Big data and visual analytics in anaesthesia and health care†

A. F. Simpao†*, L. M. Ahumada2 and M. A. Rehman†

1 Department of Anesthesiology and Critical Care, Perelman School of Medicine at the University of Pennsylvania and the Children’s Hospital of Philadelphia, 3401 Civic Center Boulevard, Suite 9329, Philadelphia, PA 19104-4399, USA
2 Enterprise Analytics and Reporting, The Children’s Hospital of Philadelphia, 1300 Market Street, Room W-8006, Philadelphia, PA 19107-3323, USA
† Corresponding author: E-mail: simpao@email.chop.edu

Editor’s key points

- Electronic health records have led to the rapid accumulation of patient-related computer data (big data).
- Data analytics applied to big data are facilitating outcomes research and quality improvement efforts.
- Visual analytics based on big data facilitate hypothesis generation and real-time clinical decision making.

While there is no rigorous definition of ‘big data’, the term is applied customarily to data sets whose size, complexity, and dynamic nature pose significant challenges to traditional data-processing tools.1, 2 Repeated observations over time and space generate most big data; examples include worldwide users’ Internet search engine queries (e.g. Google), e-commerce browsing and transactions (e.g. Amazon), and genomic sequencing in biomedical research.3, 4 While the term ‘big data’ does not apply specifically to the medical field, the role of big data has become increasingly prominent in the constantly evolving world of medicine.

Advances in computer technology, patient monitoring systems, and electronic health record systems have enabled rapid accumulation of patient data in electronic form (i.e. big data). Organizations such as the Anesthesia Quality Institute and Multicenter Perioperative Outcomes Group have spearheaded large-scale efforts to collect anaesthesia big data for outcomes research and quality improvement. Analytics—the systematic use of data combined with quantitative and qualitative analysis to make decisions—can be applied to big data for quality and performance improvements, such as predictive risk assessment, clinical decision support, and resource management. Visual analytics is the science of analytical reasoning facilitated by interactive visual interfaces, and it can facilitate performance of cognitive activities involving big data.

Ongoing integration of big data and analytics within anaesthesia and health care will increase demand for anaesthesia professionals who are well versed in both the medical and the information sciences.

Keywords: decision support systems, clinical; electronic health records; integrated advanced information management systems; medical informatics

Big data and analytics in health care

Increasing recognition of the potential utility of big data in health outcomes research has created an impetus to collect and pool EHR data in national data sets. These large data sets provide access to information on rare conditions and outcomes that are otherwise difficult to study without robust sample sizes. Examples of health-care-related big data efforts include collaborative nationwide database projects sponsored by the US Department of Health and Human Services Agency for Healthcare Research [http://www.ahrq.gov (accessed December 18, 2014)] and the UK Health and Social Care Information Centre [http://www.hscic.gov.uk (accessed December 18, 2014)]. Interest in population-level, epidemiological anaesthesia-related research has led to the creation of large anaesthesiology-specific databases, such as the Multicenter Perioperative Outcomes Group (MPOG),9 the Anesthesia Quality Institute’s (AQI) National Anesthesia Clinical Outcomes Registry (NACOR),10 the Society for Ambulatory Anesthesia database, and the Pediatric Regional Anesthesia Network.11 The realization of big data’s potential benefits (patient safety, evidence-based guidelines, and cost containment) requires validation of data quality followed by analysis of the data to enhance decisions, processes, and policies.12, 13
This approach to data analysis and application is termed ‘analytics’ making decisions by the systematic use of data combined with quantitative and qualitative analysis.\textsuperscript{14} Analytics and big data are used in health-care applications outside of traditional inpatient and outpatient EHRs, such as wearable health-monitoring systems (WHMSs) that patients use at home. Wearable health-monitoring systems consist of various sensors, actuators, and multimedia devices, enabling low-cost, non-invasive continuous monitoring of indoor and outdoor health, activity, mobility, and mental status.\textsuperscript{15} Wearable health-monitoring systems provide large amounts of continuously acquired, multivariate physiological data that necessitate data mining and analytics.\textsuperscript{16} Such systems might reduce health-care costs through disease prevention and enhance quality of life with improved disease management. In addition, WHMSs can be tailored to specific uses, such as monitoring physical activity,\textsuperscript{17} the elderly in nursing homes,\textsuperscript{18} or individuals with dementia or Parkinson’s disease.\textsuperscript{19} The feasibility of using WHMSs in the inpatient setting has also been investigated.\textsuperscript{20}

Analytics methods such as mathematical and algorithmic-based data processing, text mining, and natural language processing have been used to analyse and derive insight from data across a wide spectrum of health-care fields.\textsuperscript{21} Text mining and natural language processing retrieve and extract salient information from unstructured, semi-structured, and structured text, such as mining research articles to facilitate cancer research or detecting surgical site infections by mining EHR text, such as mining research articles to facilitate cancer remission from unstructured, semi-structured, and structured natural language processing retrieve and extract salient information to oncologists to allow physicians to tailor care for individual patients,\textsuperscript{25} better targeting of disease management and innovative patient care approaches,\textsuperscript{26} formation of searchable and accessible collections that are usable for large-scale health analytics,\textsuperscript{27} generating life-expectancy indices from EHR data,\textsuperscript{28} and real-time analysis of physiological data streams in the neonatal intensive care unit for earlier detection of deteriorating medical conditions.\textsuperscript{29} Potential perioperative applications of such methods include data aggregation combined with real-time analytics of intraoperative physiological data to guide point-of-care anaesthetic decisions.

There are powerful analytic tools that use big data in clinical decision support (CDS) systems within EHRs to help anaesthetists and other clinicians make more personalized, evidenced-based decisions; such tools can extract relevant information and provide insights in real time.\textsuperscript{30} Clinical decision support systems have been shown to reduce errors and improve clinical outcomes in certain settings, such as paediatric intensive care units,\textsuperscript{31,32} and CDS can improve performance on quality and process measures in the perioperative and medical imaging settings.\textsuperscript{33,34} Analytics can also enhance rudimentary CDS systems that are already in place (e.g. simple rules-based drug-drug interaction CDS systems that require manual reclassification as a result of suboptimal sensitivity)\textsuperscript{35,36} by using analytics techniques to query and mine the EHR for meaningful connections, and then integrating the knowledge-based rules with EHR data to improve the CDS system.\textsuperscript{37} Challenges and concerns abound regarding implementation of real-time CDS systems, such as loss of autonomy of clinicians, risk to patient privacy, and potentially basing recommendations on faulty ‘real-life’ data.\textsuperscript{38,39} Novel systems might also lack utility compared with conventional methods, such as a CDS system to mitigate postanaesthesia care delays.\textsuperscript{40} While these concerns warrant consideration, CDS systems that implement analytics techniques and predictive modelling have the potential to improve outcomes, enhance patient experiences, and reduce health-care costs.\textsuperscript{33}

Finance management is another use for analytics in health care, such as identification of billing anomalies (i.e. revenue leakage), resource allocation, cost cutting, and improving revenue. Most health-care organizations find billing anomalies using rules-based approaches and manual audits, yet these methods are time consuming and error prone. A combination of both advanced analytics approaches (e.g. machine learning and predictive modelling) and review by human billing experts can form a ‘dual approach’ to find patterns in billing records that are most likely to be associated with missing or erroneous charges; one health system used this method and reduced audit expenses by 75%.\textsuperscript{41} Analytics can be used to determine which factors within a broad network of departmental activities impact a patient’s hospital stay, and these activities can be scrutinized using analytics methods (e.g. scheduling). For example, using analytics, one institution determined that inefficiencies in the radiology department adversely extended lengths of stay beyond initial estimates. The hospital then advocated the use of proactive analytics assessments of networks of activities to enhance organizational efficiency.\textsuperscript{42} Cost-cutting measures (e.g. reducing readmission rates)\textsuperscript{43} have been targeted and measured using analytics, such as the use of EHR data to predict which patients are at a greater risk for using more resources than their peers. This is accomplished by entering EHR data into a model that is then processed by various analytic techniques in order to stratify patients as high risk.\textsuperscript{44} Revenue-generating uses of analytics include using marketing analytics and graphical information systems to target catchment areas.\textsuperscript{45,46}

A logical step after proliferation of EHR systems is application of analytics to gather meaningful information from clinical data repositories, yet even proponents of health-care big data have raised valid concerns. One challenge presented by data is not only quantitative (i.e. managing massive data files), but also qualitative; while considerable data can be generated in a rapid fashion by a hospital full of electronic patient monitors, such monitors (e.g. pulse oximetry, capnography) are subject to artifact generation.\textsuperscript{47} Application of analytics to data rife
with artifacts can produce faulty conclusions ('garbage in, garbage out'); one potential solution is the application of analytics such as machine learning and neural networks during data auditing to detect errors before analysis. Consideration must be taken for other unintended consequences that can arise from the use of analytics methods, such as re-identification of anonymized data and exclusion of disadvantaged populations.26 48

Visual analytics and big data in anaesthesia and health care

Big data pose computing challenges because of their rapid velocity, immense volume, and wide variety.1 Visual analytics (VA) tools are a category of computational tools that integrate data analytics with interactive visual interfaces, and can be used to navigate and manipulate big data.24 49 Visual analytics tools enable investigative analysis and hypothesis generation by showing connections between entities in a manner that facilitates understanding and guides selection and application of other analytics techniques.50 51 Visual analytics tools offer three main benefits compared with traditional database queries: first, the user can explore big data in a self-service ‘point-and-click’ fashion as opposed to writing database queries manually; second, complex ideas can be communicated with clarity and efficiency in visual graphs rather than the tabular data output from a traditional database query; and third, VA tools can display large amounts of filtered data in near real time.52

Several steps are required to visualize big data using VA tools. First, the data are stored in a dimensional database model that uses online analytical processing systems that were developed specifically to analyse very large data sets. Dimensional models create a unique fact table that contains all potential data transactions in addition to filters (dimensions) used to associate facts and measures throughout the database, simplifying query scripts and shortening execution times.24 53 In contrast, traditional relational database models are highly normalized (i.e. data transactions are spread across a multitude of tables), which can be suboptimal for big data analysis because of the time-consuming query scripts required to create complex reports. After data have been audited, validated, and stored in a dimensional database, VA tools are used to construct user-friendly dashboards that display histograms and graphs in order to identify and explore data trends. Visualization reduces the load on working memory, offloads cognition, and harnesses the power of human perception.54 Fundamental principles for effective VA

Fig 1 Screen shot of The Children’s Hospital of Philadelphia Perioperative Blood Transfusion visual analytics dashboard. This enables the user to explore historic blood transfusion data (based on patient characteristics and procedure type).
dashboard design include displaying critical information on a single screen, minimizing the number of objects on the screen, keeping graphical icons sparse, displaying context in abbreviated form, and using multiple colour intensities rather than multiple hues.55

Visual analytics tools are applied typically to health-care data belonging to one of three categories: business purposes, clinical operations, or scientific research in various health-care-related fields, such as genomics56 and epidemiology.57 This discussion focuses primarily on the use of VA techniques in clinical operations (i.e. patient care). Visual analytics tools are used extensively at the authors’ tertiary care paediatric hospital for exploration of patient and clinician data and CDS. Examples include a dashboard for exploring blood transfusion data filtered by patient characteristics and procedure type before ordering blood products58 (Fig. 1) and a dashboard monitoring EHR medication alerts (Fig. 2).59 Our hospital also uses VA dashboards to monitor hand hygiene compliance, nursing metrics, supply chain performance, and adherence to clinical guidelines (e.g. Febrile Infant Pathway dashboard).24 Many additional examples of VA applications in health care are available in the literature, such as for visualizing dynamic data from multiple EHRs,60 61 tracking symptom evolution during disease progression,62 performing pharmacokinetic–pharmacodynamic analysis,63 building detection models for disease surveillance,64 visualizing outcomes data,65 and enhancing health-care education.66 There is a paucity of reports detailing the use of VA in anaesthesia literature despite numerous studies describing the benefits of non-VA methods,67 anaesthesia information management system (AIMS) user interface usability,68 and near-real-time CDS.59 Indeed, a PubMed search using the terms, ‘visual analytics anaesthesia’ returned zero results on December 18, 2014.

While the use of VA tools in anaesthesia remains in its nascent stages, the first steps towards big data in anaesthesia occurred in the early 1990s,70 71 and since then proliferation of EHRs over the last two decades has increased availability of clinical anaesthesia data.72 As of May 2013, 67% of US academic hospitals had installed AIMS to generate automated electronic anaesthesia records.73 These institutions—in combination with community hospitals and ambulatory surgery centres—serve as a potential source of anaesthesia-related big data when combined with EHR data and then aggregated across multiple institutions.72 Whether or not the volume of anaesthesia data can be classified as big data remains a topic of debate.74 75 National database efforts such as the
AQA’s NACOR and MPOG will continue to produce anaesthesia big data for retrospective studies, outcomes research,76 quality assessment and management,77 and modelling for prospective studies, although anaesthetists should be aware of the pitfalls associated with database research, including suboptimal data quality, validation, and security concerns.72 76

The future of analytics in anaesthesia and health care

The proliferation of AIMS and EHRs has resulted in big data in anaesthesia and health care to be managed and analysed for various purposes, including practice management, quality improvement, and outcomes research. Emerging trends for analytics and big data in health care include facilitating population health management and value-based accountable care, detecting fraud, and using targeted communication and education campaigns to facilitate patient engagement.79 The future of anaesthesia and health-care analytics will involve ever-increasing demand for and application of sophisticated analytics methods and tools (e.g. VA dashboards) to explore and analyse data with the goals of improving patient care, increasing efficiency, optimizing resource utilization and allocation, and enhancing decision making at both clinical and enterprise levels.26 80 Patient-generated big data have been identified as a promising approach to personalized health care and medical diagnostics via data mining and the development of computer-aided diagnostic tools.81 82 Anaesthetists, through national efforts such as NACOR and MPOG, will continue to leverage big data and analytics tools to conduct research83 and transform data into information that can guide clinicians and improve patient outcomes.84

Conclusions

The improving sophistication and ubiquity of electronic monitoring systems both in the hospital and within the community feed ever-increasing amounts of anaesthesia and health care big data to clinicians, administrators, and researchers. Analytics methods offer promising tools to leverage big data to improve patient care, quality assessment, financial management, and other areas of health care. One method, visual analytics, can facilitate exploration of big data in order to generate hypotheses and guide selection and use of advanced analytics methods. Proliferation of AIMS and EHRs will increase the demand for anaesthetists who can bridge the gap between the medical and information sciences.

Authors’ contributions

A.F.S., L.M.A., and M.A.R. contributed substantially to the conception and design of this review, drafted the article, and revised it critically for important intellectual content, approved the final version to be published, and agree to be accountable for all aspects of the work, thereby ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Declaration of interest

A.F.S. and L.M.A. have no interests to declare. M.A.R. has served on the editorial board for the Journal of Clinical Monitoring and Computing, a journal that publishes anaesthesia-related articles.

References

3 Jacobs A. The pathologies of big data. Commun ACM 2009; 52: 36 – 44
11 Kheterpal S. In the land of the blind, the one-eyed man is king. Anesthesiology 2014; 120: 523 – 5
13 Murdoch TB, Detsky AS. The inevitable application of big data to health care. JAMA 2013; 309: 1351 – 2
14 Davenport TH, Harris J, Shapiro J. Competing on talent analytics. Harv Bus Rev 2010; 88: 52 – 8, 150
17 Schaeffer SE, Van Loan M, German JB. A feasibility study of wearable sensors (Basel) 2013; 13: 17472 – 500
72 Sessler DI. Big Data – and its contributions to peri-operative medicine. Anaesthesia 2014; 69: 100 – 5
74 Fisher DM. “Big data” has not come to pediatric anaesthesia. Anesthesiology 2014; 121: 204
80 Barlow RD. Great expectations for big data: will the next wave of analytics lead to a great awakening or more strife? Health Manag Technol 2014; 35: 18 – 21

Handling editor: H. C. Hemmings