Journal of Business Research 70 (2017) 287–299

Contents lists available at ScienceDirect

Journal of Business Research

Exploring the path to big data analytics success in healthcare

Yichuan Wang *, Nick Hajli *

Newcastle University Business School, 102 Middlesex Street, London, E1 7EZ, United Kingdom

ARTICLE INFO

Available online 15 August 2016

Keywords:
Big data analytics
Business value
Capability building view
Resource-based theory
Information technology source management
Health care industries

ABSTRACT

Although big data analytics have tremendous benefits for healthcare organizations, extant research has paid insufficient attention to the exploration of its business value. In order to bridge this knowledge gap, this study proposes a big data analytics-enabled business value model in which we use the resource-based theory (RBT) and capability building view to explain how big data analytics capabilities can be developed and what potential benefits can be obtained by these capabilities in the health care industries. Using this model, we investigate 109 case descriptions, covering 63 healthcare organizations to explore the causal relationships between the big data analytics capabilities and business value and the path-to-value chains for big data analytics success. Our findings provide new insights to healthcare practitioners on how to constitute big data analytics capabilities for business transformation and offer an empirical basis that can stimulate a more detailed investigation of big data analytics implementation.

1. Introduction

During the past decade, there has been a proliferation of research on health information technology (HIT), from both practitioners and academics, showing that HIT is essential for improving quality of care and financial performance (Agarwal, Gao, DesRoches, & Jha, 2010). The intensive use of HIT has generated the enormous variety of patient data that comes from medical recordings (e.g., electronic healthcare records; EHRs, biomedical data), as well as external data sources, such as insurance claims/billing, R&D laboratories, and social media (Ward, Marsolo, & Froehle, 2014). Such large-scale data galvanizes healthcare organizations toward making huge investments in big data analytics to acquire valuable insights and facilitate timely decision-making, minimize patient risk, and reduce clinical cost (Chen, Chiang, & Storey, 2012).

Computer scientists emphasize that big data analytics is capable of processing an immense volume, variety and velocity (3V) of data across a wide range of healthcare platforms, and has tremendous benefits on medical functions (Jiang et al., 2014; Srinivasan & Arunasalam, 2013). Big data analytics is increasingly advocated as one of the most important information technology (IT) innovations for healthcare organizations (Raghupathi & Raghupathi, 2014). Compared to other industries, such as retailing and banking industries, however, healthcare industry lags behind in taking advantage of thoughtful analytical tools and methods (Ferranti, Langman, Tanaka, McCall, & Ahmad, 2010; Fihn et al., 2014). Healthcare organizations are struggling with the implementation of big data analytics although they invest in numerous analytics technologies with the hope for healthcare transformation (Murdoch & Detsky, 2013; Shah & Pathak, 2014). Evidence from a survey also shows that 60% of healthcare organizations surveyed fail to develop a clear, integrated enterprise strategy and vision for analytics deployment across a broad range of functions (Deloitte Center for health Solutions, 2015). One of the reasons for these failures is a lack of understanding on the economic potential of big data analytics (Groves, Kayyali, Knott, & Kuiken, 2013; Murdoch & Detsky, 2013). Indeed, big data analytics is a double-edged sword for IT investment, potentially incurring huge financial costs for healthcare organizations due to poor governance (Watson, 2014). On the other hand, with appropriate governance, it has the potential to equip organizations with the tools they need to harness the mountains of heterogeneous data, information, and knowledge that they routinely gather (Bardhan, Oh, Zheng, & Kirksey, 2015; Basole et al., 2015; Bates, Saria, Ohno-Machado, Shah, & Escobar, 2014), support a wide range of medical functions at a lower cost (Raghupathi & Raghupathi, 2014), and develop a new portfolio of business strategies for their products and services.

According to a systemic review on the current state of big data research, by Wamba, Akter, Edwards, Chopin, and Gnanzou (2015), the constantly growing body of research on big data has mostly focused on addressing technical issues. However, organizations will not acquire the full benefits of leveraging big data analytics unless they are able to address managerial challenges effectively (Mcafee & Brynjolfsson, 2012), orchestrate strategic choices and resource configurations (Xu, Frankwick, & Ramirez, 2016), as well as understand the managerial, economic, and strategic impact of big data analytics (Raghupathi & Raghupathi, 2014; Ward et al., 2014). Without reasonable justifications, not only it is difficult to help healthcare practitioners focus priorities and

* Corresponding authors.
E-mail addresses: yi-chuan.wang@ncl.ac.uk (Y. Wang), nick.hajli@ncl.ac.uk (N. Hajli).

http://dx.doi.org/10.1016/j.jbusres.2016.08.002
0148-2963/

efforts on driving value from the adoption of big data analytics, but it also cannot find sufficient evidence of how big data analytics investment can pay off (Murdoch & Detsky, 2013; Shah & Pathak, 2014). Moving a deeper understanding on the ways and means to create business value from big data analytics will result in reducing a resistance to adopt big data analytics and an ineffective use of analytics. Thus, exploring the path to big data analytics success for healthcare transformation is currently one of the most discussed topics in the fields of computer science, information systems (IS) and healthcare informatics. This study seeks answers to the following research question: *How healthcare organizations can capture business value from big data analytics?*

To answer this question, we base our exploratory analysis on a theoretical model - namely the big data analytics-enabled business value (BDAE-BV) model to explain how big data analytics capabilities can be developed and what potential benefits can be obtained by these capabilities in healthcare organizations. Specifically, we use the resource-based theory and capability building view to link big data architectural components, through analytics-enabled IT capabilities, to a big data analytics-specific benefits framework. This model was subsequently validated on a broader empirical basis by using 109 case descriptions of big data analytics implementation. Our findings offered theoretical and practical insights on big data analytics in the healthcare context; this can enrich the understanding of big data analytics' business value creation and can also provide guidance and evidence for healthcare practitioners for their business case justifications.

2. Theoretical foundation for deriving big data analytics-enabled business value model

The theoretical foundation of our BDA-BV model comprises of two elements: resource-based theory (RBT) and capability building view. During the last two decades, RBT has been the principal theoretical foundation for explaining how resources can be transformed into a sustained competitive advantage (Barney, 1991, 2001). RBT assumes that a firm can be profitable as long as it can exploit a bundle of valuable, rare, inimitable, and non-substitutable (VRIN) resources in a highly competitive market (Barney, 1991). Drawing on the RBT, much of the works in the IS field have argued the different types of IT resources (e.g., physical, technical and human IT resources) can add value to firms' operations (Bharadwaj, 2000; Doherty & Terry, 2009; Karimi, Somers, & Bhattacherjee, 2007; Lin & Wu, 2014; Melville, Kraemer, & Gurbaxani, 2004). However, several research commentaries criticize RBT, stating that it lacks explanatory power on how IT resources are orchestrated, how specific IT systems can create unique and idiosyncratic IT capabilities and how they ultimately lead to competitive advantage gains (Kim, Shin, Kim, & Lee, 2011; Kohli & Grover, 2008; Mukhopadhyay, Kekre, & Kalathur, 1995).

Capability building view has been utilized to complement the pitfalls of RBT (Bharadwaj, 2000; Doherty & Terry, 2009; Karimi et al., 2007; Santhanam & Hartono, 2003; Saraf, Langdon, & Gosain, 2007; Wang, Liang, Zhong, Xue, & Xiao, 2012). Capability building refers to "the ability of firms to build unique competencies that can leverage their resources" (Karimi et al., 2007, p. 223). Capability building view suggests that firms have to build capabilities by selecting and deploying resources and assembling these resources into synergetic combinations, thereby transforming inputs into valuable outputs (Karimi et al., 2007; Weill & Vitale, 2002). Teece, Pisano, and Shuen (1997) have argued that such capabilities cannot easily be bought; they must be built (p. 529). Applying the capability building view in the IS field, Bharadwaj (2000) extend the notion of capabilities to a firm's IT function and defined a firm's IT capability as its "ability to mobilize and deploy IT-based resources in combination or copresent with other resources and capabilities" (p. 160). Kohli and Grover (2008) further suggest that IT capabilities are often created by combining specific physical IT artefacts, human, and technological resources.

Drawing on the capability building view, a proliferation of research has explored IT functional capabilities by certain basic IT architecture, IT functionalities or system software, arguing that such capabilities can lead to better strategic value and organizational performance (Iyer & Henderson, 2010; Rai, Pavlou, Im, & Du, 2012; Ravichandran, Lertwongsatien, & Lertwongsatien, 2005; Mueller, Viering, Legner, & Riempp, 2010; Pavlou & El Sawy, 2010). By breaking down IT-leveraging capabilities into its three underlying IT system components, for example, Pavlou and El Sawy (2010) examine whether the effective use of project and resource management systems, organizational memory systems, and cooperative work systems can achieve organizational capabilities and competitive advantage in the new product development. Meanwhile, by developing a comprehensive service-oriented architecture economic potential model (SOA-EPM), Mueller et al. (2010) identify a set of SOA capabilities (e.g., reusability, interoperability, and flexibility), derived from its design principles for enhancing organizational performance.

In the healthcare context, Anand and Wamba (2013) propose a comprehensive model to assess the business value of radio frequency identification (RFID) applied in healthcare and elucidate how capabilities of RFID improve process level effects (i.e., automational, informational, and transformational) resulting in the gain of organizational performance. Singh, Mathiassen, Stachura, and Astapova (2011) disentangle the relationship between different types of IT-enabled capabilities and improved clinical and financial outcomes. By conducting a longitudinal study on a home care provider, they found that the abilities formed by remote patient monitoring (RPM) and home healthcare devices can facilitate the formation of transactional and transformational dynamic capabilities and performances. Ghosh and Scott (2011) describe how analytical capabilities facilitate data-driven decision making. Their case study shows that Veterans Health Administration's (VHA) big data analytics systems allow aggregating patient data to establish measurable improvements, which help healthcare managers to allocate resources (e.g., determine the resource utilisation for the facility and geographic distribution of patients' support service needed) and choose future treatments and policies (e.g., assess the outcomes of policy initiatives and develop medical protocols).

Guided theoretically by these aforementioned studies, we view big data analytics architecture as a specific technical IT resource based on the RBT. It is characterized by a set of big data analytics architectural components (i.e., data aggregation, data processing, and data visualization). Each big data analytics architectural component is constituted by the specific big data analytics tools and functionalities that are used to transform healthcare data from various sources into meaningful clinical insights through big data analytics tools. Building on the IT capability building view, each component could be logically expected to generate big data analytics capabilities and these capabilities are expected to induce the business value. We thus link the logical paths among the big data analytics architectural components, big data analytics capabilities, and potential business value driven by these capabilities.

The conceptualization of our model is illustrated in Fig. 1. The solid boxes in Fig. 1 are left blank at this stage, since the logic path between big data analytics capabilities and benefit dimensions is part of our exploratory work. Later, they are filled by first identifying big data analytics capabilities, then they are linked to the benefit sub-dimensions, based on the analysis of big data implementation cases. In the following subsections, we first elaborate on each big data analytics architectural component, followed by the definition of big data analytics capabilities, and the conceptualization of big data analytics' business value.

3. The constructs of big data analytics-enabled business value model

3.1. Big data analytics architecture as a technical IT resource

To identify the components of big data analytics architecture, we review over 10 big data analytics architectures from academic literature

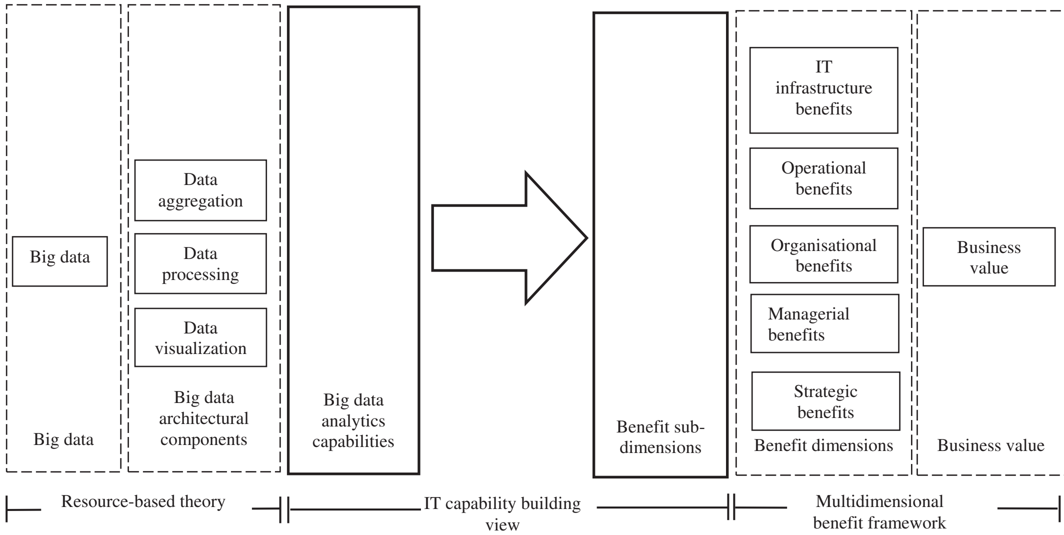

Fig. 1. Big data analytics-enabled business value (BDA-BV) model. Note: The solid box indicates the path to business value, driven by big data analytics capabilities, which is the primary focus of this exploratory study.

and technology tutorial (e.g., Hu, Wen, Chua, & Li, 2014; Jagadish et al., 2014; Phillips-Wren, Iyer, Kulkarni, & Ariyachandra, 2015; Raghupathi & Raghupathi, 2014; Ward et al., 2014; Watson, 2014). Our review starts with Ward et al.'s (2014) big data analytics architectural framework applied in the context of healthcare. This framework elucidates how clinical data in terms of four architectural layers that begin with data generation and continue through data extraction and data analysis to visualization and reporting, listing the tools and functionalities that are used in each architectural component. With these components in mind, we invite four IT experts (two practitioners and two academics) to participate in the five-round discussions for finalizing this architecture. During the discussions, we agree that big data analytics architecture is rooted in the concept of information life cycle management (i.e., collection, repository, process, and dissemination). It consists of an information flow logical framework that starts with data capture, proceeds via data transformation, and culminates with data consumption. This review generally affirmed Ward et al.'s framework, apart from the need to integrate data generation and data extraction under a single dimension – data aggregation – because big data analytics systems typically use data warehousing and ELT tools to capture, aggregate, and ready data from various sources for processing (Raghupathi & Raghupathi, 2014). Therefore, big data architecture in this study is comprised of three major architectural components: (1) data aggregation, (2) data processing, and (3) data visualization, as described below.

Data aggregation component aims to collect heterogeneous clinical data from multiple sources and transform various sources of data into certain data formats (Ward et al., 2014). In this component, data will be intelligently aggregated by three key functionalities from data aggregation tools: acquisition, transformation, and storage (Raghupathi & Raghupathi, 2014). First, data acquisition is used to effectively collect and extract data from external sources and all of the health system's components throughout the healthcare units (Phillips-Wren et al., 2015). Second, during data transformation, transformation engines are capable of moving, cleaning, splitting, translating, merging, sorting, and validating data. These transformation engines make data consistent, visible, and easily accessible for analysis. Healthcare context (data such as that typically contained in a patient record) would be extracted from EHR systems and subsequently converted into a specific standard data format, sorted by the specified criterion (e.g., patient name, location, or medical history); the record is then validated against data quality rules. Finally, the cleaned data is loaded into the target databases, such as Hadoop distributed file systems (HDFS) or in a Hadoop cloud for further processing and analysis. The data storage principles are based on compliance regulations, data policies, and access controls; data storage methods can be implemented and completed in batch processes or in real time.

The second architectural component – data processing – aims to process all kinds of data and perform appropriate analyses for harvesting insights (Ward et al., 2014). This is particularly important for transforming patient data into meaningful information that supports evidence-based decision making and meaningful use practices for healthcare organizations. In a simple taxonomy for analytics, developed by Delen (2014), there are three main kinds of analytics: descriptive, predictive, and prescriptive analytics, each distinguished by the type of data and the purpose of the analysis.

Descriptive analytics provides the ability to describe the data in summary form for exploratory insights and to answer "what has happened in the past?" questions (Phillips-Wren et al., 2015; Watson, 2014). In hospital settings, descriptive analytics is useful, as it allows healthcare practitioners to understand past patient behaviors and how these behaviors might affect outcomes from their EHR database. It also provides high-speed parallel processing, scalability, and optimization features, geared toward big data analytics, and offers a private and secure environment for confidential patient records (Wang, Kung, Ting, & Byrd, 2015). *Predictive analytics* allow users to predict or forecast the future for a specific variable, based on the estimation of probability (Phillips-Wren et al., 2015; Watson, 2014). Hadoop/MapReduce is one of the most commonly used predictive analytics-based software products, which integrates the analytical approaches, such as natural language processing (NLP), text mining, and natural networks in a massively parallel processing (MPP) environment. In general, predictive analytics provides the ability to process large volumes of data in batch form cost-effectively, allowing the analysis of both unstructured and structured data, as well as supporting data processing in near real time or real time (Belle et al., 2015). More importantly, predictive analytics enables users to develop predictive models in a flexible and interactive manner to identity causalities, patterns, and hidden relationships between the target variables for future predictions. *Prescriptive analytics* is a relatively new kind of analytics, which uses a combination of optimization-, simulation-, and heuristics-based predictive modeling technique, such as business rules, algorithms, machine learning, and computational modeling procedures (Delen, 2014). Whereas predictive analytics suggests "what will occur in the future (Watson, 2014, p. 1251)", prescriptive analytics offers the optimal solutions or possible courses of action to help users understand what to do in the future (Phillips-Wren et al., 2015; Watson, 2014). Prescriptive analytics can continually re-predict and automatically improve prediction accuracy by taking in new datasets (a combination of structured, unstructured patient data, and business rule) to develop more thorough decisions regarding the diagnoses and treatments (Riabacke, Danielson, & Ekenberg, 2012).

The third architectural component is data visualization. This component generates outputs, such as various visualization reports, real-time information monitoring, and meaningful business insights, derived from the analytics components to users in the organization. Three key functionalities are included. The first functionality yields general clinical summaries, such as historical reporting, statistical analyses, and time series comparisons. Such reporting can be utilized to provide a comprehensive view to support the implementation of evidence-based medicine (Ghosh & Scott, 2011), in order to guide diagnostic and treatment decisions (Fihn et al., 2014). Second, data visualization – a critical big data analytics feature – tends to extrapolate meaning from external data and perform visualization of the information (e.g., interactive dashboards and charts). In healthcare, these visualization reports support physicians and nurses' daily operations, and help them to make faster, better, evidence-based decisions (Roski, Bo-Linn, & Andrews, 2014). Third, real-time reporting, such as alerts and proactive notifications, real time data navigation, and operational key performance indicators (KPIs), can be sent to interested users or be made available in the form of dashboards in real time.

3.2. The definition of big data analytics capabilities

Several definitions of big data analytics capability have been developed in the literature. In general, big data analytics capability refers to the ability to manage a huge volume of disparate data to allow users to implement data analysis and reaction (Hurwitz, Nugent, Hapler, & Kaufman, 2013). Wixom, Yen, and Relich (2013) indicate that big data analytics capabilities for maximizing enterprise business value should encompass speed to insight, which is the ability to transform raw data into usable information; it should also encompass pervasive use, which is the ability to use big data analytics across the enterprise. With a lens of analytics adoption, LaValle, Lesser, Shockley, Hopkins, and Kruschwitz (2011) categorize big data analytics capabilities into three levels: aspirational, experienced, and transformed. The former two levels of analytics capabilities focus on using big data technologies to achieve cost reduction and operation optimization. The last level of capability is aimed to drive customer profitability and make targeted investments in niche analytics.

In this study, we define big data analytics capability through an information lifecycle management (ILM) view. Storage Networking Industry Association (2009) describes ILM as "the policies, processes, practices, services, and tools used to align the business value of information, with the most appropriate and cost-effective infrastructure from the time that information is created through to its final disposition (p. 2)." Generally, data, regardless of its structure in a system, has been followed this cycle, starting with collection, through repository and process, and ending up with dissemination of data. The concept of ILM helps us to understand all of the phases of the information life cycle in big data architecture (Jagadish et al., 2014; Kung, Kung, Jones-Farmer, & Wang, 2015). Therefore, with a view of ILM, we define big data analytics capability in the context of health care as:

> *-the ability to acquire, store, process and analyze large amounts of health data in various forms, and deliver meaningful information to users, which allows them to discover business values and insights in a timely fashion.*

3.3. Conceptualizing the business value of big data analytics

We use a multidimensional IS benefit framework, developed by Shang and Seddon (2002) to categorize the potential benefits of big data analytics, which ultimately lead to business value. Shang and Seddon's (2002) framework – built on a large bogy of previous research – has determined five benefit dimensions, including IT infrastructure benefits, operational benefits, organizational benefits, managerial benefits, and strategic benefits.

We argue that there are three main reasons for choosing Shang & Seddon's benefit dimensions to serve as the outcome of our model. First, part of our research's exploratory work was to provide a specific set of benefit sub-dimensions. This framework will help us to classify the benefits of big data analytics into proper categories, which enhances our understanding of big data's business value. Second, this benefit framework was designed for managers to assess the benefits of their firms' enterprise information systems; it has been refined by many studies related to ERP systems and specific IS architectures (Esteves, 2009; Gefen & Ragowsky, 2005; Mueller et al., 2010). Third, Shang and Seddon (2002) provide a clear guide for assessing and classifying benefits from enterprise systems. Further, Mueller et al. (2010) suggests the ways in which the IS benefit framework can be validated through the case descriptions, which is helpful for our study. Overall, this framework is suitable as a more generic and systemic model for categorizing the benefits of a big data analytics system.

4. Research method

To test the validity of the model in Fig. 1, we employed content analysis as our research approach. Kohli and Grover (2008) suggest that the better way to increase a broader understanding of how companies' IT investments payoff is to learn from their success stories and observe their best practices. Thus, it is appropriate to analyze secondary data from real-world implementation cases by using content analysis to understand the idea of how big data analytics capabilities and benefits will be obtained. Numerous studies have employed the secondary data (e.g., case materials) to elaborate on exploring business values of a specific information system (e.g., Mueller et al., 2010; Peppard, Weill, & Daniel, 2007). In current studies, where no big data analytics constructs are available, inductive content analysis is particularly appropriate for our study to generate categories and subcategories inductively from the case materials, and explore the relationships among them.

4.1. Data collection and selection

Our cases were drawn from case materials on current and past big data projects from multiple sources, such as journal databases, print publications, press coverage, case collections, and reports from companies, vendors, and consultants. The following case selection criteria were applied: (1) the case presents a real-world implementation of big data analytics in healthcare; (2) it clearly describes the software or techniques they introduce and benefits obtained from big data analytics. Technical reports released by IT vendors or white papers without particular case examples were removed from our data set. Meanwhile, to avoid the selection bias, we widely searched for the big data cases from major IT vendors, such as IBM, SAP, Intel, Microsoft, EMC, CISCO, Oracle and Siemens, as well as academic journal databases (i.e., ABI/INFORM Complete, Web of Science, and IEEE Xplore Digital Library). We identified 112 cases from these sources and reviewed these case descriptions to ensure they met our inclusion criteria. In this phase, one case description was eliminated due to a lack of company information and two of them were technical case studies, which only described the novel analytics technologies being developed. The final data set consists of 109 case descriptions covering 63 organizations specifically related to the healthcare industries. Of these sources, we classified 41 sources (37.61%) as reports released by companies or IT vendors, 33 sources (30.28%) as originating from the journals or conference proceedings, and 35 sources (32.11%) as press coverage and print publications, including healthcare institute reports and case collections.

4.2. Data analysis and procedures

A three-phase process for inductive content analysis (i.e. preparation, organizing, and reporting) that developed by Elo and Kyngäs (2008) was performed to ensure a better understanding of big data analytics capabilities and their benefits in the healthcare context.

The purpose of the preparation phase is to make sense of the coding process, in terms of coding unit of analysis, the level of analysis, and the purpose of evaluation (Elo & Kyngäs, 2008). After meeting five times to discuss coding process and reconcile codes, we selected the "themes" (informative and persuasive nature of case material) as the coding unit of analysis, primarily looking for the expressions of an idea, which can be sentences, paragraphs, or a portion of a page (Minichiello, Aroni, Timewell, & Alexander, 1990). The level of analysis in this study is the organization that engages in big data analytics implementation. The purpose of evaluation is to identify the causal relationships among the big data analytics architectural components, capabilities, and business value in the healthcare context from an organizational perspective, in order to develop a big data-enabled business value model. To organize the coding process, we developed a coding instruction containing of the definitions of the sub-constructs of the BDAE-BV model (see Appendix A), outline, examples of the coding procedures, and a guideline for using and administering the data sheets to reach certain reliability requirements (Krippendorff, 2012; Strauss & Corbin, 1998). After making sense of coding process, analysis is conducted in the second phase.

The goal of the second phase is to organize the statements captured from the case descriptions. We created a Microsoft Excel spreadsheet to assign a code to each statement. The first coder purposively highlighted the statements, which can inform the research questions under investigation, which completely describes the path-to-value chains about 1) *how big data analytics architectural components create the big data analytics capabilities* and 2) *how these big data analytics capabilities lead to obtain potential benefits*, while reading through all 109 case descriptions several times. These statements consist of the descriptions of technical solution, business benefit, the functionalities or tools of a specific analytics approach, and the ways in which they are applied to healthcare services or clinical operations. We treated statements in the text of the case descriptions as evidence of support for the patterns and connections of constructs in our model. To achieve neutral or unbiased results, an audit trail and audit process was used at this stage, which will increase the accuracy of classification and interrater reliability (Hsieh & Shannon, 2005). The second coder who has over 15 years working experience with a multinational technology and consulting corporation headquartered in the United States, and is currently involved in several big data projects, went through the same process independently.

During the coding process (see the coding examples in Appendix B), we followed the coding process suggested by Strauss and Corbin (1998), including open coding, axial coding, and selective coding. In the open coding process, the coders broke down, examined, and categorized the statements into predetermined conceptual model categories. The coders also used different color highlights to distinguish each concept relating to the constructs and its underlying items. In order to record these items, we have rephrased the statements from case studies rather than use direct quotes. This is because the quotes are generally too long and difficult to comprehend. Both coders provided rich background knowledge and industrial experience in classifying these statements into the constructs and items with similar meaning. In the axial coding process, the coders reread the statement to explore the pair-wise connections and path-to-value chains, and to develop more precise explanations of what big data analytics architectural components, capabilities, and benefits are, what cause them, and the benefits that arise because of them. Pair-wise connection was determined by examining the absolute number of times a construct was used to describe a path-to-value chain between big data and business value. Path-to-value chain was determined by examining unique combinations of constructs from the conceptual model's different layers and identifying the chains found in multiple cases. In the selective coding, the coders focused more on finalizing the codes by comparing and contrasting other similarly coded constructs (Urquhart, 2007). Frequency analysis was used to evaluate the importance associated with a construct, pair-wise connection, and path-to-value chain based on the repeated appearance of statements (Weber, 1990).

Finally, this coding process resulted in a total of 247 path-to-value chains. The two coders agreed on 82% of the categorisation. Most discrepancies occurred between the two coders on the categories of analytical capability and predictive capability. Once conflict occurred, the two coders reassessed each case and eventually arrived at a consensus. Ensuring interrater reliability led to much discussion and debate (Schilling, 2006). Ensuring interrater reliability led to the elimination of 9 chains. The final set comprises 238 chains. After consolidating the results of the coding and ensuring data quality, we analyzed our data set in respect of two questions:

- Which strong, pair-wise connections occur between the constructs of BDAE-BV model in the analyzed case materials?
- Which are the most prominent path-to-value chains that link big data and benefits?

5. Research results and discussion

From the analysis of the 109 cases and the resulting 238 coded chains, we present our results in the following subsections according to three distinct perspectives: (1) the total number of occurrences of the constructs, (2) the distribution of pair-wise connections between the constructs of BDAE-BV model, and (3) the distribution of path-to-value chains connecting all the constructs of BDAE-BV model.

5.1. Discussion of constructs

5.1.1. Big data analytics architectural components

Big data analytics capabilities are derived from the various tools and functionalities of big data analytics. Table 1 shows that big data analytics capabilities are mainly triggered by a data processing component (coded as part of 103 chains). This is followed by a data aggregation (74) and data visualization component (61). We break down three big data analytics architectural components, which displays the number of occurrence in the case materials for each component. Numerous cases (e.g., Kaiser Permanente Northern California and Dutch long-term care institution) indicate that descriptive analysis, online analytical processing (OLAP), and data mining is useful tools for analyzing structured patient data from the multiple perspectives.

Furthermore, our results also show that data aggregation is one of the critical big data analytics features, which can be used to improve the standardization of health care data (Shah & Pathak, 2014). An example is the Dutch long-term care institution that has been collecting incident events that include attributes, such as client, department, date and time, type of incident, cause, location, physical damage and mental damage (Spruit, Vroon, & Batenburg, 2014). Using data aggregation tools allow them to capture and extract the data from multiple sources and locations and integrate into their system. This helps them explore the causes of occurred ancient events from the relational databases.

5.1.2. Big data analytics capabilities

Among the five big data analytics capabilities identified, analytical capability was the primary capability for big data analytics (coded as part of 74 chains), followed by speed to insights capability (51), predictive analytics capability (46), interoperability capability (40) and traceability (27). Breaking down five big data analytics capabilities as shown in Table 2, which displays the number of occurrence in the case materials for each capability, we find that the ability to analyze semi-

Table 1
Breaking down big data analytics architectural components.

Big data analytics architectural components	Tools being used in the cases	The number of occurrence
Data aggregation	• Middleware • Data warehouse tools (SQL database, NoSQL database, and cloud-based database) • Extract-transform-load (ELT) tools • Hadoop distributed file system (HDFS)	74
Data processing	• Apache Hadoop/Map Reduce • Statistical analysis • OLAP • Predictive modeling • Social media analytics • Machine learning • Adhoc querying analysis • Text mining/ NLP	103
Data visualization	• Visual dashboards/systems • Reporting systems/interfaces	61
Total		238

structured and unstructured data is mentioned most often. Unstructured and semi-structured data in healthcare refers to information that can neither be stored in a traditional relational database nor fit into predefined data models, such as XML-based electronic healthcare records (EHRs), clinical images, medical transcripts, and lab results. Prior research points out that the ability to analyze unstructured data plays a pivotal role in the success of big data in healthcare settings since 80% of health data is unstructured data (Murdoch & Detsky, 2013; Russom, 2011). The main difference in the analysis process between big data analytics management systems and traditional management systems is that the former has a unique ability to analyze semi-structured or unstructured data that reveals important correlation patterns, which were previously difficult or impossible to determine (Watson, 2014). One of our cases, Leeds Teaching Hospitals in the UK analyze approximately one million unstructured case files per month, and have identified 30 distinct scenarios where there is room for improvement in either costs or operating procedures by taking advantage of natural language processing. This unstructured data analytical capability enables Leeds to improve efficiency and control costs through identifying costly healthcare services such as unnecessary extra diagnostic tests and treatments (Intel, 2013).

Speed to decisions capability allows healthcare organizations to automatically generate warmings or notices and send it to clinicians in due course (17) and generate clinical summary and presents it using visual dashboards/systems (16). Reports generated by big data analytics engines are distinct from transitional IT architectures as they facilitate the assessment of past and current operational environments across all organizational levels. Visualization reports are normally generated after near-real-time data processing and displayed on healthcare performance dashboards which assist healthcare analysts to recognize emerging healthcare issues such as medical errors, potential patient safety issues and appropriate medication use.

Predictive analytics capability manifested in 46 occurrences. However, most of occurrences come from the cases reported by the IT vendors. The cases from academic literature indicate that the application of predictive analytics to health care fields is still in its earliest stages. One of our cases demonstrated the difficulty in developing a reliable predictive model without the ability to exploit large quantity of valuable dataset (Spruit et al., 2014). Similarly, Amarasingham, Patzer, Huesch, Nguyen, and Xie (2014) conclude that due to the difficulty to customize legacy healthcare information systems for predictive models it limits the quality of predictions. They further suggest that predictive models may not

Table 2
Breaking down five big data analytics capabilities in health care.

Business analytics capabilities	Items	The number of occurrence in the cases	
Traceability	Access to historic lab results, current patient medications, and immunization history	14	27
	Track patient data based on the rules that built on hospital claims data	8	
	Monitor patient conditions on a daily basis	5	
Analytical capability	Analyze semi-structured and unstructured data (e.g., imaging data, the meaning and context of human language and voice)	22	74
	Run broad multidisciplinary studies that extract important insights from large amounts of healthcare data.	14	
	Identify correlations and patterns from diverse data to gain new insights	13	
	Analyze information in near-real or real time	10	
	Automate a process for continuous patient safety monitoring	5	
	Compare "what if" scenarios	4	
	Provide comparative interpretation of similar patient cases over time	3	
	Detect fraud, abuse, waste, and errors in care claims	3	
Speed to decisions capability	Automatically notify clinicians of critical issues (e.g., the appropriate dosing of anti-coagulation medication.	17	51
	Generate detailed reporting in visual ways	16	
	Provide actionable insights to decision-makers in near-real or real time	12	
	Generates proactive clinical recommendations for conditions	6	
Predictive analytics capability	Allow to predict patient behaviors	21	46
	Identify the patients who need extra follow up	14	
	Generate a set of predictions about the effectiveness of various treatment options for patients based on unique characteristics.	7	
	Perform "what if" analysis using predictive modeling	4	
Interoperability Capability	Integrate heterogeneous data from multiple hospital systems and devices	17	40
	Integrate data from other hospitals, clinics and data sources	14	
	Allow the information to be shared in the cloud-based data warehouse with other institutions	9	
Total		238	

Table 3
Breaking down the potential benefits driven by business analytics in health care.

Potential benefits of big data	Items	The number of occurrence in the cases	
IT infrastructure benefits	Reduce system redundancy	17	55
	Avoid unnecessary IT costs	14	
	Transfer data quickly among healthcare IT systems	12	
	Process standardization among various healthcare IT systems	6	
	Simplify IT management	4	
	Reduce IT maintenance costs regarding data storage	2	
Operational benefits	Improve the quality and accuracy of clinical decisions	38	128
	Process a large number of health records in seconds	21	
	Enable proactive treatment before the condition worsens	12	
	Reduce the number of unnecessary treatments	11	
	Shorten the time of diagnostic test	8	
	Meaningful use of EHRs	8	
	Reduce the rate of readmission	8	
	Reduce the time of patient travel	7	
	Reductions in surgery-related hospitalizations	7	
	Immediate access to clinical data for analysis	5	
	Explore inconceivable new research avenues	3	
Organizational benefits	Deliver a seamless, coordinated and consentient patient experience across all of its facilities	12	37
	Improve cross-functional communication and collaboration among administrative staff, researchers, clinicians and IT staff	9	
	Detect interoperability problems much more quickly than traditional manual methods	7	
	Drive full adoption of EHRs across organizational boundaries	5	
	Enable to share data with other institutions and add new services, content sources and research partners	4	
Managerial benefits	Gain insights quickly about emerging trends	6	11
	Provide members of the board and heads of department with sound decision-support information on the daily clinical setting	3	
	Optimization of business growth-related decisions	2	
Strategic benefits	Provide a comprehensive view of treatment delivery for meeting future need	4	7
	Use business analytics as a competitive differentiator	3	

respond to changes in EHRs, therefore requires IT personnel to manually refine the predictive rules which lowers the efficiency and productivity.

5.1.3. Potential benefits of big data analytics

Our results (see Table 3) show that the more compelling benefits of big data analytics are operational benefits (128) and IT infrastructure (55). The results also indicate that *improve the quality and accuracy of clinical decisions* (38), *process a large number of health records in seconds* (21), and *enable proactive treatment before the condition worsens* (12) are the elements with high frequency in the category of operational benefits. *Reduce system redundancy* (17), *avoid unnecessary IT costs* (14), and *transfer data quickly among healthcare IT systems* (12) are the elements most mentioned in the category of IT infrastructure benefit. This finding implies that big data analytics has a twofold potential, as it is implemented in an organization. It not only improves IT effectiveness and efficiency, but supports the optimization of clinical operations.

5.2. Discussion of pair-wise connections

In the next level of analysis, we looked at the pair-wise connections between any two of our BDAE-BV model's elements, which contribute to a deeper understanding of the causal connections among them. Our results reveal that there are some pair-wise connections, which we deem to be particularly prominent. With respect to the connections between big data analytics architectural layers and big data analytics capabilities, as shown in Table 4, 43% of big data analytics capabilities were formed by the data processing layer. This layer mainly leads to analytical capability (53 connections) and predictive analytics capability (21 connections). This aspect underlines that a data processing component in big data solutions is the critical role in enhancing healthcare organizations' various analytics capabilities. On the other hand, the data visualization layer leads to speed to decision capability (39 connections) and predictive analytics capability (22 connections). This shows that providing the meaningful information in a visual way enables managers to make decisions with ease.

Our results regarding the connection between big data analytics capability and benefits dimensions reveal that different big data analytics capabilities and various combinations bring different benefits (see Table 5). One particular big data analytics capability – analytical capability – is associated with all five potential benefits, with a total of 74 connections: operational benefits (43 connections), IT infrastructure benefits (14 connections), and managerial benefits (12 connections). Speed to decisions capability has the second highest count of connections (51), which contributes to operational benefits (29 connections) and managerial benefits (12). Predictive capability leads in five dimensions of benefits and it has 46 links in total. On the other hand,

Table 4
Number of pair-wise connections linking big data analytics architectural components with big data analytics capabilities.

Business analytics capabilities	Big data analytics architectural components			
	Data aggregation	Data processing	Data visualization	Total
Traceability	10	17	0	27
Analytical capability	21	53	0	74
Speed to decisions capability	2	10	39	51
Predictive analytics capability	3	21	22	46
Interoperability Capability	38	2	0	40
Total	74	103	61	238

Table 5
Number of pair-wise connections linking business analytics capabilities with benefit dimensions.

Benefit dimensions	Business analytics capabilities					
	Traceability	Analytical capability	Speed to decisions capability	Predictive analytics capability	Interoperability capability	Total
IT infrastructure benefits	10	14	5	3	23	55
Operational benefits	17	43	29	28	11	128
Managerial benefits	0	12	12	7	6	37
Organizational benefits	0	3	4	4	0	11
Strategic benefits	0	2	1	4	0	7
Total	27	74	51	46	40	238

interoperability capability could potentially bring both IT infrastructure benefits (23 connections) and operational benefits (11 connections). Based on this results, we find that healthcare organizations needs to focus on developing the analytical, speed to decisions and predictive analytics capabilities that in order to gain operational and managerial benefits, while for IT infrastructure benefits, interoperability and analytical capabilities are needed. All five big data analytics capabilities could bring IT infrastructure and operational benefits, but only analytical, speed to decisions, and predictive analytics capabilities are associated with managerial, organizational, and strategic benefits. Overall, 77% of chains show that IT infrastructure and operational benefits can be acquired by the use of big data analytics. However, our results also demonstrate that big data analytics have a limited ability to help healthcare organizations gain organizational, strategic, or managerial benefits as of now. Thus, in health care, more advanced applications and maturing analytical processes, along with alignment between the business and analytics strategies, the formation of decision-making culture, strong committed sponsorship, and people skilled in the use of analytics, are necessary for big data solutions to achieve their full potential (Wang, Kung, & Byrd, 2016; Watson, 2014).

5.3. Discussion of path-to-value chains

Three path-to-value chains were observed most frequently as shown in Fig. 2. Three of the path-to-value chains were observed most frequently (coded 71 times in total). We did not present any process link from predictive capability, because the frequency count is below the cut-off point (20 occurrences) chosen.

The first of these chains leads from data processing component, through analytical capability, to operational benefits (26 occurrences). Analytical capability refers to the analytical techniques typically used in a big data management system to process data with an immense volume (from terabytes to exabytes), variety (from text to graph), and velocity (from batch to streaming) via unique data storage, management, analysis, and visualization technologies (Chen et al., 2012). An analytical process in a big data analytics system starts by acquiring data from both inside and outside the healthcare sectors, storing it in distributed database systems, filtering it according to specific discovery criteria, and then analyzing it to generate meaningful outcomes for users to facilitate decision-making.

Having analytical capability allows healthcare organizations to identify patterns of care and discover associations from massive healthcare

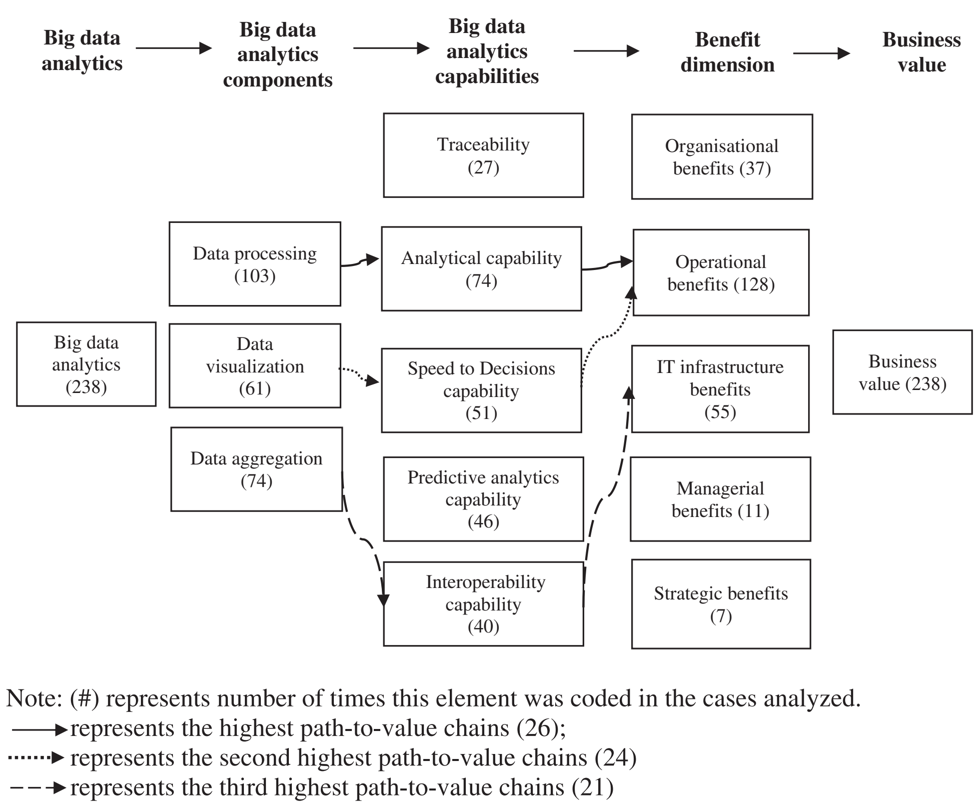

Fig. 2. Big data analytics-enabled business value model Including the analysis of constructs in the cases.

records, thus providing a broader view for evidence-based clinical practice. Analytical approaches provide solutions to parallel process large data volumes, manipulate real-time, or near real time data, and capture all patients' visual data or medical records. In doing so, such approaches can identify previously unnoticed patterns in patients related to hospital readmissions and support a better balance between capacity and cost. Interestingly, our many cases shows that analyzing patient preference patterns also helps hospitals to recognize the utility of participating in future clinical trials and identify new potential markets. For example, through taking advantage of social media resources such as comments and feedback, hospitals can gain access to more information on current patient treatment conditions (e.g., side effects and hospitalization), and thus detect rapidly increasing interest in specific markets that will enable them to develop high quality health care that meets the needs of patients.

The second, which starts with data visualization component and moves through speed to decisions capability to operational benefits, is equally significant (24 occurrences). Speed to decisions refers to the ability to quickly transform raw data into meaningful information for supporting decision making and action taking (Wixom et al., 2013). Healthcare organizations with successful analytics initiatives are more likely to yield actionable information and knowledge, such as historical reporting, executive summaries, drill-down queries, statistical analyses, and time series comparisons. Such information is deployed in real time (e.g., medical devices' dashboard metrics), while other information (e.g., daily reports) will be presented in summary form. Healthcare organizations take advantage of this information to uncover comprehensive solutions to pre-set questions, in order to support the implementation of evidence-based medicine, as well as detecting advanced warnings for disease surveillance, and developing personalized patient care (Chawla & Davis, 2013; Jee & Kim, 2013).

Moreover, the new insights can be discovered by an assessment of past and current operation environment all organizational levels and a systemic and comprehensive perspective. This enables managers to recognize feasible opportunities for customer relationship improvement, particularly regarding long-term strategic decisions (Persson & Ryals, 2014). For instance, Premier Healthcare Alliance collects data from different departmental systems and sends it to a central data warehouse. After near-real-time data processing, the reports generated are then used to help users recognize emerging healthcare issues such as patient safety and appropriate medication use. By expeditiously correlating clinical research, patient data, and clinical practice guidelines, WellPoint, one of the largest health benefits companies in the United States, not only could pre-approve a doctor's urgent request for coverage of a treatment in near real-time, but also could provide recommendations on a confidence scale to support nurses' decisions on treatment for each patient in seconds. Such a speed to insight capability allows WellPoint to expedite their decision-making process for each patient, and thus, improve service quality and reduce considerable costs (WellPoint, 2014).

The final chain, which goes from a data aggregation layer, through interoperability capability, to IT infrastructure benefits, is slightly less common (21 occurrences). In general, interoperability capability in a healthcare information system is defined as the ability to integrate data and process, in order to support collaboration and other healthcare activities (Sadeghi, Benyoucef, & Kuziemsky, 2012). Traditionally, legacy healthcare systems have been designed in an iterative manner which focused on business process automation, not care delivery (Eastwood, 2013). As a result, medical errors frequently originate from the fragmentation of clinical data and the lack of standardization of health care data among labs, hospital systems, financial IT systems and EHRs (Institute for Health Technology Transformation, 2013; Shah & Pathak, 2014). Data normalization is an essential process to solve this problem. This process can be accomplished in Hadoop/MapReduce, which is aimed to associate similar types of data from disparate systems into a single data in an attempt to minimize the rate of redundant procedures. Once clinical data has been normalized among healthcare systems, this allows big data analytics platforms to facilitate the transformation of asynchronous patient information and dissemination of patient data in an interoperable way that authorized users can gain easier access to a broad array of data.

Big data analytics system, which provides interoperability capability in a healthcare environment could be an important solution, not only in the early detection and prevention of adverse events from multiple medical functions or locations, such as medication errors, lack of patient monitoring, and patient deterioration, but also in enabling the identification of differences in patient characteristics among different referral workflows (Dolin, Rogers, & Jaffe, 2015). A typical example in our case is Rhode Island Quality Institute (RIQI), which they build an interoperable infrastructure to permit the seamless exchange of data among various systems and aggregate individual patient data into a comprehensive and meaningful view of health care. RIQI use the aggregated quality data and analytics to improve clinical workflow by adjusting resources to cope with inconsistent IT function workflows and identify disease patterns across the state (Institute for Health Technology Transformation, 2013). Therefore, we argue that interoperability capability, driven by big data analytics, can enhance IT infrastructure benefits for healthcare organizations.

6. Implications

6.1. Theoretical implications

Big data-related technologies are the most influential IT innovations in the last decade. Resulted from such phenomenon, research has been focused on the technical side, instead of the managerial and/or strategic views, which further hinders the progress of big data analytics research. A unique theoretical contribution of this paper is its conceptualization of path-to-value chains to drawing business value from big data analytics in the context of healthcare. We use RBT and capability building view as our base to develop a generic BDAE-BV model for research/theory and then validated it empirically in the context of healthcare. Given the lack of models to explore the business value of big data analytics in health care, our research model underpins that big data analytics is a technical IT source with a potential to create analytics-based IT capabilities that, in turn, support the acquisition of healthcare performance. Our exploratory study reveals the essential constructs, connections, and path-to-value chains for understanding big data analytics implementation. In doing so, this model is among the first attempts to systematically capture the complex relations that link big data analytics resources and its capabilities and the associated business value. Compared to recent studies that proposed a big data analytics process model (Seddon, Constantinidis, & Dod, 2012; Tamm, Seddon, & Shanks, 2013), our theoretical model elucidates how big data analytics capabilities were formed by three big data architectural components, and how these capabilities lead to the benefits, which improves the general understanding of big data analytics' business value. To the best of our knowledge, this is a first study that took such a unique approach toward integrating the most prominent theories, by applying the new perspectives to a current IT innovation to show the causal chains of IT business value theoretically and empirically.

We frame our research in healthcare industries, because different industries have different needs or goals of using big data technology solutions. It is best to test a generic model in a specific context. Most prior works focused on a technological understanding of big data rather than identifying the business value of big data analytics in healthcare settings. There has not been sufficient evidence of big data analytics investment benefits (Murdoch & Detsky, 2013; Shah & Pathak, 2014). These findings could be a starting point to study how big data analytics constitutes business value in practice and offer an empirical basis that can stimulate a more detailed investigation of big data analytics implementation.

6.2. Practical implications

Healthcare is paying attention to the implementation of big data analytics and how such innovations can help optimize the quality of care and improve their financial performance due to the rising needs for cost effectiveness and high service quality demands (Raghupathi & Raghupathi, 2014). The BDAE-BV model has identified five big data analytics capabilities and provided three dominant path-to-value chains for healthcare organizations to facilitate their analytics capabilities and show how these capabilities can be implemented to gain business value. Healthcare managers may use them as a template to build their capability portfolio of big data analytics according to their immediate and future plans. Moreover, the insights we gained provide healthcare practitioners with the easy-to-follow scenarios for assessing the benefit of big data analytics and support them in defining their approach to its implementation as they claim that the application of big data analytics to health care is inevitable (Murdoch & Detsky, 2013).

The fourth path-to-value chain we identified in the analytics, which goes from data processing component through predictive analytics capability to operational benefits, was below our cut-off criteria. Yet, it still provides practical insights. Although the application of predictive analytics to health care fields is still in its infancy and the difficulties were demonstrated by many studies (Bardhan et al., 2015; Shmueli & Koppius, 2011; Spruit et al., 2014), failure to adequately utilize predictive analytics tools in analyzing large-scale clinical data results in missed opportunities to offer more insights into clinical decision making process. Predictive analytics helps healthcare managers disentangle the complex structure of clinical cost, identify best clinical practices, and gain a broader understanding of future healthcare trends based on knowledge of patients' lifestyles, habits, disease management and surveillance (Groves et al., 2013). An example form our case selection, Beth Israel Deaconess Medical Center indicates that predictive analytics supports home health care by predicting patient illness, quickly deploying nurse to supplement the care no matter where the patient suffers a health emergency, avoiding expensive emergency department visits, and collaborating with local healthcare providers for care coordination (Halamka, 2014). Therefore, healthcare managers should encourage their employees to establish predictive models that fit in their business situations to discover new ideas in relation to healthcare service and market opportunities.

7. Limitations and future research directions

While this is one of the first studies on exploring the paths to big data analytics success, limitations are inevitable. First, one of the major challenges to validate conceptual model from cases is the case selection. Although several studies have relied on case materials to explore the value of emerging technologies (e.g., Mueller et al., 2010; Seddon et al., 2012), the case materials we chosen for developing the model mainly came from IT vendors and healthcare institute reports and case collections. There is, therefore, a potential bias, as vendors usually showcase their success projects to promote their products. To address this limitation, we selected cases from both the academic databases and IT vendors' case descriptions. Although the relative absence of academic sources in our case collection is due to the incipient nature of big data adoption in health care, they may provide more rigorous and objective statements. Given the growing number of research on healthcare analytics, more cases from academic databases can be included to validate our BDAE-BV model. Future research may consider the further analysis of the data set with respect to additional variables such as hospital size and institution type (public or private).

Second, different industries have different needs or goals of using big data analytics technology solutions. We targeted healthcare for this study, hence, the results are industry-specific. Future research can apply the logic of the BDAE-BV model to other particular industries. Different big data analytics capabilities and outcomes may surface.

The final limitation concerns the coder bias. To address this bias, we first defined the constructs and sub-constructs in our model based on the literature to ensure coders' common understanding of each concept. Then, we followed Strauss and Corbin's (1998) coding process to ensure interrater reliability. Although we have carefully analyzed and interpreted the data to explore the patterns from case descriptions, the coding process still can be improved by more rigorous coding approaches in the future research. In addition, given the growing number of healthcare organizations adopting big data technologies, the sample frame for collecting primary data is larger. Examining the relationships in the BDAE-BV model with a quantitative analysis method could offer stronger empirical evidence.

8. Conclusions

Drawing on the RBT and capability building view, this study illustrates how healthcare organizations' big data analytics capabilities can be developed by big data analytics architecture, thereby leading to create business value. Our findings explore the three path-to-value chains to reach big data analytics success by analyzing secondary data consisting of big data cases specifically in the healthcare context. From analyzing our data set, we used the number of occurrences of constructs, pair-wise connections, and path-to-value chains as indicators for discovering the remarkable paths driving big data analytics to gain business value rather than using the single case. Therefore, this study provides new insights to healthcare practitioners on how to constitute big data analytics capabilities for business transformation and offer an empirical basis that can stimulate a more detailed investigation of big data analytics implementation.

Acknowledgements

The authors would like to thank the editors and the two anonymous reviewers for a most constructive and developmental review process. The authors also thank Professor Terry Anthony Byrd, and Dr. LeeAnn Kung for their insightful comments and suggestions to improve this article. In addition, we would like to thank Dr. Ting from IBM for providing his knowledge and practical experience in assisting the coding process.

Appendix A. Defining the initial sub-constructs of BDAE-BV model.

Sub-constructs	Definitions	Sources
Data aggregation	The tools use to transform different types of healthcare data into a data format that can be read by the data analysis platform.	Raghupathi and Raghupathi (2014)
Data processing	The tools use to process all kinds of data and perform appropriate analyses to harvest insights and decisions	Ward et al. (2014)
Data visualization	The tools used to produce reports about daily healthcare services to aid managers' decisions and actions	Ghosh and Scott (2011)

Appendix A (*continued*)

Sub-constructs	Definitions	Sources
Traceability	Integrate and track the patient data from all of the IT components throughout the various healthcare service units	Wang et al. (2016)
Analytical capability	The ability to process clinical data with an immense volume (from terabytes to exabytes), variety (from text to graph) and velocity (from batch to streaming) by using descriptive analytics techniques	Watson (2014); Seddon et al. (2012); Cao, Duan, and Li (2015)
Speed to decision capability	The ability to effectively generate outputs regarding patients, care process and service to guide diagnostic and treatment decisions	Groves et al. (2013); Wixom et al. (2013)
Predictive capability	The ability to explore data and identify useful correlations, patterns and trends and extrapolate them to forecast what is likely to occur in the future	Negash (2004); Hurwitz et al. (2013)
Interoperability capability	The ability to integrate data and process to support collaboration and other healthcare activities.	Sadeghi et al. (2012)
IT infrastructure benefits	Sharable and reusable IT resources that provide a foundation for present and future business applications	Shang and Seddon (2002)
Operational benefits	The benefits obtained from the improvement of operational activities	
Managerial benefits	The benefits obtained from business management activities which involve allocation and control of the firms' resources, monitoring of operations and supporting of business strategic decisions	
Strategic benefits	The benefits obtained from strategic activities which involve long-range planning regarding high-level decisions	
Organizational benefits	The benefits arise when the use of an enterprise system benefits an organization in terms of focus, cohesion, learning, and execution of its chosen strategies.	

Appendix B. Coding example.

Statements	Open (underlined) and axial (*italic*) codes	Themes emerging from selective coding
(Case name: London's Royal Free Hampstead National Health Service (NHS) Trust) The RAPIDComm 4.0 solution integrates oversight of blood gas, urinalysis, and diabetes care testing and helps drive compliance in all of these areas," says Hall. "We can generate audit trails, process QC data, record patient results while maintaining patient privacy, and limit instrument access to only trained and certified operators. All of this happens on a common platform that supports different types of point-of-care test equipment. The benefits of integrated connectivity also make it easier to build a case for funding to connect more analyzers. Ten of our 35 CLINITEK® analyzers are capable of connectivity, pending the installation of network points. (a case study provided by Siemens Healthcare Diagnostics)	<u>Data aggregation Component</u> <u>Capability-interoperability</u> *Integrate heterogeneous data from multiple hospital systems and devices* (*cause*) <u>Operational benefit</u> *Immediate access to clinical data for analysis (effect)*	Comparing this statement to other similarly coded elements (i.e., data aggregation component, interoperability capability and operational benefit). Two coders agreed on interoperability capability driven by data aggregation component can lead to improved operational benefit Recording this statement as one of the path-to-value chain: **Data aggregation→interoperability→operational benefit**
(Case name: Memorial Healthcare System) The vendor vetting system, known as VETTED, was built based on Memorial Healthcare System's existing enterprise content management platform from IBM and business partner Information Management Consultants (IMC). By adding IBM i2 analytics capabilities, Memorial Healthcare now has greater visibility into its vendor community and accounting staff can complete vetting activities within a few hours. The system's analytics connect the dots among vendor companies and between individuals, vendor companies and physicians to help uncover potential fraudulent behavior. (a case study provided by IBM)	<u>Data processing component</u> <u>Capability-analytical</u> *Identify correlations and patterns from diverse data to gain new insights* (*cause*) <u>Managerial benefit</u> *Gain insights quickly about emerging trends* (*effect*)	Comparing this statement to other similarly coded elements (i.e., data processing component, analytical capability and managerial benefit). Two coders agreed on analytical capability driven by data processing component can lead to improved managerial benefit. Recording this statement as one of the path-to-value chain: **Data processing→analytical capability→managerial benefit**

References

Agarwal, R., Gao, G., DesRoches, C., & Jha, A. K. (2010). Research commentary - the digital transformation of healthcare: Current status and the road ahead. *Information Systems Research*, *21*(4), 796–809.

Amarasingham, R., Patzer, R. E., Huesch, M., Nguyen, N. Q., & Xie, B. (2014). Implementing electronic health care predictive analytics: Considerations and challenges. *Health Affairs*, *33*(7), 1148–1154.

Anand, A., & Wamba, S. F. (2013). Business value of RFID-enabled healthcare transformation projects. *Business Process Management Journal*, *19*(1), 111–145.

Bardhan, I., Oh, J. H., Zheng, Z., & Kirksey, K. (2015). Predictive analytics for readmission of patients with congestive heart failure. *Information Systems Research*, *26*(1), 19–39.

Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, *17*(1), 99–120.

Barney, J. B. (2001). Resource-based theories of competitive advantage: A ten-year retrospective on the resource-based view. *Journal of Management*, *27*(6), 643–650.

Basole, R. C., Braunstein, M. L., Kumar, V., Park, H., Kahng, M., Chau, D. H. P., ... Thompson, M. (2015). Understanding variations in pediatric asthma care processes in the emergency department using visual analytics. *Journal of the American Medical Informatics Association*, *22*(2), 318–323.

Bates, D. W., Saria, S., Ohno-Machado, L., Shah, A., & Escobar, G. (2014). Big data in health care: Using analytics to identify and manage high-risk and high-cost patients. *Health Affairs*, *33*(7), 1123–1131.

Belle, A., Thiagarajan, R., Soroushmehr, S. M., Navidi, F., Beard, D. A., & Najarian, K. (2015). Big data analytics in healthcare. *BioMed Research International*, *2015*, 1–16.

Bharadwaj, A. S. (2000). A resource-based perspective on information technology capability and firm performance: An empirical investigation. *MIS Quarterly*, *24*(1), 169–196.

Cao, G., Duan, Y., & Li, G. (2015). Linking business analytics to decision making effectiveness: A path model analysis. *IEEE Transactions on Engineering Management*, *62*(3), 384–395.

Chawla, N. V., & Davis, D. A. (2013). Bringing big data to personalized healthcare: A patient-centered framework. *Journal of General Internal Medicine*, *28*(3), S660–S665.

Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly*, *36*(4), 1165–1188.

Delen, D. (2014). *Real-world data mining: Applied business analytics and decision making*. FT Press.

Deloitte Center for health Solutions (2015). *Health system analytics: The missing key to unlock value-based care*. Deloitte Development LLC.

Doherty, N. F., & Terry, M. (2009). The role of IS capability in delivering sustainable improvements to competitive positioning. *Journal of Strategic Information Systems*, *18*(2), 100–116.

Dolin, R. H., Rogers, B., & Jaffe, C. (2015). Health level seven interoperability strategy: Big data, incrementally structured. *Methods of Information in Medicine*, *54*(1), 75–82.

Eastwood, B. (2013). Is healthcare IT interoperability (almost) here? (Retrieved January 12, 2015 from CIO Web site) http://www.cio.com/article/2386750/vertical-industries/is-healthcare-it-interoperability-almost-here.html?null

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing*, *62*(1), 107–115.

Esteves, J. (2009). A benefits realisation road-map framework for ERP usage in small and medium-sized enterprises. *Journal of Enterprise Information Management*, *22*(1/2), 25–35.

Ferranti, J. M., Langman, M. K., Tanaka, D., McCall, J., & Ahmad, A. (2010). Bridging the gap: Leveraging business intelligence tools in support of patient safety and financial effectiveness. *Journal of the American Medical Informatics Association*, *17*(2), 136–143.

Fihn, S. D., Francis, J., Clancy, C., Nielson, C., Nelson, K., Rumsfeld, J., ... Graham, G. L. (2014). Insights from advanced analytics at the veterans health administration. *Health Affairs*, *33*(7), 1203–1211.

Gefen, D., & Ragowsky, A. (2005). A multi-level approach to measuring the benefits of an ERP system in manufacturing firms. *Information Systems Management*, *22*(1), 18–25.

Ghosh, B., & Scott, J. E. (2011). Antecedents and catalysts for developing a healthcare analytic capability. *Communications of the Association for Information Systems*, *29*(1), 395–410.

Groves, P., Kayyali, B., Knott, D., & Kuiken, S. V. (2013). *The "big data" revolution in healthcare: Accelerating value and innovation*. McKinsey & Company.

Halamka, J. D. (2014). Early experiences with big data at an academic medical center. *Health Affairs*, *33*(7), 1132–1138.

Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, *15*(9), 1277–1288.

Hu, H., Wen, Y., Chua, T. S., & Li, X. (2014). Toward scalable systems for big data analytics: A technology tutorial. *IEEE Access*, *2*, 652–687.

Hurwitz, J., Nugent, A., Hapler, F., & Kaufman, M. (2013). *Big data for dummies*. Hoboken, New Jersey: John Wiley & Sons.

Institute for Health Technology Transformation (2013). *Transforming health care through big data: Strategies for leveraging big data in the health care industry*. New York, NY: Institute for Health Technology Transformation.

Intel (2013). Leeds teaching hospitals uses big data to study accident and emergency trends. (Available at) http://www.intel.com/content/dam/www/public/us/en/documents/case-studies/big-data-xeon-processor-e5-core-i5-leeds-technology-study.pdf

Iyer, B., & Henderson, J. C. (2010). Preparing for the future: Understanding the seven capacities of cloud computing. *MIS Quarterly Executive*, *9*(2), 117–131.

Jagadish, H. V., Gehrke, J., Labrinidis, A., Papakonstantinou, Y., Patel, J. M., Ramakrishnan, R., & Shahabi, C. (2014). Big data and its technical challenges. *Communications of the ACM*, *57*(7), 86–94.

Jee, K., & Kim, G. H. (2013). Potentiality of big data in the medical sector: Focus on how to reshape the healthcare system. *Healthcare Informatics Research*, *19*(2), 79–85.

Jiang, P., Winkley, J., Zhao, C., Munnoch, R., Min, G., & Yang, L. T. (2014). An intelligent information forwarder for healthcare big data systems with distributed wearable sensors. *IEEE Systems Journal*, *PP*(99), 1–9.

Karimi, J., Somers, T. M., & Bhattacherjee, A. (2007). The role of information systems resources in ERP capability building and business process outcomes. *Journal of Management Information Systems*, *24*(2), 221–260.

Kim, G., Shin, B., Kim, K. K., & Lee, H. G. (2011). IT capabilities, process-oriented dynamic capabilities, and firm financial performance. *Journal of the Association for Information Systems*, *12*(7), 487–587.

Kohli, R., & Grover, V. (2008). Business value of IT: An essay on expanding research directions to keep up with times. *Journal of the Association for Information Systems*, *9*(1), 23–39.

Krippendorff, K. (2012). *Content analysis: An introduction to its methodology*. Sage Press.

Kung, L., Kung, H. J., Jones-Farmer, A., & Wang, Y. (2015). Managing big data for firm performance: A configurational approach. *Proceeding of the Twenty-First Americas Conference on Information Systems, Puerto Rico*.

LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2011). Big data, analytics and the path from insights to value. *MIT Sloan Management Review*, *52*(2), 21–31.

Lin, Y., & Wu, L. Y. (2014). Exploring the role of dynamic capabilities in firm performance under the resource-based view framework. *Journal of Business Research*, *67*(3), 407–413.

Mcafee, A., & Brynjolfsson, E. (2012). Big Data: The Management Revolution. *Harvard Business Review*, 59–68.

Melville, N., Kraemer, K., & Gurbaxani, V. (2004). Review: Information technology and organizational performance: An integrative model of IT business value. *MIS Quarterly*, *28*(2), 283–322.

Minichiello, V., Aroni, R., Timewell, E., & Alexander, L. (1990). *In-depth interviewing: Researching people*. Hong Kong: Longman Cheshire.

Mueller, B., Viering, G., Legner, C., & Riempp, G. (2010). Understanding the economic potential of service-oriented architecture. *Journal of Management Information Systems*, *26*(4), 145–180.

Mukhopadhyay, T., Kekre, S., & Kalathur, S. (1995). Business value of information technology: A study of electronic data interchange. *MIS Quarterly*, *19*(2), 137–156.

Murdoch, T. B., & Detsky, A. S. (2013). The inevitable application of big data to health care. *The Journal of the American Medical Association*, *309*(13), 1351–1352.

Negash, S. (2004). Business intelligence. *Communications of the Association for Information Systems*, *13*, 177–195.

Pavlou, P. A., & El Sawy, O. A. (2010). The "Third Hand": IT-enabled competitive advantage in turbulence through improvisational capabilities. *Information Systems Research*, *21*(3), 443–471.

Peppard, J., Weill, P., & Daniel, E. (2007). Managing the realization of business benefits from in investments. *MIS Quarterly Executive*, *6*(1), 1–11.

Persson, A., & Ryals, L. (2014). Making customer relationship decisions: Analytics v rules of thumb. *Journal of Business Research*, *67*(8), 1725–1732.

Phillips-Wren, G., Iyer, L. S., Kulkarni, U., & Ariyachandra, T. (2015). Business analytics in the context of big data: A roadmap for research. *Communications of the Association for Information Systems*, *37*, 448–472.

Raghupathi, W., & Raghupathi, V. (2014). Big data analytics in healthcare: Promise and potential. *Health Information Science and Systems*, *2*(1), 3.

Rai, A., Pavlou, P. A., Im, G., & Du, S. (2012). Interfirm IT capability profiles and communications for cocreating relational value: Evidence from the logistics industry. *MIS Quarterly*, *36*(1), 233–262.

Ravichandran, T., Lertwongsatien, C., & Lertwongsatien, C. (2005). Effect of information systems resources and capabilities on firm performance: A resource-based perspective. *Journal of Management Information Systems*, *21*(4), 237–276.

Riabacke, M., Danielson, M., & Ekenberg, L. (2012). State-of-the-art prescriptive criteria weight elicitation. *Advances in Decision Sciences*, *2012*, 1–24.

Roski, J., Bo-Linn, G. W., & Andrews, T. A. (2014). Creating value in health care through big data: Opportunities and policy implications. *Health Affairs*, *33*(7), 1115–1122.

Russom, P. (2011). Big data analytics. *TDWI best practices report, fourth quarter*.

Sadeghi, P., Benyoucef, M., & Kuziemsky, C. E. (2012). A mashup based framework for multi level healthcare interoperability. *Information Systems Frontiers*, *14*(1), 57–72.

Santhanam, R., & Hartono, E. (2003). Issues in linking information technology capability to firm performance. *MIS Quarterly*, *27*(1), 125–153.

Saraf, N., Langdon, C. S., & Gosain, S. (2007). IS application capabilities and relational value in interfirm partnerships. *Information Systems Research*, *18*(3), 320–339.

Schilling, J. (2006). On the pragmatics of qualitative assessment: Designing the process for content analysis. *European Journal of Psychological Assessment*, *22*(1), 28–37.

Seddon, P. B., Constantinidis, D., & Dod, H. (2012). *How does business analytics contribute to business value?* Proceeding of Thirty Third International Conference on Information Systems, Orlando, USA.

Shah, N. D., & Pathak, J. (2014). Why health care may finally Be ready for big data. (Retrieved January 12, 2015 from Harvard Business Review) https://hbr.org/2014/12/why-health-care-may-finally-be-ready-for-big-data

Shang, S., & Seddon, P. B. (2002). Assessing and managing the benefits of Enterprise systems: The business Manager's perspective. *Information Systems Journal*, *12*(4), 271–299.

Shmueli, G., & Koppius, O. R. (2011). Predictive analytics in information systems research. *MIS Quarterly*, *35*(3), 553–572.

Singh, R., Mathiassen, L., Stachura, M. E., & Astapova, E. V. (2011). Dynamic capabilities in home health: IT-enabled transformation of post-acute care. *Journal of the Association for Information Systems*, *12*(2), 163–188.

Spruit, M., Vroon, R., & Batenburg, R. (2014). Towards healthcare business intelligence in long-term care: An explorative case study in the Netherlands. *Computers in Human Behavior*, *30*, 698–707.

Srinivasan, U., & Arunasalam, B. (2013). Leveraging big data analytics to reduce healthcare costs. *IT Professional*, *15*(6), 21–28.

Storage Networking Industry Association (2009). *The information lifecycle management maturity model*. Storage Networking Industry Association Press.

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Tamm, T., Seddon, P., & Shanks, G. (2013). Pathways to value from business analytics. *Proceeding of Thirty Fourth International Conference on Information Systems, Milan, Italy*.

Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic Management Journal*, *18*(7), 509–533.

Urquhart, C. (2007). The evolving nature of grounded theory method: The case of the information systems discipline. In A., & K. (Eds.), *The SAGE handbook of grounded theory* (pp. 339–359). Los Angeles, CA: Sage Publications.

Wamba, S. F., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. (2015). How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study. *International Journal of Production Economics*, *165*, 234–246.

Wang, Y., Kung, L., & Byrd, T. A. (2016). Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. *Technological Forecasting and Social Change*. http://dx.doi.org/10.1016/j.techfore.2015.12.019.

Wang, Y., Kung, L., Ting, C., & Byrd, T. A. (2015). Beyond a technical perspective: Understanding big data capabilities in health care. *Proceedings of 48th Annual Hawaii International Conference on System Sciences (HICSS), Kauai, Hawaii*.

Wang, N., Liang, H., Zhong, W., Xue, Y., & Xiao, J. (2012). Resource structuring or capability building? An empirical study of the business value of information technology. *Journal of Management Information Systems*, *29*(2), 325–367.

Ward, M. J., Marsolo, K. A., & Froehle, C. M. (2014). Applications of business analytics in healthcare. *Business Horizons*, *57*(5), 571–582.

Watson, H. J. (2014). Tutorial: Big data analytics: Concepts, technologies, and applications. *Communications of the Association for Information Systems*, *34*(1), 1247–1268.

Weber, R. P. (1990). *Basic content analysis*. Newburry Park, CA: Sage Publications.

Weill, P., & Vitale, M. (2002). What IT infrastructure capabilities are needed to implement e-business models? *MIS Quarterly Executive*, *1*(1), 17–34.

WellPoint (2014). *Opening keynote address at the business health agenda 2014*. WellPoint Press (Available at: http://www.antheminc.com/prodcontrib/groups/wellpoint/documents/wlp_assets/pw_e213326.pdf).

Wixom, B. H., Yen, B., & Relich, M. (2013). Maximizing value from business analytics. *MIS Quarterly Executive*, *12*(2), 111–123.

Xu, Z., Frankwick, G. L., & Ramirez, E. (2016). Effects of Big Data analytics and traditional marketing analytics on new product success: A knowledge fusion perspective. *Journal of Business Research*, *69*(5), 1562–1566.