital operates 24 hours per day, an online survey provides an opportunity to have a cross section of employee types and professions included.

An open style of questioning combining qualitative and quantitative approaches (Kaplan and Duchon 1988), was used with a neutral tone. The response variables are a combination of the 7-point Likert scale (where 1="Strongly Disagree", 4="Neutral" and 7="Strongly Agree") whereby categorical, binary and open-ended text were utilised. The survey instrument was peer reviewed and pilot-tested by a group of data management professionals prior to launching.

The questions in the survey were divided into three parts:

- i. **General demographic information.** This section was designed to capture basic demographic information about the participants, including their degree of data quality education and awareness.
- ii. **Data quality perceptions.** Research indicates that users of health IS are more likely to use the systems if the perceived value (i.e. of benefit to them) is high (Pai and Huang 2011). Studies on data quality in health have also highlighted this need for clear definition of roles and accountabilities in regards to data and records management, also recommending training or the need for specific qualifications to be obtained for some roles (Craswell *et al.* 2014a; Health Information and Quality Authority 2011, p. 17). Accordingly, in this set of questions we examined the respondents' knowledge about data quality, in terms of understanding its meaning and importance. We explored their perceptions of the importance of each of the data quality dimensions, and the importance of data quality to a set of primary and secondary use functions. Finally, we explored roles and responsibilities for managing data, and specific EHR use.
- iii. **Challenges of managing data.** An open-ended question at the end of the protocol was dedicated to capturing perceived challenges with managing data in the given context. We opted for an open-ended question as it allowed us to elicit feedback that structured questions may not (Reja *et al.* 2003). This also allowed us to tap into emotional aspects of data quality to further elicit attitudes (Greer 1988), which is an important aspect of our study.

### 3.2 Study Population & Sampling Procedure

The study cohort (n=678) was identified by hospital management as users of the obstetric EHR. The study cohort consisted of Nurses (24%), Midwives (63%), and other health professionals such as clinical facilitators, case managers and caseload managers (13%).

An anonymous online survey link, cover letter and participant information sheet were distributed via email to the study population by the hospital's Chief Medical Officer. A follow up reminder was sent each week during the survey period, running for the whole of November 2018. A total of 127 responses were received, of which 79 were complete (incomplete responses have been excluded from data analysis).

### 3.3 Data Analysis

Upon completion of the collation of these results, statistical analysis was performed using IBM SPSS v.25. The characteristics of the respondents were analysed using descriptive statistics. Cronbach's alpha correlation coefficient was used to measure the internal consistency of the scale variables with  $\alpha=0.9$ ; considered to be a good result (George and Mallery 2016). The open-ended question, 'What are your 3 biggest challenges with managing data?' was coded based on themes in the text. Where the themes aligned with data quality characteristics, as defined by Zhang *et al.* (2014), the data quality characteristic was used. Where a specific data quality characteristic could not be identified (e.g. time, culture, etc.), the theme was captured instead.

## 4 RESULTS

The results of this study were collated and categorised using the respondents selected profession. A list was provided which specified each of the professions (Nurse, Midwife, Other) to ensure accuracy. This question allowed for data analysis by profession (Figure 1: Characteristics of the survey respondents). Of the 79 responses, the distribution of the responses by profession was 33 Midwives, 14 Nurses and 7 indicated a combined profession of "Nurse and Midwife". The remaining 25 chose "Other" as their profession, representing all other clinical roles. The survey results also indicate that 94% of Nurses and Midwives have formal tertiary qualifications. This is not unexpected as a relevant tertiary education would be a requirement of these roles.

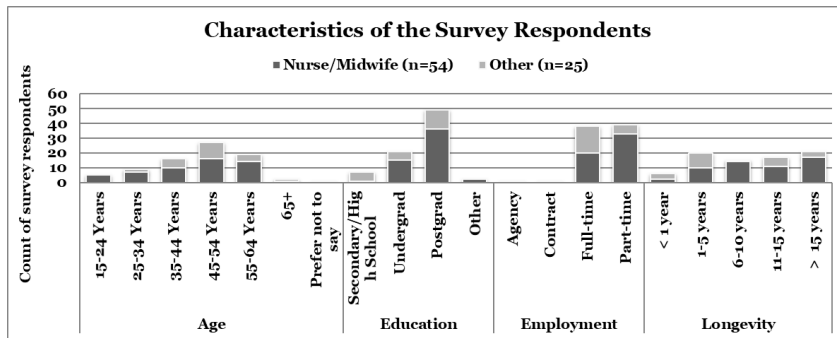


Figure 1: Characteristics of the survey respondents – Nurse & Midwife vs Other role types

Our analysis of this data indicates that 57% of the Nurses and Midwives were over the age of 45, and 51% had been working in the same hospital for more than ten years. In addition, 61% of respondents were part-time.

#### 4.1 Data Quality Perceptions

Respondents were initially asked who they felt should be responsible for data quality (Figure 2: Perception of responsibility for data management); multiple responses were possible. ‘Self’ and ‘Manager’ are ranked highest, with ‘Everyone’ ranked in the bottom 3. What we can ascertain from this result is a sense that individuals know they should be responsible for data within their area, but in a broader sense, that the responsibility sits with other individuals.

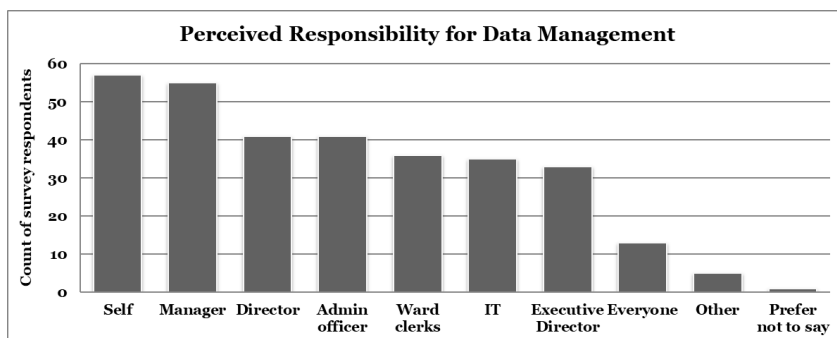


Figure 2: Perception of responsibility for data management ranked by count of responses

Respondents were asked to identify the importance of quality data to areas of primary and secondary use (Figure 3: Importance of data quality to various hospital functions). When it came to patient care and research use of data, we see agreement that quality data was most important. When the use related to billing however, inclusive of reporting and rostering, we see an increased number of neutral responses. We posit from the analysis of this question that staff do not have a clear understanding of how the information is used, beyond the immediate primary use of caring for patients.

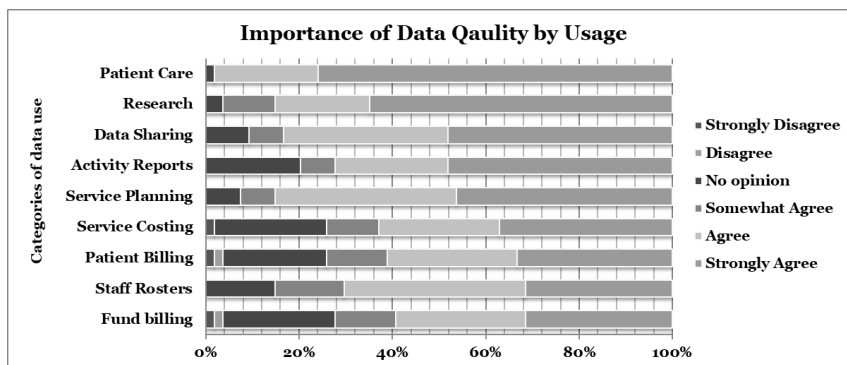


Figure 3: Importance of data quality to various hospital functions

This finding suggests that additional contextual data quality education may be beneficial to help the data producers understand the value of the secondary use of data (Botsis *et al.* 2010).

Respondents were asked about their understanding of what data quality means, and if they had received data quality education. Further questions in this section explored the importance of each of the dimensions of data quality in general (Figure 4: Perceived importance of data quality dimensions). Accuracy, Availability & Accessibility, and Reliability & Credibility were the highest responses, while Consistency and Validity received the lowest rankings. Some responses (Consistency, Usability & Interpretability for example) have strong neutral responses. 77.7% of Nurses and Midwives responded positively that they were aware of what data quality means and 46.3% indicated they had received data quality education. Of those who indicated they had received data quality education, 17.7% responded that they had received formal education, and 39.2% responded that they learned about data quality on the job. These results indicate that most of the Nurses and Midwives cohort have not had formal data quality education, but that most have managed to develop awareness of the concept; primarily through workplace learning.

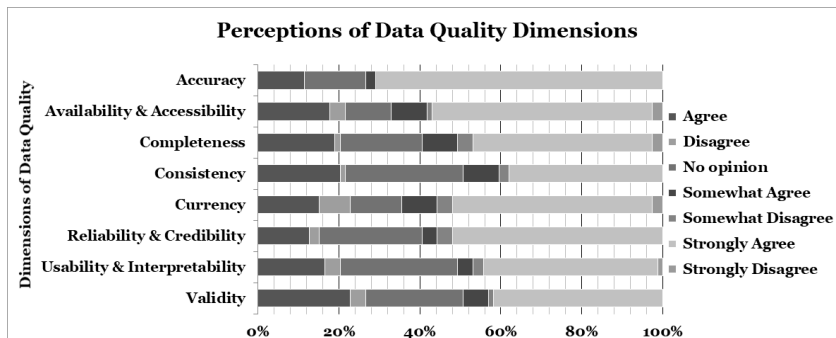


Figure 4: Perceived importance of data quality dimensions

The respondents were then asked to indicate if they felt data quality education would be beneficial for themselves or their team. 96% of Nurses and Midwives felt data quality education would benefit them, and 90% felt data quality education would benefit their team. This high response indicates a significant opportunity to improve data quality through providing training and suggests that Nurses and Midwives are not confident in their knowledge about data quality, yet most recognise its importance.

## 4.2 Data Quality Challenges

Our analysis of the open-ended responses (Figure 5: Challenges of data quality mapped to data quality dimensions) indicates which data quality dimensions were the most challenging. In total, we coded 113 discreet feedback items. In comparing these responses with the perceived importance of data quality dimensions (Figure 4: Perceived importance of data quality dimensions), Accuracy appears to be the most important, and the most challenging.

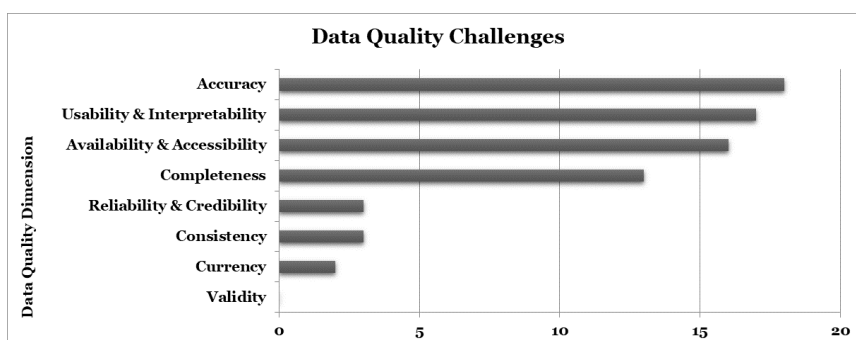


Figure 5: Challenges of data quality mapped to data quality dimensions

Availability and Accessibility also scored high, identifying that getting access to information is both important and a challenge. While Usability & Interpretability scored high in the coded challenges, it was ranked third from last in terms of perceived importance. This mismatch between importance and use is interesting as it displays a difference between perceived use, and actual use, which could be a potential avenue for further investigation.

In addition to the coded data quality impacts, a number of non-data quality related responses were provided as reasons for why there are data quality challenges (Figure 6: Themes affecting data quality).

Whilst all challenges are interesting, the top five challenges account for 72% of the results, thus we focus on these below.

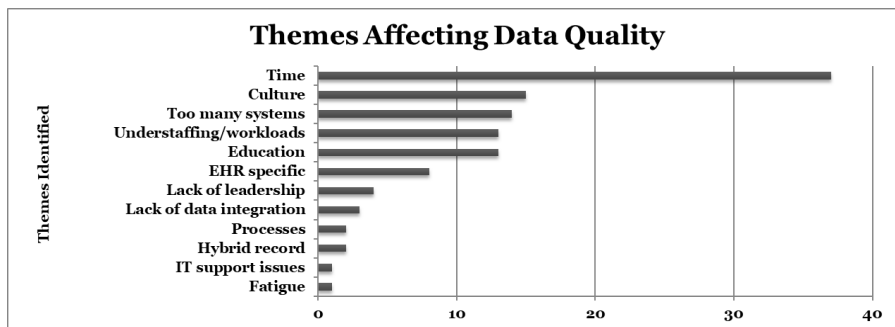


Figure 6: Themes affecting data quality

**Time.** Lack of time was indicated as the strongest detractor from providing quality data. For example, one respondent stated:

*“Being time poor and having excessive work load, lack of support from team, understaffing, acuity, time to collect data and staff attitudes to staff collection.”*

According to our data, time is affected by staffing levels, workloads, duplication of effort and patient acuity. Respondents also indicated that fatigue (in relation to workloads and understaffing) effects the ability to enter high quality data. Culture

There were also comments that clinical documentation isn’t properly completed and correctly entered into the obstetric EHR before handover, leading to information gaps. For example:

*“Incomplete or missing information for privately managed patients... in their poorly written medical notes. This makes it more difficult to capture accurate data.”*

*“Having other clinicians agree that the EHR provided antenatal history is not the job of only one department. That is the patient is admitted, it is the responsibility of the clinician caring for her. Sometimes the OB doesn't ask all that we do or share their appointment notes. Not all data collected in the EHR is visible in-patient record, and some clinicians don't access the EHR data until discharge.”*

**Information systems:** Respondents stated there were too many IS where they must input data, which often causes duplication of data. For example, one respondent indicated:

*“Unbooked patients who come into birth suite in labour - it's impossible to fill in the whole of the EHR. It takes time to enter data into the EHR and often you have to repeat the same information more than once.”*

Respondents also indicated that the EHR archetype model needed to be adjusted to better align with the clinical care pathways, making data entry easier. Better integration between IS was suggested by respondents to reduce duplicate data entry.

**Education:** Knowing how to access data, and what to do with it once they found it was reported. Comments suggested that some staff may not understand the importance of the data and therefore not put as much care into their data entry. Specific training in the EHR was suggested, as evidenced by this example comment:

*“Knowing how to do it, accessing the data, need more education regarding obtaining quality data, lack of education around data quality and its importance, lack of leadership in ensuring data quality is a strategic goal lack of investment in tools to assist with data quality and making data visible to the business.”*

**Understaffing/Workloads:** Commonly associated with Time in the responses was the tension between patient record management and patient acuity. Our analysis indicates that the higher the patient acuity, the less time the employee has for data entry, as exemplified by:

*“Time, under staffed, casual staff, inadequate staffing levels, time constraints, secondary to high acuity of patients, different systems/duplicate documentation, constant changes in processes...”*



## 5 DISCUSSION AND OUTLOOK

The aim of this study was to uncover detractors of high-quality data within EHRs by understanding the data quality perceptions and challenges that Nurse and Midwifery staff experience when using an EHR, within a healthcare setting. Through this study, we add to the current body of knowledge by offering further empirical insights in a field where minimal insights exist.

Our findings align with other studies in identifying data accuracy, reliability and timeliness as the dimensions of most importance to our study participants (Charnock 2019; Chen et al. 2014). The findings also indicated that time (workload), culture and the need for duplicated data entry are detractors from high-quality data; particularly within the context of a hybrid patient record. We also identified a lack of formal data quality education to be a cause of lower data quality outcomes. Uncertainty about roles and responsibilities of data management is also as a challenge when trying to escalate or resolve data quality issues.

The effects of time were identified as an issue in a similar study of midwives and perinatal data collection officers (Craswell *et al.* 2014b) where the researchers found that “participants reported that when they were busy or pressed for time, they enter less data into the perinatal data record” (P297). Ahituv *et al.* (1998) examined the effects of time pressures on the overall level of data quality, and demonstrated that as time pressures increase, individuals determine for themselves which pieces of information are more or less important than others, and look to reduce effort in some tasks, such as searching. This aligns with the notion that cognitive-fit (Vessey 2006) stems from a cost-benefit trade-off between the task to be performed, the availability of information and the level at which the information matches the task alongside the amount of time available to complete the task. As the complexity of a symbolic task increases, the decision maker may switch cognitive strategies from symbolic to spatial to reduce the time element (Vessey 1994). Similarly, the study by Murphy (2009) suggested that “a re-assessment of aspects, such as staffing ratios and work schedules, of those tasked with manual data acquisition” (P1885) may be required. Our study confirms time as an influencer of data quality in EHRs.

Our key findings indicate that although Nurses and Midwives are aware of data quality importance, they may not be confident in their level of knowledge (as seen by the indicated need for DQ training). This finding is also supported by suggestions in the open-ended responses that data quality knowledge was not prevalent and further education about data quality would be beneficial. Given that manual data entry remains a primary source of data acquisition for health IS (Sukumar *et al.* 2015), education about data quality dimensions should be included in any education involving information systems use. A key example that was drawn from these results includes Completeness and Accuracy as two data quality dimensions that measure the degree to which data matches its real-world object. Through the utilisation of these dimensions, educating data entry operators could frame learning to yield a greater understanding of the importance and value of quality data when instructing, which would lead to improved data quality outcomes. In the study of the New Zealand health system by Kerr *et al.* (2006) the researchers noted:

*“...being aware of the importance of both education and training at all necessary stages and levels - training is essential to effect the culture change that must accompany the realisation of the importance of data quality” (P111)*

Although education and training are recommended in data quality frameworks, EHR data quality education, and its effect, has not been studied thus far. This makes for a large gap in knowledge, and an identified opportunity for further research. The identification that education about secondary data usage (Safran *et al.* 2007) might yield improvements to data quality outcomes has become paramount to the suggestion of further investigation requirements. By doing so, a better understanding can be achieved surrounding the effects of data quality education on data quality outcomes (via a study of behaviours and outcomes).

A study examining the unintended consequences of EHR systems implementation (Vanderhook and Abraham 2017) summarises that “increased observation during implementation should be put on people, their values, norms, and the culture...” (P224), and notes that “the social and technical elements are deeply rooted and intertwined” (P224). There is little literature about the effects of culture on data quality outcomes, however, a study by Shanks and Corbitt (1999) stated that “social level data quality problems are highly relevant for organisations that are concerned with enterprise-wide or global information systems.” (P795). Through our survey results we find empirical support for culture as an influencer on data quality outcomes. It is notable however that further work is required to refine these behavioural aspects in order to form a model and approach that can be utilised in any organisation.

This study is not without limitations. The perception survey response rate was relatively low (n=79) and most respondents for the Nurse and Midwifery cohort were part-time. Their views may not represent the other role types adequately (full-time, contract, casual etc.); therefore, the results cannot rule out self-selection and non-response biases. Most of the respondents work within the same hospital, thus it is possible that people who have completed the survey could have influenced others. Despite these limitations, the responses were of good internal reliability ( $\alpha=.9$ ). The study was undertaken in a large, busy hospital and therefore subject to forces outside of the control of the research team. The results align with general research about data quality in healthcare, as well as specific research about perinatal data, Nursing and Midwifery. As noted by Dixon-Woods *et al.* (2012, P882) there is no 'silver bullet' to addressing data quality in healthcare, and multiple approaches are likely to be required. This includes data quality education, defining roles and responsibilities for data management, allowing more time for data entry or exploring technology solutions to assist with data entry (to maximise accuracy and improve accessibility), and ensuring input screens for manual data entry match the clinical pathways. These findings can be used in conjunction with other data quality frameworks and assessments to create a more holistic approach to identifying and managing data quality in healthcare.

## 6 REFERENCES

- Ahituv, N., Igarbaria, M., and Sella, A. V. 1998. "The Effects of Time Pressure and Completeness of Information on Decision Making," *Journal of Management Information Systems* (15:2), pp. 153–172. (<https://doi.org/10.1080/07421222.1998.11518212>).
- Almutiry, O., Wills, G., and Crowder, R. 2013. "Toward a Framework for Data Quality in Electronic Health Record," *International Conference E-Society 2013, 13-16 March 2013*, pp. 365–370.
- Batini, C., Cappiello, C., Francalanci, C., and Maurino, A. 2009. "Methodologies for Data Quality Assessment and Improvement," *ACM Computing Surveys* (41:3), pp. 1–52. (<https://doi.org/10.1145/1541880.1541883>).
- Beale, T. 2001. "Archetypes: Constraint-Based Domain Models for Future- Proof Information Systems," *OOPSLA 2002 Workshop on Behavioural Semantics* (21), pp. 1–69. (<https://doi.org/10.1.1.147.8835>).
- Botsis, T., Hartvigsen, G., Chen, F., and Weng, C. 2010. "Secondary Use of EHR: Data Quality Issues and Informatics Opportunities.," *AMIA Joint Summits on Translational Science Proceedings. AMIA Joint Summits on Translational Science* (2010), American Medical Informatics Association, pp. 1–5. (<http://www.ncbi.nlm.nih.gov/pubmed/21347133> <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3041534>).
- Charnock, V. 2019. "Electronic Healthcare Records and Data Quality," *Health Information and Libraries Journal* (36:1), pp. 91–95. (<https://doi.org/10.1111/hir.12249>).
- Chen, H., Hailey, D., Wang, N., and Yu, P. 2014. "A Review of Data Quality Assessment Methods for Public Health Information Systems," *International Journal of Environmental Research and Public Health* (11:5), pp. 5170–5207. (<https://doi.org/10.3390/ijerph110505170>).
- Craswell, A., Moxham, L., and Broadbent, M. 2013. "Computerisation of Perinatal Data Entry: Is the Data Accurate?," in *Studies in Health Technology and Informatics* (Vol. 188), pp. 7–13. (<https://doi.org/10.3233/978-1-61499-266-0-7>).
- Craswell, A., Moxham, L., and Broadbent, M. 2014a. "Shared Responsibility for Electronic Records: Governance in Perinatal Data Entry," *Studies in Health Technology and Informatics* (204), pp. 19–24. (<https://doi.org/10.3233/978-1-61499-427-5-19>).
- Craswell, A., Moxham, L., and Broadbent, M. 2014b. "Does Use of Computer Technology for Perinatal Data Collection Influence Data Quality?," *Health Informatics Journal* (22:2), pp. 293–303. (<https://doi.org/10.1177/1460458214556372>).
- Dixon-Woods, M., McNicol, S., and Martin, G. 2012. "Ten Challenges in Improving Quality in Healthcare: Lessons from the Health Foundation's Programme Evaluations and Relevant Literature," *BMJ Quality and Safety* (21:10), pp. 876–884. (<https://doi.org/10.1136/bmjqs-2011-000760>).
- ECRI Institute PSO. 2015. "Wrong-Record , Wrong-Data Errors with Health IT Systems," *PSO Navigator* (7:2), pp. 1–10.

- Evans, J. R., and Mathur, A. 2005. "The Value of Online Surveys," *Internet Research* (15:2), pp. 195–219. (<https://doi.org/10.1108/10662240510590360>).
- George, D., and Mallery, P. 2016. *IBM SPSS Statistics 25 Step by Step : A Simple Guide and Reference*, London, UNITED KINGDOM: Routledge. (<http://ebookcentral.proquest.com/lib/uql/detail.action?docID=4455907>).
- Greer, J. 1988. "What Do Open-Ended Questions Measure?," *Academic.Oup.Com*. (<https://academic.oup.com/poq/article-abstract/52/3/365/1851766>).
- Haegemans, T., Snoeck, M., Lemahieu, W., Stumpe, F., and Goderis, A. 2016. "Towards a Theoretical Framework to Explain Root Causes of Errors in Manually Acquired Data," *International Conference on Information Quality*, 15:1-15:10.
- Haug, A., Zachariassen, F., and van Liempd, D. 2011. "The Costs of Poor Data Quality," *Journal of Industrial Engineering and Management* (4:2), pp. 168–193. (<https://doi.org/10.3926/jiem.2011.v4n2.p168-193>).
- Health Information and Quality Authority. 2011. *International Review of Data Quality*, (April), pp. 1–59.
- Houben, S., Frost, M., and Bardram, J. E. 2015. "Collaborative Affordances of Hybrid Patient Record Technologies in Medical Work," *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing - CSCW '15*, New York, New York, USA: ACM Press, pp. 785–797. (<https://doi.org/10.1145/2675133.2675164>).
- Hubona, G. S., and Geitz, S. 1997. "External Variables, Beliefs, Attitudes and Information Technology Usage Behavior," *Proceedings of the Hawaii International Conference on System Sciences* (3:1), pp. 21–28. (<https://doi.org/10.1109/HICSS.1997.661560>).
- Jackson, J. D., Yi, M. Y., and Park, J. S. 2013. "An Empirical Test of Three Mediation Models for the Relationship between Personal Innovativeness and User Acceptance of Technology," *Information and Management* (50:4), Elsevier B.V., pp. 154–161. (<https://doi.org/10.1016/j.im.2013.02.006>).
- Jayawardene, V., Sadiq, S., and Indulska, M. 2013. "An Analysis of Data Quality Dimensions," *ITEE Technical Report No. 2013-01*, pp. 1–32. ([http://espace.library.uq.edu.au/view/UQ:312314/n2013-01\\_TechnicalReport\\_Jayawardene.pdf](http://espace.library.uq.edu.au/view/UQ:312314/n2013-01_TechnicalReport_Jayawardene.pdf)).
- Kaplan, B., and Duchon, D. 1988. "Combining Qualitative and Quantitative Methods in Information Systems Research: A Case Study," *MIS Quarterly* (12:4), p. 571. (<https://doi.org/10.2307/249133>).
- Keene, D. J., Lamb, S. E., Tutton, E., Handley, R., and Willett, K. 2018. "Electronic Health Record Usability Issues and Potential Contribution to Patient Harm Downloaded From : By a Western Michigan University Medical School User on 04 / 11 / 2018 Downloaded From : By a Western Michigan University Medical School User on 04 / 11," *Jama* (9), pp. 11–13.
- Kerr, K. 2006. "The Institutionalisation of Data Quality in the New Zealand Health Sector," *Thesis*, The University of Auckland.
- Korteisto, T., Kaila, M., Komulainen, J., Mäntyranta, T., and Rissanen, P. 2010. "Healthcare Professionals' Intentions to Use Clinical Guidelines: A Survey Using the Theory of Planned Behaviour," *Implementation Science* (5:1), BioMed Central, p. 51. (<https://doi.org/10.1186/1748-5908-5-51>).
- Mays, J. A., and Mathias, P. C. 2019. "Measuring the Rate of Manual Transcription Error in Outpatient Point-of-Care Testing," *Journal of the American Medical Informatics Association* (26:3), pp. 269–272. (<https://doi.org/10.1093/jamia/ocy170>).
- Murphy, G. D. 2009. "Improving the Quality of Manually Acquired Data: Applying the Theory of Planned Behaviour to Data Quality," *Reliability Engineering and System Safety*, pp. 1881–1886. (<https://doi.org/10.1016/j.res.2009.05.008>).
- Nelson, R. R., Todd, P. A., and Wixom, B. H. 2005. "Antecedents of Information and System Quality: An Empirical Examination within the Context of Data Warehousing," *Journal of Management Information Systems* (21:4), pp. 199–235. (<https://doi.org/10.1080/07421222.2005.11045823>).
- Orr, K. 1998. "Data Quality and Systems Theory," in *Communications of the ACM* (Vol. 41), ACM, February 1, pp. 66–71. (<https://doi.org/10.1145/269012.269023>).

- Pai, F.-Y., and Huang, K.-I. 2011. "Applying the Technology Acceptance Model to the Introduction of Healthcare Information Systems," *Technological Forecasting and Social Change* (78:4), pp. 650–660. (<https://doi.org/10.1016/j.techfore.2010.11.007>).
- Reja, U., Manfreda, K. L., Hlebec, V., and Vehovar, V. 2003. "Open-Ended vs. Close-Ended Questions in Web Questionnaires," *Developments in Applied Statistics* (19:August 2014), pp. 159–177. ([http://www.websm.org/uploadi/editor/Reja\\_2003\\_open\\_vs\\_close-ended\\_questions.pdf](http://www.websm.org/uploadi/editor/Reja_2003_open_vs_close-ended_questions.pdf)).
- Safran, C., Bloomrosen, M., Hammond, E., Labkoff, S., Markel-Fox, S., Tang, P. C., and Detmer, D. E. 2007. "Toward a National Framework for the Secondary Use of Health Data: An AMIA White Paper," *J Am Med Inform Assoc* (14:1), pp. 1–9. (<https://doi.org/10.1197/jamia.M2273.Introduction>).
- Shanks, G., and Corbitt, B. 1999. "Understanding Data Quality: Social and Cultural Aspects," *10th Australasian Conference on Information Systems*, (1998), pp. 785–797. ([https://www.researchgate.net/profile/Brian\\_Corbitt/publication/228607385\\_Understanding\\_data\\_quality\\_Social\\_and\\_cultural\\_aspects/links/00b4952bof62468150000000.pdf](https://www.researchgate.net/profile/Brian_Corbitt/publication/228607385_Understanding_data_quality_Social_and_cultural_aspects/links/00b4952bof62468150000000.pdf)).
- Shulte, F., and Fortune, E. F. 2018. "Death By A Thousand Clicks: Where Electronic Health Records Went Wrong," *Kaiser Health News*, pp. 1–21. (<https://khn.org/news/death-by-a-thousand-clicks/>, accessed June 28, 2019).
- Strong, D. M., Lee, Y. W., and Wang, R. Y. 1997. "Data Quality in Context," *Communications of the ACM* (40:5), pp. 103–110. (<https://doi.org/10.1145/253769.253804>).
- Sukumar, S. R., Natarajan, R., and Ferrell, R. K. 2015. "Quality of Big Data in Health Care," *International Journal of Health Care Quality Assurance* (28:6), pp. 621–634. (<https://doi.org/10.1108/IJHCQA-07-2014-0080>).
- Sundararaman, A., and Ramanathan, S. V. 2016. "Open Research Issues and Emerging Research Directions in Data Quality for Public Health," *Proceedings of the 21st International Conference on Information Quality (ICIQ-2016)*.
- Vanderhook, S., and Abraham, J. 2017. "Unintended Consequences of EHR Systems: A Narrative Review," *Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care* (6:1), pp. 218–225. (<https://doi.org/10.1177/2327857917061048>).
- Vessey, I. 1994. "The Effect of Information Presentation on Decision Making: A Cost-Benefit Analysis," *Information and Management* (27:2), pp. 103–119. ([https://doi.org/10.1016/0378-7206\(94\)90010-8](https://doi.org/10.1016/0378-7206(94)90010-8)).
- Vessey, I. 2006. "The Theory of Cognitive Fit," *Human-Computer Interaction and Management Information Systems: Foundations* (5), pp. 141–183. (<https://books.google.com.au/books?hl=en&lr=&id=s53XBgAAQBAJ&oi=fnd&pg=PA141&dq=cognitive+fit+theory&ots=Wql7BT4mo1&sig=zongTXYmENWVHZKn8zVAKWSbVnc>).
- Zhang, R., Jayawardene, V., Indulska, M., Sadiq, S., and Zhou, X. 2014. "A Data Driven Approach for Discovering Data Quality Requirements," *ICIS 2014 Proceedings*, pp. 1–10. (<http://aisel.aisnet.org/icis2014/proceedings/DecisionAnalytics/13>).

**Copyright:** © 2019 Downey, Indulska & Sadiq. This is an open-access article distributed under the terms of the [Creative Commons Attribution-NonCommercial 3.0 Australia License](https://creativecommons.org/licenses/by-nc/3.0/), which permits non-commercial use, distribution, and reproduction in any medium, provided the original author and ACIS are credited.