Implications and Modelling of Data Quality on Confidence of Clinical Decision Support: a Conceptual Framework

George DESPOTOU[a,1], Richard W. JONES[b], Konstantinos KATZIS[c], Eda BILICI[a], Zhao LEI[a] , Sarah LIM CHOI KEUNG[a] , Theodoros N. ARVANITIS[a].
[a] *Institute of Digital Healthcare, WMG, University of Warwick, UK.*
[b] *School of Engineering, University of Warwick, UK.*
[c] *Department of Computer Science and Engineering, European University, Cyprus.*

Abstract. Integrated care paradigms depend on multiple sources of data. The quality of data used in decision-making will ultimately affect the delivered care to the patient. Quality includes several dimensions, which may affect the result. This paper presents how data quality dimensions may affect the delivered service, and propose a conceptual framework for the classification of confidence in data used in clinical decision-making for integrated care.

Keywords. Clinical decision support, decision-making confidence, data quality, integrated care

Introduction

In recent years, there has been the impetus for patients to self-manage their healthcare, for both physical and mental health issues. Advances in the way humans interact with computers, such as gesture and voice recognition, has allowed easier modes of interaction with IT applications, requiring very little technical skill. Furthermore, the evolution of mobile devices has provided a platform for integrating sensors, as well as applications personalised to the user [1, 2, 3]. The ability of the patient to interact with the healthcare service, has resulted in the increasing design and adoption of new paradigms of healthcare. These paradigms include active consideration of data communicated by the patient, such as self-taken measurements, patient reports, as well as entries by healthcare professionals who have examined the patient. Data can be produced by continuous or intermittent monitoring, and constitute part of the patient's personal healthcare record (PHR). This enriched PHR enabled to make decisions about the care plan of the patient.

The quality of data used in decision-making will ultimately affect the delivered care to the patient, and includes several dimensions. For example, consider a healthcare expert using out of date values to make an assessment. This may result in a decision based on data that are no longer valid, hence potentially adversely affecting the patient. The heterogeneous sources in the new integrated healthcare paradigms, may affect the quality dimensions of data. Different manufacturers of devices used for self-quantification, skill

1 Corresponding Author: George Despotou: g.despotou@warwick.ac.uk

of the patient to use devices and report values, as well different IT applications, may compromise data quality. This affects the confidence with which decisions are made.

This paper takes an example of clinical decision making, amenable to new healthcare paradigms, and illustrates how data can be evaluated using the quality dimensions. The work then proposes a classification, which can be used by healthcare professionals, to understand the impact quality may have on the confidence with which they make decisions.

1. Data considerations in clinical decision making

Figure 1 presents an example excerpt of a clinical guideline, – guidance CG127 from the UK National Institute for Health and Care Excellence (NICE) for diagnosis and management of hypertension disorder, which already relies on multiple data sources, including home monitoring [4].

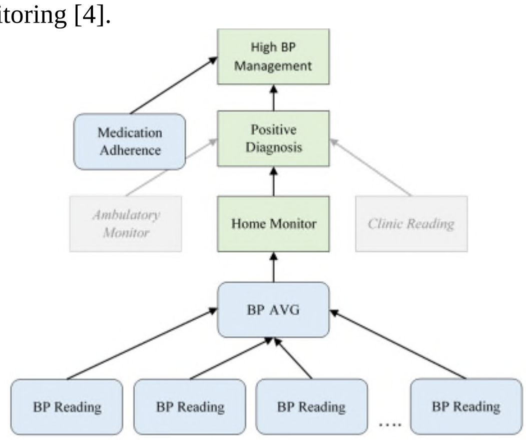

Figure 1. Data representation of guideline UK NICE – CG127

The disorder (excerpt in this example) can be diagnosed in three different ways: a measurement at a clinic, through an ambulatory monitor (i.e. a 24/7 wearable monitor), as well as patient home-based monitor. A positive diagnosis using home monitors, requires multiple readings in specific intervals (see Fig.2), creating an overall average measurement. This average measurement value can then be used to diagnose the disorder. This intent of these requirements is to remove potential deviating readings due to the equipment, its use, or readings when the state of the patient is not known (e.g., after exercise), which may skew the end-result. Nevertheless, even with these requirements, readings can be divergent due to common failures, such as skill of measurement, as well as a consistently uncalibrated equipment.

Use the average value of at least 14 measurements taken during the person's usual waking hours to confirm a diagnosis of hypertension. When using home blood pressure monitoring (HBPM) to confirm a diagnosis of hypertension, ensure that: for each blood pressure recording, two consecutive measurements are taken, at least 1 minute apart and with the person seated and blood pressure is recorded twice daily, ideally in the morning and evening and blood pressure recording continues for at least 4 days, ideally for 7 days. [UK NICE - CG127]

Figure 2. BP data requirements for UK NICE – CG127

The decision making regarding the disorder is further affected by patient originating data, during the disorder management stage. Moreover, the healthcare professional will need to evaluate the effectiveness of the patient's medication under the current care plan. This implies an understanding of medication adherence, which is something not considered explicitly by the guidelines.

For people who are already having treatment with bendroflumethiazide or hydrochlorothiazide and whose blood pressure is stable and well controlled, continue treatment with the bendroflumethiazide or hydrochlorothiazide. [UK NICE - CG127]

Figure 3. Patient data dependency for hypertension management UK NICE – CG127

In a healthcare service that expects the patient to report adherence (e.g., via a mobile application), adherence data will have a crucial role on subsequent decision-making. The implications of this become more pronounced in complex services, such as patient-centric, multi-morbidity management, which may have to integrate multiple guidelines.

2. Impact of data quality attributes on healthcare service

Data possess a number of attributes, representing different issues, which may affect the purpose for which they are used; in this case the function of clinical decision support [5]. Table 1 illustrates how a representative set of these attributes can affect how data are used by a service.

Table 1. Data quality dimensions applied to the hypertension disorder example

Data Attribute	Example Issue with Attribute
Accuracy	BP reading divergent from actual BP value of the patient
Validity	BP reading captured in an unexpected way such as different units, or extreme values (e.g. swapping systolic with diastolic reading)
Reliability	Equipment used not fit for purpose, or patient not trained to capture BP reading accurately
Timeliness	BP reading older than required, taken at wrong interval
Relevance	Use of BP instead of BPAVG
Completeness	Stored value missing units or clinical identifier

Ultimately, data may contribute to only a few risks, in this case: 1) wrong diagnosis (false positive, false negative) and 2) mismanagement of treatment. Understanding the data attribute that may be the reason for concern, allows us to draw more meaningful conclusions, about how we can use them (i.e. data). For example, consider the timeliness of BP not being according to the guidelines (i.e. >1 min between consecutive values), but taken less than a minute apart. If acceptability of that value is seen as binary it will be rejected; however, a qualitative evaluation of the value may still result in meaningful conclusions.

3. Qualitative confidence assessment

Simply masking data with decisions (e.g., rejection of BPAVG due to a value not meeting the quality criteria) may not offer sufficient resolution to healthcare professionals, often in a position where they need to make decisions based on imperfect knowledge [6] Context of how the information is used is also important. For example, lack of confidence (due to insufficient quality) may be acceptable if the decision is to ask a patient, based on their PHR values, to visit a doctor, than if the decision is about which medication is more suitable for a patient. A framework for classifying and directing professionals to evaluating quality and subsequent confidence in a decision can be a useful tool. Figure 4 proposes a traffic light classification for data, based on the suitability of the data for intended use (e.g., timeliness requirements from a guideline), and the information (or lack) of about the quality attributes of the data. The classification accepts use of data with other than intended quality, given some usefulness for the patient. Additionally it considers whether there is sufficient information about data quality.

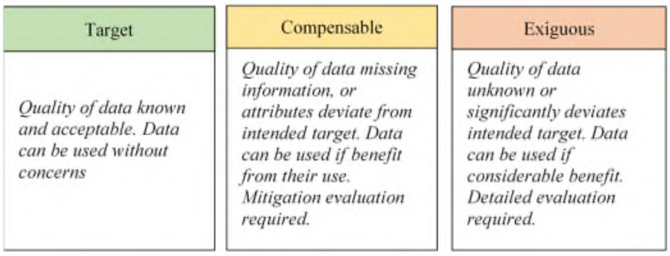

Figure 4. Traffic light confidence framework

Confidence gaps, as well as corrective action taken, are attributes that can be recorded in a PHR (future aims of this work). This allows traceability on what professionals perceived acceptable trade-offs of data quality, as well as provide a clear audit trail for subsequent patient examinations.

Conclusions

New care models employ decision making dependant on multiple integrated sources, of not well known pedigree and quality standards. Data quality consists of a number of dimensions that may affect the service to the patient, also resulting in risks. Use of data for decision making should enable healthcare professionals to understand the trade-offs of data quality and incorporate this knowledge to their decision making. the authors are working towards a EHR/EPR compliant IT implementation of the framework.

References

[1] Research2Guidance, ‘mHealth App Developer Economics 2016’, 2016

[2] FDA, Mobile Medical Applications, Guidance for Industry and Food and Drug Administration Staff, 2015, Available at: https://www.fda.gov/downloads/MedicalDevices/.../UCM263366.pdf

[3] G. Bafhtiar1, V. Bodinier, G. Despotou, M.T.Elliott, N. Bryant, T.N. Arvanitis, ‘Providing Patient Home Clinical Decision Support using Off-the-shelf Cloud-based Smart Voice Recognition. WIN Health Informatics Network Annual Conference 2017.

[4] UK NICE, Hypertension in adults: diagnosis and management, Clinical guideline [CG127, Available at: https://www.nice.org.uk/guidance/cg127

[5] Cai, L. & Zhu, Y., (2015). The Challenges of Data Quality and Data Quality Assessment in the Big Data Era. Data Science Journal. 14, p.2. DOI: http://doi.org/10.5334/dsj-2015-002

[6] E. H. Shortliffe and B. G. Buchanan, “A model of inexact reasoning in medicine,” Mathematical Biosciences, vol. 23, no. 3–4, pp. 351–379, Apr. 1975.