2017

Precision of EMR Data: The Case For a Drug and Alcohol Service

Siyu Qian
University of Wollongong, sqian@uow.edu.au

Esther Munyisia
La Trobe University

Ping Yu
University of Wollongong, ping@uow.edu.au

David M. Hailey
University of Wollongong, dhailey@uow.edu.au

David Reid
Illawarra Shoalhaven Local Health District, Drug And Alcohol Service

Publication Details

Qian, S., Munyisia, E., Yu, P., Hailey, D. M. & Reid, D. (2017). Precision of EMR Data: The Case For a Drug and Alcohol Service. 16th World Congress on Medical and Health Informatics (MedInfo2017)

Precision of EMR Data: The Case For a Drug and Alcohol Service

Abstract
Error-laden data can negatively affect clinical and operational decision making, research findings and funding allocation. This study examined the number and types of data errors in an electronic medical record (EMR) system in a Drug and Alcohol service. Specifically, errors in service data were examined. 9,379 errors were identified from ten error reports generated between March 2015 and May 2016, three months after the implementation of the EMR system. These errors were grouped into four types: mismatched data fields (60.5%), duplicate medical record error (3.2%), date/time error (8.8%) and blank field error (27.4%). The errors can be prevented by adding functions such as alert messages in the EMR system. How and why the errors occur need to be investigated in future studies.

Keywords
data:, service, alcohol, emr, drug, precision, case

Disciplines
Engineering | Science and Technology Studies

Publication Details
Qian, S., Munyisia, E., Yu, P., Hailey, D. M. & Reid, D. (2017). Precision of EMR Data: The Case For a Drug and Alcohol Service. 16th World Congress on Medical and Health Informatics (MedInfo2017)

Precision of EMR Data: The Case For a Drug and Alcohol Service

Siyu Qian[a], Esther Munyisia[b], Ping Yu[a], David Hailey[a], David Reid[c]

[a] *School of Computing and Information Technology, University of Wollongong, Wollongong, New South Wales, Australia*
[b] *School of Psychology and Public Health, La Trobe University, Melbourne, Victoria, Australia*
[c] *Illawarra Shoalhaven Local Health District, Drug and Alcohol Service, New South Wales, Australia*

Abstract

Error-laden data can negatively affect clinical and operational decision making, research findings and funding allocation. This study examined the number and types of data errors in an electronic medical record (EMR) system in a Drug and Alcohol service. Specifically, errors in service data were examined. 9,379 errors were identified from ten error reports generated between March 2015 and May 2016, three months after the implementation of the EMR system. These errors were grouped into four types: mismatched data fields (60.5%), duplicate medical record error (3.2%), date/time error (8.8%) and blank field error (27.4%). The errors can be prevented by adding functions such as alert messages in the EMR system. How and why the errors occur need to be investigated in future studies.

Keywords: Electronic Medical Records; Error; Drug and Alcohol Service

Introduction

The adoption and use of health information systems (HIS) have made large amount of digital data available for use in clinical and operational decision making [1], research [2] and funding allocation [3]. However, data errors may jeopardize the realisation of these purposes. For example, Ward et al. found that data errors in HIS time stamp can compromise the ability of an emergency department to accurately determine its operational performance [1].

Inherent data errors in clinical research databases may negatively impact the research findings [4, 5]. In a study comparing the estimation results of mortality rates using an error-free database and an error-seeded database, authors found that the overestimated mortality rates are typical results of using the latter database [6]. The estimates can be more than double the true value [6].

Error-laden data can also lead to mis-allocation of healthcare funding. In a study examing the consequences of miscoding in a hospital in Australia, authors found that about 16% of inpatient cases discharged from a specialised surgical unit during a six-month time period were miscoded. This led to an approximately $575,000 underpayment to the hospital [3].

Understanding the types and number of errors in an organization's HIS is thus useful for managers to develop strategies to prevent errors. This will ensure that the HIS truly supports organisational performance measurement, decision making, research and funding allocation.

Data errors identified by previous studies include time stamp error [1], miscoding [3], missing data [7], data transfer error [8], spelling error [9], duplicate records [10, 11], drop-down menu selection error [10] and inconsistencies between data fields [12].

Causes of these errors are related to HIS design and how HIS are used [8]. For example, the drop-down menu selection error may be caused by too many items in the drop-down list or items being too close together [13]. Spelling error may be attributed to healthcare providers documenting in a rush without proofreading [9].

Alcohol is one of the major risks for both physical and social health. Excessive consumption of alcohol can cause a wide range of harms including road accidents, domestic and public violence, family breakdown, crime, liver disease and brain damage [14]. Between 2014 and 2015, more than 115,000 Australians received over 170,000 treatment episodes from publicly-funded Drug and Alcohol (D&A) service [15].

In 2013, the public health in New South Wales (NSW), Australia moved from lump-sum funding of D&A services to funding based on treatment activities. The effectiveness of this new model of funding is substantially relied on the precision of data recorded in a HIS. Inaccurate recording may not only affect the amount of funding allocated to a D&A service, but also the managerial decisions made using this data, for example in policy making, service planning, research and education [16]. To our knowledge, however, no study has investigated the precision of D&A service data under this model of funding.

Therefore, this study aimed to investigate the number and types of errors in service data in a D&A service in NSW, Australia.

Methods

Study setting

This study was conducted in a D&A service in NSW, Australia.

Data source

An electronic medical record (EMR) system was implemented in December 2014. Service activity data were extracted by performance unit staff from the EMR system and uploaded to Web Non-Admitted Patient (WebNAP), a system that reported outpatient activities to NSW Ministry of Health for use in state health policy decision making including funding.

The WebNAP system matched the uploaded activity data with its predefined activity classification. When the data did not match the classification, the WebNAP system identified an error and automatically recorded it in an error report. For example, the WebNAP activity classification showed that D&A was a community service, therefore D&A healthcare providers should always choose community as their setting type. If a provider chose hospital, an error occurred.

Datasets

Ten error reports generated between March 2015 and May 2016, three to 14 months after implementation of the system, were provided by the information manager in the D&A service. This relates to approximately 150,000 records per year for 60,000 patients. On past experience, errors would be expected in 7% of the records with more than one error in 2% of these.

An error report contains the following data fields in an excel spreadsheet: clinic name, healthcare worker's name, client name, client medical record number, error description, appointment date/time, service date/time, referral date/time, referral receipt date/time, service type, provider type, setting type, modality of care, financial class, Department of Veterans' Affairs (DVA) card type and DVA card number.

Names of healthcare workers and clients were deidentified to maintain confidentiality.

Data analysis

Data on error description was extracted for analysis by the researchers. The errors are labelled based on the feedback from the WebNAP system and were further grouped to higher level categories. The number of each type of error was counted.

Ethics approval

Ethics approval to conduct this study was granted by the joint Health and Medical Human Research Ethics Committee of the University of Wollongong and the Illawarra & Shoalhaven Local Health District. Access to the error reports was granted by the service manager of the D&A service. Consent was obtained from healthcare providers for their error reports to be used in this study.

Results

Types of Errors in D&A Service Data

The identified errors were grouped into four types: 'mismatched data fields', 'duplicate record', 'date/time error' and 'blank field' (see Table 1).

There were two types of 'mismatched data fields'. One was 'service option error' and the other was 'DVA information error'. 'Service option error' occurred when one of the three data entry fields- provider type, modality of care and setting type- did not match what was set up in the WebNAP classification. Provider type described the job role of a healthcare provider e.g. a registered nurse, a psychologist or a counsellor. Modality of care was the means for delivery of a service e.g. telephone, email or face-to-face meeting between a healthcare provider or a group of healthcare providers and a client. Setting type was the location where the service was provided e.g. hospital, community or home.

The 'DVA information error' occurred when the information provided in the two data entry fields - DVA card details and financial classification - did not match with each other. There were two situations. One was 'DVA card details supplied but financial classification is not DVA' and the other was 'financial classification is DVA but missing DVA card details'.

'Duplicate medical record' occurred when more than one encounter was created for a client in the EMR system. However, a client could only have one active encounter at a time.

The 'date/time error' was related to referral date/time, referral receipt date/time, service start date/time and service end date/time. These date/times include time information. The logical order of these date/times must be as follows: the referral date/time should be earlier than the referral receipt date/time which should be earlier than the service start date/time and the service end date/time. When this order was turned around, an error occurred.

The last data error type was 'blank field' which occurred when a data entry field was blank. The identified blank fields were provider type, funding source, financial classification and address fields including post code, suburb and street.

Table 1– Number and types of errors in service data in the EMR system in a D&A service.

Error types	Number	%
Total	**9379**	**100**
Mismatched data fields	**5675**	**60.5**
Service option error (provider type, modality of care or setting type does not match what is in WebNAP classification.)	5650	60.2
*DVA information error (DVA card details supplied but financial classification is not DVA, or financial classification is DVA but missing DVA card details)	25	0.3
Duplicate medical record A second encounter is created in the EMR, but a client can only have one active encounter at a time	**303**	**3.2**
Date/time error	**834**	**8.8**
Service start date/time is after service end date/time	591	6.3
Referral receipt date/time is after service date/time or before referral date/time	145	1.5
Referral date/time is after service start date/time	98	1
Blank field	**2567**	**27.4**
Blank provider type	2473	26.4
Blank post code	36	0.4
Blank suburb	21	0.2
Blank street	21	0.2
Blank funding source code	10	0.1
Blank financial classification	6	0.1

*DVA: Department of Veterans' Affairs

Number of Errors in Each Error Type

Overall, 9,379 errors were identified from the error reports (see Table 1). 60.5% of the errors were 'mismatched data fields', of which majority was 'service option error'. 27.4% were 'blank field' with 'blank provider type' being the error that occurred most frequently. 'Date/time error' accounted for 8.8% of the total number of errors. 'Service start date/time is after service end date/time' was the major error which accounted for 6.3%. 'Duplicate medical record' was the least frequently occurring error which accounts for 3.2% of the total errors.

Discussion

As errors in operational data can be pervasive in the immediate period after the implementation of HIS [7], this study focused on analysing error reports for an EMR system three to 14 months after its implementation. The large amount of errors found in this study may be due to the learning curve of healthcare providers. Further study will investigate trend in data errors over time.

We classified errors into four types. This is useful for investigating causes of error and developing different prevention and mitigation strategies required for different error types.

Causes of these errors might be related to the EMR system design issue, how the system was used, the environment in which it was used or a combination of them. For example, this study found that 'service option error' was the most common error in the D&A service data and wrong selection of 'setting type' was one immediate reason for it. In the WebNAP reporting system, the D&A service was mapped as a community setting. That means all healthcare providers working in the D&A service must select 'community' for setting type, regardless where the service was provided. This would ensure that funding was allocated to D&A service.

Some of the healthcare providers, although employed to work for the D&A service, were responsible for providing this specialist service in a hospital setting. A D&A healthcare provider might select 'hospital' for the setting type, because he or she provided the service there. This would result in a 'service option error' because the setting type of 'hospital' was not mapped in the WebNAP system for the D&A service. If this error was not corrected, funding would be misplaced to the hospital setting, instead of D&A service.

The term 'setting type' may confuse healthcare providers, especially new employees, on whether it means the location of service provided or the healthcare service by which the healthcare provider was employed. This could be one reason for the error to occur.

Another reason might be the design issue of the WebNAP system. The system used 'setting type' to determine which healthcare service would get funding. Actually, it could have used the information about the healthcare service by which the healthcare provider was employed to determine the correct 'setting type'. This suggested that the system designer did not really understand how healthcare providers worked and how funding was allocated. Extraction of correct data elements would eliminate the error and ensure correct funding allocation.

A mix of system design issue and the environmental condition under which end users use the system may contribute to the occurrence of error. This study found that 'date/time error' accounted for almost 9% of the total number of errors. Date/time data were required from three forms that D&A healthcare providers used in the EMR system: intake form, assessment form and clinical note. The intake form was used when a client first contacted the D&A service. The assessment form was used to assess the person after intake. The clinical note was used in the subsequent visits of the client. The entire journey of the person with the D&A service from intake to discharge is an 'encounter'. Each contact of the person with the D&A service may contain serveral 'services', e.g. counselling, rehabilitation or supervised medication administration. The three forms were used at different stages of the encounter. The intake form and the assessment form were used in the beginning of this encounter. The clinical note was used multiple times until the closure of the encounter.

Within each form, the service start date/time must be before the service end date/time. However, the way to record service start and end date/time is different among the three forms (see Table 2). The intake form automatically populated service *start* date/time from the computer, but required healthcare providers to manually enter *end* date/time. The assessment form required both *start and end* date/time to be manually entered. The clinical note required a manual entry of *start* date/time, but automatically filled in the *end* date/time. This mixed ways of recording date/time opens opportunities for error.

Another reason for the error to occur might be the environmental condition under which healthcare providers recorded data. Sitting with a client and recording data in a computer at the same time may increase the probablity of making error. Automatic time recording may have better data accuracy than manual time recording [17]. A check of date/time by the EMR system at the data entry stage may also help to prevent the erroneous data from being recorded.

Table 2– How service start and end date/time is recorded.

Forms	Service start date/time	Service end date/time
Intake form	Auto-populated, can be adjusted manually	Manually entered
Assessment form	Manually entered	Manually entered
Clinical note	Manually entered	Auto-populated, can be adjusted manually

A lack of alert message function in the system may open opportunity for error. For example, although the percentage of 'duplicate medical record' is not high in this study, duplicate records can mislead healthcare providers in clinical decision making because they may miss important information that exists in a different record [11]. This error may also cause confusion in information retrieval [18]. Disruptive pop-up alert message [4] built into an EMR system may help to decrease this error.

The 'blank provider type' is the second common error found in this study. This error might be caused by system dysfunction. It may also be due to the inability of the system administrator to keep up with the workload of correctly mapping the providers in the WebNAP system–because of the high turnover of D&A staff, or lack of information on changes in roles for staff members or new staff in the service. Further study is needed to investigate how and why this error occurred.

Since the service data were first extracted by the performance unit staff before submitting to WebNAP, it was possible that the errors reported in this study could also be injected during the process of data extraction.

With the increased use of EMR, research using EMR data has been prospering, for example in EMR phenotyping [19], clinical workflow modelling [20] and disease prediction [21]. The strength of evidence from the secondary analysis of EMR data can be hindered by errors contained in these data [4, 5]. In the case mentioned in the introduction section, the study comparing the estimation results of mortality rates using an error-free database and an error-seeded database, the analysis result using the error-laden data was more than double the true results using the error-free data [6]. This shows the negative impact of erroneous data in reducing accuracy of data

analytics and lead to invalid findings. Therefore, data errors and their causes need to be identified, reduced or eliminated to ensure high data quality to provide accurage evidence for research and health decision making [22].

Limitation

We did not analyse all the error reports generated after the implementation of the EMR system. This was because there was no health information manager at the D&A service for a period of three months. Consequently no error reports were downloaded in this period. However, the errors made during that time period rolled over to the following months' reports, so we think that we had all errors analysed. Directly interviewing the healthcare providers who made the errors would give further insights about the reasons for the errors made, which would be beneficial for learning to prevent errors in the EMR system.

Conclusion

Error-laden data can jeopardize clinical and operational decision making [1], research findings [6] and funding allocation [3]. To manage errors, it is paramount for D&A service to understand the nature and extent of error and the environment that induce error [23].

This study investigated the number and types of errors in service data in a D&A service. The identified errors were grouped into four error types: 'mismatched data fields' (60.5%), 'duplicate medical record' (3.2%), 'date/time error' (8.8%) and 'blank field' (27.4%). The top three most frequent sub errors were 'service option error', 'blank provider type' and 'service start date/time is after service end date/time'. The results from this study underscore the importance of understanding errors in EMR data. Further study will investigate the trend in data errors overtime and how and why errors occur in the EMR system.

Acknowledgements

We are grateful for the healthcare providers at the D&A service who provided consent for their error reports to be used in this study. We thank the director and service managers for their support and approval of this study. The health information manager Mr. Will Fernadez is acknowledged for supporting the data collection process.

References

[1] M.J. Ward, W.H. Self, and C.M. Froehle, Effects of Common Data Errors in Electronic Health Records on Emergency Department Operational Performance Metrics: A Monte Carlo Simulation, *Academic Emergency Medicine* **22** (2015), 1085-1092.

[2] M.A. Wildeman, J. Zandbergen, A. Vincent, C. Herdini, J.M. Middeldorp, R. Fles, O. Dalesio, E. Van der Donk, and I.B. Tan, Can an online clinical data management service help in improving data collection and data quality in a developing country setting?, *Trials* **12** (2011).

[3] P. Cheng, A. Gilchrist, K.M. Robinson, and L. Paul, The risk and consequences of clinical miscoding due to inadequate medical documentation: a case study of the impact on health services funding, *Health Information Management Journal* **38** (2009), 35-46.

[4] K. Haerian, J. McKeeby, G. Dipatrizio, and J.J. Cimino, Use of clinical alerting to improve the collection of clinical research data, *AMIA. Annual Symposium proceedings / AMIA Symposium. AMIA Symposium* **2009** (2009), 218-222.

[5] S. Goldberg, A. Niemierko, and A. Turchin, Analysis of data errors in clinical research databases, in: *AMIA*, Citeseer, 2008.

[6] S. Gallivan, J. Stark, C. Pagel, G. Williams, and W.G. Williams, Dead reckoning: can we trust estimates of mortality rates in clinical databases?, *European Journal of Cardio-Thoracic Surgery* **33** (2008), 334-340.

[7] M.J. Ward, C.M. Froehle, K.W. Hart, and C.J. Lindsell, Operational data integrity during electronic health record implementation in the ED, *American Journal of Emergency Medicine* **31** (2013), 1029-1033.

[8] D.F. Sittig and H. Singh, A red‐flag‐based approach to risk management of EHR‐related safety concerns, *Journal of Healthcare Risk Management* **33** (2013), 21-26.

[9] F. Hussain and U. Qamar, Identification and correction of misspelled drugs' names in electronic medical records (EMR), in: *ICEIS 2016 - Proceedings of the 18th International Conference on Enterprise Information Systems*, 2016, pp. 333-338.

[10] K.E. Walsh, W.G. Adams, H. Bauchner, R.J. Vinci, J.B. Chessare, M.R. Cooper, P.M. Hebert, E.G. Schainker, and C.P. Landrigan, Medication errors related to computerized order entry for children, *Pediatrics* **118** (2006), 1872-1879.

[11] A.B. McCoy, A. Wright, M.G. Kahn, J.S. Shapiro, E.V. Bernstam, and D.F. Sittig, Matching identifiers in electronic health records: implications for duplicate records and patient safety, *BMJ quality & safety* **22** (2013), 219-224.

[12] H. Singh, S. Mani, D. Espadas, N. Petersen, V. Franklin, and L.A. Petersen, Prescription errors and outcomes related to inconsistent information transmitted through computerized order entry: a prospective study, *Archives of Internal Medicine* **169** (2009), 982-989.

[13] R. Khajouei and M. Jaspers, The impact of CPOE medication systems' design aspects on usability, workflow and medication orders, *Methods of Information in Medicine* **49** (2010), 3.

[14] Australian Institute of Health and Welfare, Alcohol and other drugs, 2016. Available: http://www.aihw.gov.au/alcohol-and-other-drugs/

[15] Australian Institute of Health and Welfare, Alcohol and other drug treatment services (AODTS), 2016. Available: http://www.aihw.gov.au/alcohol-and-other-drugs/aodts/

[16] B.G. Audit Commission, Data quality matters: learning from audit commission work on data quality in local public services, in, 2015.

[17] R.A. Marjamaa, P.M. Torkki, M.I. Torkki, and O.A. Kirvelä, Time accuracy of a radio frequency identification patient tracking system for recording operating room timestamps, *Anesthesia and Analgesia* **102** (2006), 1183-1186.

[18] S. Qian and P. Yu, Fitting clinical workflow: The case for wound care in a residential aged care home, in: *Studies in Health Technology and Informatics,* H. Grain, F. Martin-Sanchez, and L.K. Schaper, eds., IOS press, Melbourne, Australia, 2014, pp. 130-136.

[19] T. Zheng, W. Xie, L. Xu, X. He, Y. Zhang, M. You, G. Yang, and Y. Chen, A machine learning-based framework to identify type 2 diabetes through electronic health records, *International Journal of Medical Informatics* **97** (2017), 120-127.

[20] Y. Chen, W. Xie, C.A. Gunter, D. Liebovitz, S. Mehrotra, H. Zhang, and B. Malin, Inferring Clinical Workflow Efficiency via Electronic Medical Record Utilization, *AMIA Annual Symposium Proceedings* **2015** (2015), 416-425.

[21] R. Kop, M. Hoogendoorn, L.M. Moons, M.E. Numans, and A. ten Teije, On the advantage of using dedicated data mining techniques to predict colorectal cancer, in: *Conference on Artificial Intelligence in Medicine in Europe*, Springer, 2015, pp. 133-142.

[22] S. Hoffman and A. Podgurski, Big Bad Data: Law, Public Health, and Biomedical Databases, *Journal of Law, Medicine and Ethics* **41** (2013), 56-60.

[23] R.L. Helmreich, On error management: lessons from aviation, *British Medical Journal* **320** (2000), 781.

Address for correspondence
Associate Professor Ping Yu, School of Computing and Information Technology, University of Wollongong, Wollongong, New South Wales, 2522, Australia. Email: ping@uow.edu.au. Work phone: +61 02 4221 416.