

LETTER

Open Access

2015, big data in healthcare: for whom the bell tolls?

Sven Van Poucke^{1*}, Michiel Thomeer² and Admir Hadzic³

The health care sector generates bountiful data around the clock, which can paradoxically complicate our quest for information, knowledge, and 'wisdom' [1]. It may be prudent that medical end-users consider seriously a fundamental change that would allow us to gain full value from the 'big data' that the health care section is generating [2]. Proponents of the big data revolution suggest that the value for physicians rests on the added information provided by big data analysis. Indeed, supplementary information could clarify areas for improvement, such as optimization of treatments, reduced adverse events and readmission rates, earlier identification of those patients whose health is worsening, and more efficient identification of populations in need. Recent cloud computing has even turned computing and software into commodity services, and such big data processing seems to be forging a technology revolution [3,4]. However, opponents of the big data revolution speculate that validation and impact analyses of big data in health care are still in their infancy, and approaches such as Google's baseline study may thus not be effective in preventing disease, and possibly even lead to unnecessary, if not harmful, interventions [5].

The value of any kind of data is greatly enhanced when it exists in a form that allows for integration with other data [6]. One problem with large data sets in general is the risk for 'GIGO' - garbage in, garbage out - that requires very careful and thoughtful investigation to rule out the many errors of large-scale data capture before any of it can be used. Thus, an essential step for data integration is the annotation of multiple bodies of data using common controlled vocabularies or 'ontologies' that incorporate accurate representations of biological reality [7]. Data mining in health care is not new, and initiatives for data acquisition and analysis, storage and retrieval have all been presented before [8,9]. Yet, to our knowledge,

subcommittees addressing ontology have not been established by any medical specialty.

As clinicians, we apply general principles of risk stratification and risk modification to individual patients based on our education and experience. The proliferation of biomedical research makes it difficult to keep abreast of current knowledge, so clinical decision support technologies that are based on data mining techniques are knocking at our doors. Although their implementation seems inevitable, the lack of standardization continues [9]. A dramatic paradigm shift toward controlled ontologies is needed in order to optimize the technologies that integrate big data into medical decision making and practice.

Competing interests

SVP and MT declare that they have no competing interests. AH consulted and advised for Skypharma, GE, Sonosite, Codman & Shrutleff, Inc. (Johnson and Johnson), Cadence, Pacira Pharmaceuticals, Baxter and BBraun Medical. Research funding from Glaxo Smith-Kline Industries, Pacira, Baxter. AH receives royalty income from BBraun Medical.

Authors' contributions

SVP, MT, and AH contributed equally in writing this letter. All authors read and approved the final manuscript.

Acknowledgements

All authors (SVP, MT, AH) declare that this letter was written with no funding. No role was played by any funding body in the design, collection, analysis, and interpretation of data; in the writing of the manuscript; and in the decision to submit the manuscript for publication.

Author details

¹Department of Anesthesiology, Intensive Care, Emergency Medicine, Pain Therapy, Ziekenhuis Oost-Limburg, Schiepse Bos 6, 3600 Genk, Belgium.

²Department of Respiratory Medicine, Ziekenhuis Oost-Limburg, Schiepse Bos 6, 3600 Genk, Belgium. ³Department of Anesthesiology, Ziekenhuis Oost-Limburg, Schiepse Bos 6, 3600 Genk, Belgium.

Published online: 02 April 2015

References

1. Henry NL. Knowledge management: a new concern for public administration. *Public Administration Rev.* 1974;34:189-96.
2. Groves P, Kayyali B, Knott D, Van Kuiken S. The 'big data' revolution in healthcare. Accelerating value and innovation. McKinsey Global Institute: New York, NY; 2013.

* Correspondence: svanpoucke@gmail.com

¹Department of Anesthesiology, Intensive Care, Emergency Medicine, Pain Therapy, Ziekenhuis Oost-Limburg, Schiepse Bos 6, 3600 Genk, Belgium
Full list of author information is available at the end of the article

3. Pathak J, Shah ND. Why health care may finally be ready for big data. *Harvard Business Review*. 2014. <https://hbr.org/2014/12/why-health-care-may-finally-be-ready-for-big-data>.
4. Herland M, Khoshgoftaar TM, Wald R. A review of data mining using big data in health informatics. *J Big Data*. 2014;1:2.
5. Diamandis EP. The hundred person wellness project and Google's baseline study: medical revolution or unnecessary and potentially harmful over-testing? *BMC Med*. 2015;3:5.
6. Daghavan D. Big data: not really the same as level 1 data. *Oncology (Williston Park)*. 2015;29:70, 72.
7. Smith B, Ashburner M, Rosse C, Bard J, Bug W, Ceusters W, et al. The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. *Nat Biotechnol*. 2007;25:1251–5.
8. Lee J, Scott DJ, Villarroel M, Clifford GD, Saeed M, Mark RG. Open-access MIMIC-II database for intensive care research. *Conf Proc IEEE Eng Med Biol Soc*. 2011;2011:8315–8.
9. Dejam A, Malley BE, Feng M, Cismondi F, Park S, Samani S. The effect of age and clinical circumstances on the outcome of red blood cell transfusion in critically ill patients. *Crit Care*. 2014;18:487.