# Leveraging Big Data Analytics to Improve Quality of Care in Healthcare Organisations: A Configurational Perspective

**Article** *in* British Journal of Management · May 2019

DOI: 10.1111/1467-8551.12332



**4 authors:**

Yichuan Wang
The University of Sheffield

Leeann Kung
Georgia Southern University

Suraksha Gupta
Kent Business School

Sena Ozdemir
Lancaster University

BJM BRITISH JOURNAL of MANAGEMENT
BAM British Academy of Management
*British Journal of Management, Vol. 30, 362–388 (2019)*
DOI: 10.1111/1467-8551.12332



# Leveraging Big Data Analytics to Improve Quality of Care in Healthcare Organizations: A Configurational Perspective

Yichuan Wang, LeeAnn Kung,[1] Suraksha Gupta[2] and Sena Ozdemir[3]
Sheffield University Management School, University of Sheffield, Conduit Rd., Sheffield S10 1FL, UK, [1]Department of Information Systems, College of Business, Georgia Southern University, [2]Newcastle University London, 102 Middlesex St., London E1 7EZ, UK, and [3]Coventry University, Business School, Jaguar Building, Priory Street, Coventry CV1 5FB, UK
Corresponding author email: yichuan.wang@sheffield.ac.uk

**Big data analytics (BDA) is beneficial for organizations, yet implementing BDA to leverage profitability is fundamental challenge confronting practitioners. Although prior research has explored the impact that BDA has on business growth, there is a lack of research that explains the full complexity of BDA implementations. Examination of how and under what conditions BDA achieves organizational performance from a holistic perspective is absent from the existing literature. Extending the theoretical perspective from the traditional views (e.g. resource-based theory) to configuration theory, the authors have developed a conceptual model of BDA success that aims to investigate how BDA capabilities interact with complementary organizational resources and organizational capabilities in multiple configuration solutions leading to higher quality of care in healthcare organizations. To test this model, the authors use fuzzy-set qualitative comparative analysis to analyse multi-source data acquired from a survey and databases maintained by the Centres for Medicare & Medicaid Services. The findings suggest that BDA, when given alone, is not sufficient in achieving the outcome, but is a synergy effect in which BDA capabilities and analytical personnel's skills together with organizational resources and capabilities as supportive role can improve average excess readmission rates and patient satisfaction in healthcare organizations.**

## Introduction

Constantly increasing large volumes of data in various formats (from electronic health records (EHRs)) and other data sources such as pharmaceutical events, insurance claims/billing and R&D laboratories) is challenging healthcare organizations' data management capabilities. The need for better data management is not unique to healthcare, but it is more vital in healthcare because it concerns patients' well-being, which is more important than the bottom line in other industries. Indeed, excellent data management could facilitate reliable predictions of patient behaviour, medical knowledge creation and clinical practice improvements (Kallinikos and Tempini, 2014; Raghupathi and Raghupathi, 2014). However, many healthcare organizations are suffering from a lack of data standards and integration, data overload issues and barriers to the collection of high-quality data that result in billing errors, medical mistakes and generating unnecessary costs (Ward, Marsolo and Froehle, 2014). Data quality depends not only on its own features, but also on the business environment using the data, including business processes and business users. Only the data that conform to the relevant uses and meet requirements can be considered qualified (or good quality) data. Big data analytics (BDA) is increasingly being endorsed for its potentially crucial role in addressing these challenges in healthcare sectors. Yet, our knowledge regarding how BDA can be implemented into practice and how it impacts on organizational performance still remains limited.

Prior research has explored the impact that BDA has on business growth through the lens of resource-based theory (RBT), the knowledge-based view (KBV) and information-processing view (IPV). Several scholars have drawn on RBT to conceptualize a BDA capability by orchestrating tangible and intangible big data and human resources to business process and to examine its direct effects on operational and strategic-level performance (Akter *et al.*, 2016; Gupta and George, 2016; Trkman *et al.*, 2010; Wamba *et al.*, 2017; Wang and Hajli, 2017). Proponents of KBV consider BDA application an effective tool to acquire and harness knowledge that enables firms to create organizational agility and competitive advantage (Côrte-Real, Oliveira and Ruivo, 2017; Wang and Byrd, 2017; Xu, Frankwick and Ramirez, 2016). Some scholars use IPV to explain how BDA can help firms manage task complexity and respond to environmental changes through the mechanisms of information-processing (Cao, Duan and Li., 2015; Srinivasan and Swink, 2018). However, no prior research is capable of explaining the full complexity of BDA implementations or examines how and under what conditions BDA can achieve organizational performance from a holistic perspective in the healthcare context. This leads to our research question: What configurations of BDA capabilities, complementary organizational resources and organizational capabilities lead to improved healthcare performance?

In the attempt to answer this question, we first propose a conceptual model with a set of BDA success elements. A set of BDA capability that consists of BDA technological and human resources from the existing literature is identified. We then go on to explore other organizational elements (i.e. complementary organizational resources and organizational capabilities) to be considered as the potential impact of BDA on healthcare performance.

Second, we draw on the configuration theory approach (El Sawy *et al.*, 2010) to explain how BDA and other organizational elements simultaneously combine to achieve healthcare performance (i.e. low average excess readmission and high total performance score (TPS)). Configuration theory emerged from organizational research and strategic management (Fiss, 2007; Fiss, Cambré and Marx, 2003). The core concept of this theory, configuration, is defined as 'a specific combination of causal elements or conditions that generate an outcome of interest' (El Sawy *et al.*, 2010, p. 838). This approach allows us to understand how organizations can create business value from BDA by exploring the complex patterns and combinations of interconnected elements. Given that BDA's business value generation is a complex process resulting from multi-way interactions among multiple elements, we argue that configuration theory provides an excellent anchor to explain the creation of BDA's business value and explore the configurational effects of BDA capability and organizational elements on improving quality of care in healthcare. We accordingly use a set-theoretic configurational method – fuzzy-set qualitative comparative analysis (fsQCA) – as our data analysis approach.

Our study contributes to the management literature in three ways. First, this research proposes a conceptual model with a configurational lens to explicate the complexity of big data analytics implementation. To the best of our knowledge, as yet, no previous studies have considered the complex interactions among BDA and the organizational elements driving organizational performance in the healthcare context. Second, the configurations that we identified provide evidence regarding the ways in which the different relational aspects interact with each other to create high performance in healthcare. This thus extends and deepens understanding of how big data analytics can be implemented into practice. It could be a useful guidance for practitioners, outlining a variety of paths that they can follow, depending on their specific circumstances. Finally, from a methodological standpoint, this study contributes by exemplifying complementarities of fsQCA and regression-based methods. The regression-based method is suitable for explaining the causal paths through which BDA impacts organizational performance, whereas fsQCA provides a deeper understanding of the complex, non-linear and synergistic effects of BDA and organizational elements on organizational performance.

## Theoretical background and research model

*Brief review: path to big data analytics success*

Big data was first defined in terms of its volume, velocity and variety (3Vs). Then a fourth V was added, 'veracity', which refers to data accuracy that relates to quality. After which it became possible to develop more sophisticated data analysis software to fulfil the needs of handling the information explosion according to the way it is accessed, searched, processed and managed (Gandomi and Haider, 2015). While volume for big data does not have a threshold for measurement, as its form can vary depending on the time and style of its collection, it refers to the size, dimension or magnitude measured in terabytes or petabytes (Demchenko *et al.*, 2013). Big data is also explained in terms of variety that explains the structural heterogeneity in a data set as structured when found in forms of spreadsheets or relational databases; or unstructured data in the form of videos, audios, images, text or tables (Gandomi and Haider, 2015). Velocity in reference to big data reflects the speed, rate or cost at which data are generated using smartphones or other technological advancements such as biometric technologies to be analysed (Demchenko *et al.*, 2013). Scholars such as Demchenko *et al.* (2013) have expanded the traditional 3Vs-based definition of big data to its value-generation capacity from the analysis based on volume and variety of data available to the analyst.

Veracity is directly related to data quality, as it refers to the inherent biases, noise and abnormality in data. Veracity also includes data consistency (defined by the statistical reliability of data) and data trustworthiness (based on data origin, data collection and processing methods and security infrastructure). These data quality issues, in turn, impact data integrity and data accountability. This characteristic of big data presents its importance and challenges, especially in healthcare, in needing a high level of data analytical capability because it affects the decisions concerning patients' well-being.

The literature has proposed several conceptual frameworks to explain how to implement BDA in organizations grounded in RBT, KBV and IPV, as summarized in Table 1 and visualised in Figure 1. Based on RBT, much of the research on big data has found that the different types of BDA resources (e.g. physical, technical and human resources) can add value to firms' operations. These resources can develop BDA-specific capabilities that firms could use to gain meaningful insights and reshape organizational performance. For example, Seddon, Constantinidis and Dod (2012) argue that the functional fit of BDA tools and readily available high-quality data, and staff with good analytical skills, are predictors that positively influence the benefits gained from ongoing BDA improvement projects. A recent study conducted by Wamba *et al.* (2017) indicated that BDA infrastructure capability, management capability and personnel capability all have a strong effect on firm performance.

As an extension of RBT, the KBV views knowledge as a value, rare, inimitable and non-substitutable (VRIN) resource, and argues that knowledge absorption plays a critical role in acquiring new knowledge (Grant, 1996). Côrte-Real, Oliveira and Ruivo (2017) have adopted this theory to develop BDA-enabled knowledge assets, namely exogenous knowledge management, endogenous knowledge management and knowledge-sharing with partners. They provide evidence from a survey of 500 European firms to suggest that these BDA-enabled knowledge assets create firms' organizational agility, thereby strengthening their competitive advantage. With an emphasis on knowledge absorption, Wang and Byrd (2017) indicate that the effective use of data analysis and interpretation tools in healthcare units indirectly influence decision-making effectiveness through the mediating role of knowledge absorptive capacity.

The IPV argues that organizations' performance depends on their ability to process information (Galbraith, 1974). According to this view, Srinivasan and Swink (2018) suggest that the application of BDA lies at the heart of the organizational information process, since it enhances firms' ability to collect, disseminate, store, analyse and display information, all of which strengthen firms' capability to process information. To facilitate organizational information-processing capability, prior research has emphasized that organizations should design their organizational structure, mechanism and business processes in conjunction with data analysis processes, which may reduce the environmental uncertainty and ambiguity of the problem context (Kowalczyk and Buxmann, 2014; Sharma, Mithas and Kankanhalli, 2014). As regards supply chains, for example, Trkman *et al.* (2010) report

Table 1. *Summary of the literature on BDA models*

| Study | Theoretical base | Methodology – Description | Factors leading to BDA success | | | Outcomes gained by BDA success | Causality |
|---|---|---|---|---|---|---|---|
| | | | BDAC | COR | OC | | |
| Akter *et al.* (2016) | RBT | Empirical – A survey of 152 BDA professionals toward understanding the impact of BDA capabilities on firm performance | V | | | Firm performance | Linear relationship |
| Cao, Duan and Li (2015) | IPV | Empirical – A survey of 740 responses collected from UK businesses toward understanding the impact of business analytics usage on organizational decision-making effectiveness | V | V | V | Decision-making effectiveness | Linear relationship |
| Fink, Yogev and Even (2017) | Contingency theory | Mixed methods – Develop and test a model of business intelligent (BI) value creation that links BI team, BI infrastructure, operational and strategic capabilities to business value | V | | | Operational and strategic business value | Linear relationship |
| Gupta and George (2016) | RBT | Empirical – Create an instrument to measure BDA capabilities | V | | | Market and operational performance | Linear relationship |
| Popovič *et al.* (2012) | Information systems success model | Empirical – A survey-based study of the impact of business intelligence systems maturity on the quality of information content | V | V (as a moderator) | | The use of information in business process | Linear relationship |
| Srinivasan and Swink (2018) | IPV | Empirical – A survey of 191 global firms toward examining the impact of supply chain analytics capability on operational performance | V | | V (as a moderator) | Cost performance and delivery performance | Linear relationship |
| Trkman *et al.* (2010) | Supply chain operations reference model | Empirical – A survey of 310 responses from various industries toward investigating how the use of analytics on supply chain (SC) process influences SC performance | V | | | Supply chain performance | Linear relationship |
| Wamba *et al.* (2017) | RBT and dynamic capability view | Empirical – A survey of 297 Chinese IT managers toward examining factors that contribute to improved firm performance | V | | V | Firm performance | Linear relationship |

(*Continued*)

*Table 1. Continued*

| Study | Theoretical base | Methodology – Description | Factors leading to BDA success | | | Outcomes gained by BDA success | Causality |
|---|---|---|---|---|---|---|---|
| | | | BDAC | COR | OC | | |
| Wang and Hajli (2017) | RBT | Case study – Develop a BDA-enabled business value model to explore the cause-and-effect relationship between BDA capabilities and business value | V | | | Potential benefits of BDA | Linear relationship |
| Wang, Kung, and Byrd (2018a) | RBT | Case study – Identify various BDA functionalities that in combination build BDA capabilities from 26 published case studies | V | | | Potential benefits of BDA | – |
| Wixom, Yen and Relich (2013) | – | Case study – Explore two key factors and their underlying dimensions for maximizing big data analytics value in a fashion retailer case | V | | | Transactional, informational, and strategic vale | – |
| Xu, Frankwick and Ramirez (2016) | KBV | Conceptual – A theoretical framework that links traditional marketing analytics and BDA to new product success | V | | | New product success | Linear relationship |
| Côrte-Real, Oliveira and Ruivo (2017) | KBV | Empirical – A survey of 500 European firms to suggest that these BDA enabled knowledge assets create firms' organizational agility | V | | | Organizational agility | Linear relationship |
| Current study | Configuration theory | Empirical – Using a multi-source data set to examine how BDA capabilities interact with complementary organizational resources and organizational capabilities in multiple configurations to achieve quality of care | V | V | V | Quality of care | Non-linear (interaction) relationship |

BDAC: big data analytics-enabled capabilities; COR: complementary organizational resources; OC: organizational capabilities.

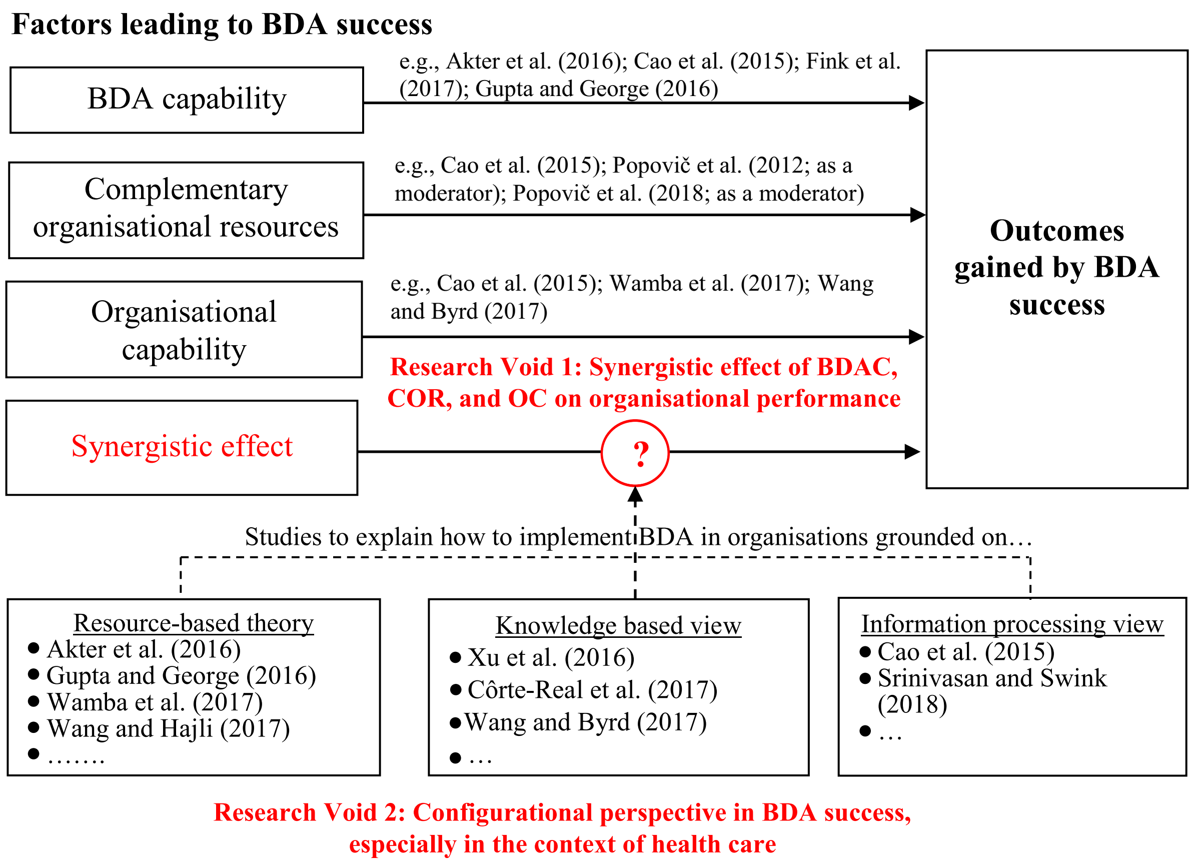


*Figure 1. Literature on determinants of big data analytics success [Colour figure can be viewed at wileyonlinelibrary.com]*

that firms with the ability to analyse and use their information within the different stages of the supply chain (i.e. plan, source, make and deliver) enjoy a superior supply chain performance as a result. In the same vein, Cao, Duan and Li (2015) have found that using BDA influences information-processing capability through the mediation of a data-driven environment, which, in turn, has a positive effect on decision-making effectiveness. These studies explore ways in which business decisions are made through the joint effects of BDA and information-processing mechanisms.

### *Research model of configurations producing organizational performance*

Business value of information technology (IT) literature (e.g. Bharadwaj, 2000; Melville, Kraemer and Gurbaxani, 2004; Nevo and Wade, 2010) contends that IT alone does not unequivocally facilitate organizational performance. Indeed, IT business value creation is a complex process that cannot be fully explained by a set of factors and regression-based methods, but instead involves the systemic and simultaneous arrangement of multiple elements. The link between IT and organizational performance is not likely to be straightforward in terms of the multi-way interactions among the IT elements (e.g. IT infrastructure and IT applications) and other organizational elements (organizational structure and culture). Researchers have emphasized that various complementarities, such as organizational culture, policies and rules, organizational structure and environmental conditions, should interact with IT to generate superior organizational performance (Fichman, 2004; Melville, Kraemer and Gurbaxani, 2004; Nevo and Wade, 2010). For instance, Tanriverdi (2006), who investigates the effects of information technology synergies, ascribes them to a combination of IT resources, namely IT infrastructure, IT strategy, IT human resource, and IT vendor management. In the management research, Zammuto *et al.* (2007) suggest that it is important to understand the process and outcomes of a combination of IT processes and organizations, since either IT or organizational aspects alone would not provide a complete picture of IT business value creation. In information systems strategy research, El Sawy *et al.* (2010) argue that acquiring strategic advantage in today's turbulent environments is complex, and IT resources alone are not sufficient to explain this complexity. Using fsQCA, El Sawy *et al.* (2010) examined how IT systems, dynamic capability and environmental turbulence interact as digital ecodynamic systems that produce strategic advantages in turbulent environments. These studies extend the theoretical perspective that 'business value should be rooted in the identification of IT resources' to encompass 'seeking … the best configuration of possible IT resources' (Schryen, 2013). Thus, we adopt configuration theory as a theoretical basis of our research model to explain the complex interactions among BDA capabilities and complementary organizational resources and organizational capabilities and their effects on each other to co-create a higher performance in a healthcare context.

To justify the inclusion of key elements in our research model, we employ the logic structure and rationale of the IT business value generation framework proposed by Melville, Kraemer and Gurbaxani (2004). This framework demonstrates how the business value of IT can be intensified by the bundling of resources (i.e. technology IT resources, human IT resources and complementary organizational resources) and the synthesis and integration of business processes. This framework expands and deepens understanding of the RBT in an IT context by specifying the underlying mechanisms driving the way in which IT resources are applied within business processes to improve organizational performance. This framework also explains that the inimitability of rare organizational resources is complementary to technological IT resources and that human IT expertise has a significant potential to improve the operational efficiency of business processes, which in turn spurs economic value for a focal firm.

Applying this framework to the BDA context, we identified: (1) six general categories of BDA capabilities – data integration, analytical, data interpretation, predictive and the technical and business skills of analytics personnel, as components of technological and human IT resources from the extant literature; (2) two complementary organizational resources – evidence-based decision-making culture and data governance; and (3) two organizational capabilities embedded in the business process – planned dynamic and improvisational capabilities. These elements can be combined in various potential configurations to determine which options result in improved healthcare performance. Figure 2 illustrates the interactions among these three configuration elements of BDA, with the intersecting orbits representing a holistic confluence that will subsequently contribute to an enhanced quality of care in healthcare. The ten elements included in our configurational analysis are described in the next section.

### *Elements of big data analytics capabilities*

Big data analytics capability is defined as the ability to acquire, store, process and analyse 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 (Wang and Hajli, 2017). We propose four dimensions of BDA capability in healthcare: (1) data integration capability; (2) analytical capability; (3) predictive capability; and (4) data interpretation capability, as described below in more detail. The key functionalities and applications in healthcare for each BDA capability are summarized in Table 2.

*Data integration capability.* Data integration capability is defined as the ability to transform diverse types of data into a data format that can be read and analysed by the data analysis platform (Wang and Byrd, 2017). Through three key functionalities of data integration in BDA systems (i.e. acquisition, transformation and storage), data can be consistent, visible, easily accessible and interoperable for analysis (Raghupathi and Raghupathi, 2014). A high level of data integration allows healthcare organizations to aggregate intelligently data such as clinical data, billing/insurance data, pharmaceutical R&D data, and patient behaviour data via extract-transform-load (ELT) tools and provides users with a comprehensive view of these data (Wang *et al.*, 2018b). To make better use of healthcare data, Hsu and Griese (2018) suggest that healthcare organizations should have centralized data deposited in standalone virtual databases linking all data silos

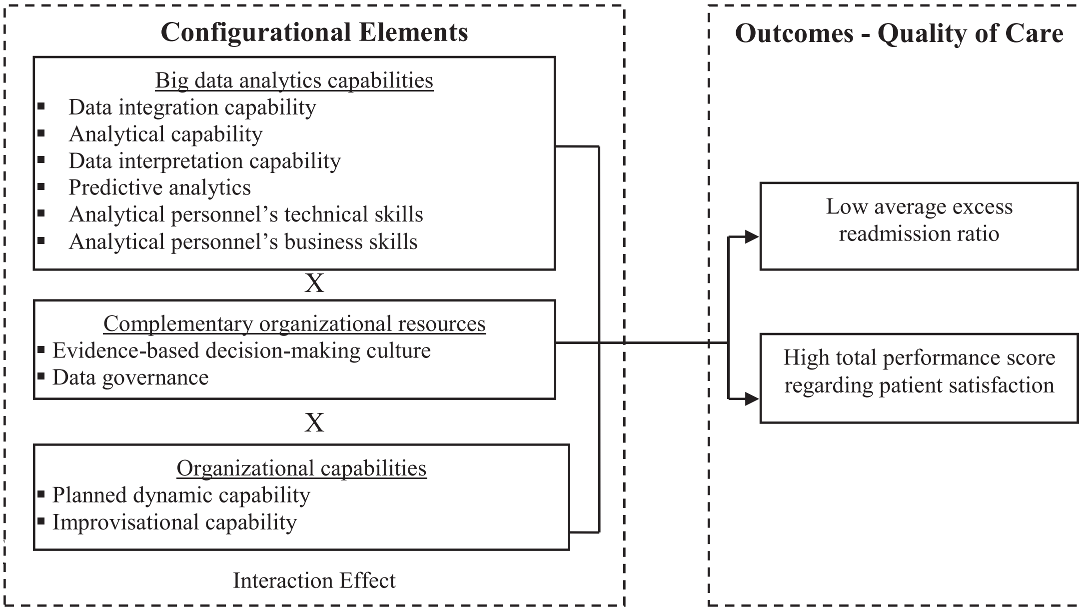


*Figure 2. Research model*

*Table 2. Key functionality and application of BDA capability*

| BDA capability | Functionalities and tools | Key applications in health care |
|---|---|---|
| Data integration | Middleware<br>Data warehouse<br>Extract-transform-load (ELT) tools<br>Hadoop distributed file system (HDFS)<br>NoSQL database | Integrate seamlessly clinical data across multiple regions or facilities in real time or near real time<br>Track medical events based on the rules that have been built on hospital claims<br>Search clinical databases for all data related to patient characteristics and conditions |
| Analytical | Basic statistical analysis<br>Online analytical processing (OLAP) | Analyse large amounts of clinical data to understand the past and current state for specific target variables<br>Explore the causes of occurred medical events from relational databases<br>Support real-time processing of multiple clinical data streams |
| Predictive | Regression techniques<br>Predictive modelling<br>Social media analytics<br>Machine learning<br>Text mining/Natural Language Processing | Examine undetected correlations, patterns, trends between specific variables of interest across regions or facilities<br>Comparison and cross-referencing of current and historical data and their outcomes to predict future trends<br>Provide actionable insights or recommendations in a format readily understood by its users |
| Data interpretation | Visual dashboards/systems<br>Reporting systems/interfaces | Generate clinical summary (or performance metrics) in real time or near real time and present in visual dashboards/systems<br>Provide system outputs for role-based decision-making |

for review by medical staff when needed. Data integration driven by BDA systems should also allow users to track the data created by devices worn by individual patients and collect them in real time or near real time, making it possible to gather location, event and physiological information, including time stamps, from each patient wearing a device. Hence, since data integration capabilities support healthcare services in value-adding ways, this is viewed as one of the key BDA capabilities in healthcare.

*Analytical capability.* Analytical capability refers to the ability to drive decisions and actions through the extensive use of data and different analytical techniques based on the specific mechanisms used for analytics, thus addressing the various needs of users and other stakeholders (Ghosh and Scott, 2011). In healthcare, the use of analytical tools that can support core clinical operations and processes is particularly important as a means of increasing the quality of care (Raghupathi and Raghupathi, 2014). Healthcare analytical systems allow users to identify patterns of care and discover associations from massive collections of healthcare records, thus providing a broader view for evidence-based clinical practice. Such analysis can identify previously unnoticed patterns in patients related to hospital readmissions and support a better balance between capacity and cost. For example, one effective analytical technique, descriptive analytics, has been widely used in BDA systems (Watson, 2014). In a hospital setting, this technique enables users to understand past patient behaviours and how these behaviours might affect outcomes based on the information stored in their database. Most importantly, the ability to analyse patient preferences helps hospitals to recognize the utility of participating in clinical trials and identify new potential markets. Data analysis can thus help to increase the efficiency of healthcare delivery, leading us to include analytical capability as a key dimension of BDA capability.

*Predictive capability.* Predictive capability is 'the process of using a set of sophisticated tools to develop models and estimations of what the environment will do in the future' (Wessler, 2013, p. 21). It is the ability to apply diverse statistical analysis methods, modelling, machine learning and data-mining to both structured and unstructured data to determine future outcomes. Predictive analysis makes it possible to cross reference current and historical data to generate context-aware recommendations that enable managers to make predictions about future events and trends. This capability relies on predictive analytical engines that incorporate a data warehouse, a predictive platform with predictive algorithms (e.g. decision trees, neural networks and logistic regression) and a predictive interface that provides feedback and recommendations to users.

Predictive capabilities can reduce the degree of uncertainty, enabling managers to make better decisions faster and hence support preventive care. The Texas Health Harris Methodist Hospital Alliance, for example, analyses information from medical sensors to predict patients' movements and thus provide needed services more efficiently. It also monitors patients' actions throughout their hospital stay to help reduce medical risk. For instance, Samorani and LaGanga (2015) combined the predictive analytics with optimization to tackle the problem of overbooking appointments, given the predictions of patients' no-show behaviour. Thus, healthcare entities with superior big data predictive capabilities should be able to leverage helpful predictive reports to improve decision-making, optimize existing operations and provide high-quality healthcare services.

*Data interpretation capability.* Data interpretation capability emphasizes the ability to produce a healthcare matrix and reports that evaluate patient care and service and identify areas for improvement. In general, data interpretation tools such as dashboards and reporting interfaces yield historical reporting, executive summaries, drill-down queries, statistical analyses and time series comparisons. These outputs can provide a comprehensive view that supports the implementation of evidence-based medicine, detects advanced warnings for disease surveillance, and helps develop personalized patient care (Ghosh and Scott, 2011).

Data interpretation tools enable data to be visualized in various formats, such as interactive dashboards and charts that support physicians and nurses' daily operations and help them to make faster and more rational evidence-based decisions (Roski, Bo-Linn and Andrews, 2014). For example, a Dutch long-term care institution has visualized the number of incidents, the locations where the incidents occurred and the type of physical damage that resulted by mining a collection of 5692 incidents that occurred over a four-year period (Sprui, Vroon and Batenburg 2014). Displaying frequency tables in the form of visual dashboards has enabled this Dutch long-term care institution to improve patient safety. Therefore, as data interpretation is a critical feature of BDA systems, we propose data interpretation capability as a key element of BDA capability.

*Technical and business skills of analytical personnel.* Davenport, Harris and Morison (2010) define analytical personnel as the members of an organization who have an analytical mindset and help derive value from BDA. Analytical staff fulfil a hybrid role that requires a broad combination of technical and soft skills and multidisciplinary knowledge domains. The skill sets for analytical personnel have been thoroughly investigated by researchers. The skills needed by well-qualified analytics personnel are summarized in Appendix A. Based on their different levels of data analytical skills, Wilder and Ozgur (2015) categorize analytical staff as data scientists, data specialists and big data analysts. Data scientists understand how to extract answers to important questions from the tsunami of unstructured information available to them (Davenport and Patil, 2012), while data specialists not only have a solid foundation in computer science, mathematics and management, but also understand how data are managed (Wilder and Ozgur, 2015). Business analysts, who often hold a title such as Chief Data Officer, are key leaders in an organization and are responsible for establishing sound governance to ensure data quality, using data-driven insights to make sound decisions, identify business opportunities and address business problems (Lee *et al.*, 2014). Managers and employees with relevant professional analytical competencies represent a crucial element for BDA success, since incorrect interpretations of the reports generated could lead to serious errors of judgment and questionable decisions. Indeed, the success of a BDA project depends on the ability of the organization's analytical staff to understand not only the overall business environment, but also the specific organizational context of the data they work with. Surprisingly little has been reported regarding the role of the analytical personnel as an enabler of BDA success in the existing literature.

The six elements of BDA capabilities discussed above are related, but distinct. The BDA capability elements by themselves may not explain the actual patterns that reflect the mechanism of the influence of BDA implementation on an outcome of interest. Instead, their interactions and combinations with other organizational elements such as complementary resources and organizational capabilities may determine their role in business value (El Sawy *et al.*, 2010; Melville, Kraemer and Gurbaxani, 2004; Ragin, 2008a). Therefore, we also examined other organizational elements that may influence healthcare performance, along with these BDA capabilities.

### *Elements of complementary organizational resources*

Companies who are eager to implement BDA to create business value must undergo adjustments or even dramatic changes in their day-to-day operations, data policies and organizational culture (Davenport, Harris and Morison, 2010; LaValle *et al.*, 2011). Complementary organizational resources are required for a successful BDA implementation (Watson, 2014). Especially in healthcare, such resources help organizations face the challenges of standardizing many different types of data across various healthcare systems and resources (Shah and Pathak, 2014). Big data analytics-enabled complementary organizational resources are regarded as a specific type of organizational resource with the aid of BDA that tend to be tacit, idiosyncratic and deeply embedded in the organization. Key complementary organizational resources in the context of BDA, such as enterprise-wide analytics orientation (Seddon, Constantinidis and Dod, 2012) and a fact-based decision-making culture (Seddon, Constantinidis and Dod, 2012; Watson, 2014), have been recognized as key drivers of superior organizational performance. In this study, we select an evidence-based decision-making culture and data governance as the primary complementary organizational resources and describe them in the next section.

*Evidence-based decision-making culture.* Organizational culture plays an important role in enabling an organization to create a business value with analytics (Kiron and Shockley, 2011). Organizational culture is defined as a set of collective values, beliefs, norms and principles that guide what happens in organizations by defining appropriate behaviour for various situations (Ravasi and Schultz, 2006). Many studies have reported that organizational culture represents a major hurdle hindering the widespread use of fact-based decision-making (e.g. Kiron and Shockley, 2011; LaValle *et al.*, 2011); shifting the decision-making process away from intuitive thinking and individual experience to 'the facts' facilitated by BDA is a challenging undertaking for an organization (Watson, 2014).

This study focuses on a particular aspect of organizational culture from a BDA perspective, namely an evidence-based decision-making culture, defined as a culture of embracing evidence-based management and embedding evidence-based decision-making in the core values and processes of the organization (Davenport, Harris and Morison, 2010). Kettinger, Zhang and Marchand (2011) describe this concept as an information-oriented culture where business executives have a heightened awareness of information and information management as they make decisions or formulate business strategies. Kiron, Ferguson and Prentice (2013) view this as a data-driven culture, defined as 'a pattern of behaviours and practices by a group of people who share a belief that having, understanding and using certain kinds of data and information plays a critical role in the success of their organisation' (Kiron, Ferguson and Prentice, 2013, p. 18). Researchers suggest that successful analytics use is most likely when an evidence-based decision-making culture is rooted in the enterprise's key business processes, and that this kind of culture would tend to inspire an organization to measure, test and evaluate quantitative evidence (Davenport, 2006; Kiron, Prentice and Ferguson, 2012). Popovič *et al.* (2012) found that an organization with an analytical decision-making culture can positively affect the quality of information provided by business intelligence systems. As noted by Ross, Beath and Quaadgras (2013), building an evidence-based decision-making culture in an organization should ensure that all decision-makers share performance metrics that originate from one undisputed source, provide decision-makers at all levels with near-real-time feedback, articulate business rules and update them with new facts when necessary and provide high-quality coaching to decision-makers on a regular basis. An evidence-based decision-making culture would allow healthcare organizations to make better use of real-time data, making more accurate diagnoses and better treatment decisions and offering more reliable care to patients.

*Data governance.* Data governance plays a critical role in BDA implementation to ensure the quality, security, privacy and life cycle of the data that is collected and stored (Khatri and Brown, 2010). Data governance that is built on IT governance aims to formulate data rules and policies and provide a vision and guidelines relating to privacy, security, life cycle and ownership of data by aligning the objectives of multiple functions (Kooper, Maes and Lindgreen, 2011; LaValle *et al.*, 2011). Typically, a data governance framework is composed of master data management (MDM), data life cycle management and data security and privacy management. Master data management is regarded as the processes, governance, policies, standards and tools for collecting, aggregating, matching, consolidating, quality-assuring, persisting and distributing data throughout an organization (Loshin, 2010). The aim of data management is to ensure that data is properly standardized, removed and incorporated to create the immediacy, completeness, accuracy and availability of data needed to support data analysis and decision-making. Data life-cycle management is the process of managing business information throughout its life cycle, from archiving data, via maintaining a data warehouse, testing and delivering different application systems, to deleting and disposing of data (Jagadish *et al.*, 2014). Data security and privacy management is the platform for providing enterprise-level data activities in terms of discovery, configuration assessment, monitoring, auditing and protection. Khatri and Brown (2010) proposed a hierarchical framework that includes five interrelated decision domains – data principles, data quality, metadata, data access and data life cycle – for assessing the effectiveness of data governance when implementing BDA in an organization.

The key to successful data governance is not technology or methods; instead, it is about practices and people in the organization and their complex ownership of the data that the BDA initiative will affect. Cao, Duan and Li (2015) describe this concept in an organization's data-driven environment as 'the organisational practices reflected by developing explicit data strategy and policy to guide analytic activities and designing its structure and process to enable and facilitate BDA activities' (Cao, Duan and Li, 2015, p. 385). Data governance can also be viewed as a set of policies, a way of working or a framework of optimizing the value of information in some sense to the decision-makers involved (Kooper, Maes and Lindgreen, 2011). As Davenport and Harris (2007) suggest, establishing robust data governance is the first step in implementing BDA. In hospitals, for example, establishing rigorous data policies and data access control mechanisms for highly sensitive healthcare data can prevent security breaches and protect patient privacy (Wang, Kung, and Byrd, 2018a). Adopting suitable data policies, standards and compliance requirements to restrict users' permissions will ensure the system satisfies healthcare regulations and creates a safe environment for the proper use of patient information. Therefore, we include data governance as an important element in achieving healthcare performance configurations.

*Elements of organizational capabilities*

To achieve the vast potential of BDA, not only will enterprise IT architectures need to change, but almost every department within a company will also undergo adjustments (Davenport, Harris and Morison, 2010). Managing BDA is not merely a simple technical issue, but also a managerial and strategic one (McAfee and Brynjolfsson, 2012). Thus, organizational capabilities have been shown to be significant predictors of BDA success (Wamba *et al.*, 2017; Wang and Byrd, 2017). In general, organizational capability is defined as the ability to adapt to ongoing changes in the business processes and functional activities of the firm (Luo, Fan and Zhang, 2012). It has also been described as 'an organisation's ability to create value in a unique way by utilizing resources' (Wu and Hu, 2012, p. 981) from the RBT perspective.

From a dynamic capability perspective, two types of distinctive organizational capabilities – planned dynamic capability and improvisational capability – have been identified from among the core business processes for boosting business value (Pavlou and El Sawy, 2010). El Sawy *et al.* (2010) have highlighted the role of IT systems in shaping these two capabilities and inducing environmental turbulence to help build a strategic advantage within digital ecosystems. Planned dynamic capability is a firm's high-level organizational ability to integrate, reconfigure, gain and renew resources to match rapidly changing market environments (Easterby-Smith, Lyles and Peteraf, 2009; Eisenhardt and Martin, 2000; Teece, Pisano and Shuen, 1997) and enhance a firm's agility (Roberts and Grover, 2012). Barreto (2010) and Teece (2007) regard dynamic capability viewing as the ability to sense and shape opportunities and threats, seize market opportunities and maintain competitiveness. In big data research, dynamic capabilities that are triggered by BDA capabilities have been identified as intermediate variables that contribute to firm performance (Côrte-Real, Oliveira and Ruivo, 2017; Wamba *et al.*, 2017).

Improvisational capability is defined as an organization's learned ability to respond to unexpected environmental turbulence quickly, simultaneously forming and executing novel solutions by reconfiguring available resources (El Sawy and Pavlou, 2008). Research from both strategic and organizational management fields has emphasized the importance of organizational improvisation in handling extreme competition, coping with changing circumstances and pursuing potential business opportunities (e.g. Bergh and Lim, 2008; Hadida and Tarvainen, 2014; Moorman and Miner, 1998). Improvisational capability plays a crucial role in building an organization's agility when reacting to market changes. Such 'spontaneous' capabilities enable organizations to make effective and real-time decisions in response to turbulence without having to go through formal planning channels. We thus include planned dynamic capability and improvisational capability as two important organizational capabilities for achieving healthcare performance with BDA.

## Research methods

*Data collection*

The healthcare industry was selected as the research context for this study for two reasons: (1) BDA implementation and the study of its effects in healthcare industries are lagging far behind other industries such as retail and banking (Raghupathi and Raghupathi, 2014); and (2) focusing on a single industry can mitigate any potential confounding effects due to industry nature and variation. We tested our model using a multi-source data set acquired from a survey and databases maintained by the Centres for Medicare & Medicaid Services (CMS).

An initial population set of 4668 senior IT executives (primarily vice presidents, CIOs and IT directors) in US hospitals, listing the facility name, job title, phone number and email address for each, was extracted from the Healthcare Information and Management Systems Society (HIMMS) database. After data cleaning to remove incomplete information and duplicates, 3307 senior information system executives remained. An online survey was specifically designed for this study. The first round of 3307 questionnaires resulted in 511 emails being blocked by their organizations' firewall and 1589 emails that were never opened; a gentle reminder was sent a week later. Of the 1207 invitations that were seen by potential respondents, 65 responses were returned, 63 of which were complete and usable for the data analysis, showing a response rate of 5.39%. According to Armstrong and Overton (1977), non-response bias was assessed by comparing the early (first 25%) and late respondents (last 25% that are equivalent to non-respondents) for every measurement using paired sample *t*-tests. The results showed no statistically significant difference ($p > 0.05$) between these two groups, indicating that non-response bias did not present a problem for this study.

The CMS data were downloaded from the Hospital Compare website.[1] This website provides information on how well hospitals provide healthcare service to their patients and allows them to compare performance metrics related to certain conditions. We extracted average excess readmission ratios (AERR) and TPSs from the CMS database to evaluate the quality of care as the outcome for this study. We were able to match CMS data to our survey data in 34 cases for AERR and 29 cases for TPS.

### *Measurement validity and reliability*

The definitions of key constructs and measurements used in the current study are presented in Appendix B. Most measurement items were adopted from the literature and modified to fit the context of this study. The measurements and items of constructs are presented in Appendix C. Except for the outcome variables, a 7-point Likert-type scale (1 = strongly disagree, 7 = strongly agree) was used for all the constructs.

The validity and reliability of measurements were assessed from the sample data set ($n$ = 63) collected for this study. All of the reliability coefficients (Cronbach's alphas) are above 0.70 (Table 4), confirming that the measurements are reliable. Convergent validity was assessed by three criteria: (1) item loading; (2) composite reliability; and (3) average variance extracted (AVE) (Fornell and Larcker, 1981). As shown in Table 2 and Appendix D, the loadings are all within acceptable ranges, and all but one item for data governance have loadings above the threshold of 0.7. The single item that drops below this level has a loading of 0.650, which exceeds the acceptable threshold of 0.6 proposed by other scholars (Chin, 1998). The composite reliability scores range from 0.85 to 0.94. Each AVE is above 0.5 (Table 4), indicating that the latent construct can account for at least 50% of the variance in these items. Two methods were employed to assess discriminant validity: (1) checking whether each item loads more highly on its assigned construct than on other constructs, as suggested by Gefen, Straub and Boudreau (2000); and (2) checking whether each construct's square root of AVE is greater than its correlations with other constructs (Table 3) (Fornell and Larcker, 1981). Each item loading in the cross-loading table (Appendix D) is markedly higher on its assigned construct than on the other constructs. The square root of the AVE is greater than all of the inter-construct correlations (Chin, 1998). Thus, our measurements demonstrate sufficient discriminant and convergent validities.

In addition, to reduce common method bias, we protected respondent–researcher anonymity, provided clear directions and separated the independent and dependent variables (Podsakoff *et al.*, 2003). We then assessed the potential effect of common method bias statistically by conducting two tests. First, Harman's one-factor test (Podsakoff and Organ, 1986) generated ten principal constructs; the unrotated factor solution shows that the first construct explains only 16.74% of the variance, indicating that our data do not suffer from high common method bias. Second, we compared correlations among the constructs. The results revealed no constructs with correlations over 0.7, whereas evidence of common method bias ought to have brought about significantly higher correlations ($r < 0.90$) (Bagozzi, Yi and Phillips, 1991). Consequently, these tests suggest that common method bias is not a major concern for this study.

### *Analysis method: fsQCA*

Fuzzy-set qualitative comparative analysis was used to explain how BDA capability, organizational resources and organizational capability elements simultaneously combined to create high quality of care. As our study is to examine the modelling of asymmetric relationships between variables, fsQCA provides several benefits to our study. First, fsQCA focuses squarely on the 'middle ground' between variable-oriented quantitative methods and case-oriented qualitative methods (Ragin, 2000, p. 22). It thus allows for evaluating case studies with few cases for standard statistical analyses. Second, regression-based analysis is limited to two-way or three-way interaction effects, while cluster analysis only discovers homogeneous patterns without control over the outcome (Fiss, 2007). Fuzzy-set qualitative comparative analysis

[1]www.medicare.gov/hospitalcompare

*Table 3. Inter-construct correlations*

| Constructs | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Data integration capability | **0.87** | | | | | | | | | |
| Analytical capability | 0.06 | **0.91** | | | | | | | | |
| Data interpretation capability | 0.19 | 0.25 | **0.94** | | | | | | | |
| Predictive capability | 0.09 | 0.20 | 0.19 | **0.89** | | | | | | |
| Technical skills | 0.37** | −0.21 | −0.19 | 0.31* | **0.88** | | | | | |
| Business skills | 0.05 | 0.02 | 0.17 | 0.25* | 0.23 | **0.81** | | | | |
| Decision-making culture | 0.14 | 0.16 | 0.17 | 0.03 | −0.11 | 0.16 | **0.88** | | | |
| Data governance | −0.09 | 0.21 | −0.27* | −0.08 | 0.11 | 0.10 | −0.26* | **0.74** | | |
| Dynamic capability | 0.34** | −0.01 | 0.10 | 0.04 | 0.32** | −0.08 | −0.18 | −0.06 | **0.86** | |
| Improvisational capability | −0.05 | −0.42** | 0.19 | 0.17 | −0.15 | 0.07 | 0.24 | −0.09 | −0.10 | **0.89** |

*Note:* $n = 63$; square root of AVE values are in bold $*p < 0.05$; $**p < 0.01$.

*Table 4. Reliability and validity measures of the research model*

| Construct | Items | Mean | SD | Loading | Composite reliability | AVE | α |
|---|---|---|---|---|---|---|---|
| Data integration capability | 3 | 4.70 | 0.99 | 0.782–0.886 | 0.90 | 0.75 | 0.85 |
| Analytical capability | 3 | 4.27 | 1.23 | 0.802–0.913 | 0.94 | 0.83 | 0.90 |
| Data interpretation capability | 2 | 4.60 | 1.55 | 0.843–0.907 | 0.94 | 0.89 | 0.89 |
| Predictive capability | 3 | 4.34 | 1.06 | 0.832–0.867 | 0.92 | 0.79 | 0.87 |
| Technical skills | 4 | 5.13 | 1.17 | 0.792–0.865 | 0.94 | 0.78 | 0.90 |
| Business skills | 3 | 4.52 | 0.98 | 0.731–0.853 | 0.85 | 0.65 | 0.79 |
| Decision-making culture | 3 | 3.80 | 1.27 | 0.778–0.922 | 0.91 | 0.78 | 0.86 |
| Data governance | 5 | 3.63 | 0.92 | 0.650–0.812 | 0.85 | 0.55 | 0.84 |
| Dynamic capabilities | 4 | 3.55 | 1.28 | 0.715–0.915 | 0.92 | 0.74 | 0.88 |
| Improvisational capabilities | 3 | 3.58 | 1.17 | 0.827–0.898 | 0.92 | 0.80 | 0.87 |
| Average excess readmission ratio | – | 0.999 | 0.058 | – | – | – | – |
| Total performance score | – | 40.603 | 11.452 | – | – | – | – |

takes the perspective that cases are composed of combinations of theoretically relevant attributes and that the relationships between these attributes and the outcome of interest can be understood through an examination of the subset relations (Ragin, 2000, 2008b). Fuzzy-set qualitative comparative analysis could be the best approach to deal with multi-way interactions and examine how variables systemically combine to create outcomes (Misangyi *et al.*, 2017).

*Calibration.* A critical step in a fsQCA analysis is to carefully convert data into measures of set membership using theoretical or substantive knowledge external to the empirical data – a process known as calibration. This process transforms interval scale values to fuzzy-set membership scores based on three qualitative anchors: full membership; full non-membership; and the crossover point of maximum ambiguity regarding membership in the set of interest (Rai, Patnayakuni and Seth, 2006). We followed Ragin (2008a) in calibrating fuzzy-set memberships. For each calibration, we set thresholds based on industry common standards, if available, extant theory or substantive knowledge. We used the direct method of calibration in the fsQCA software to transform the measures into set memberships (e.g. Fiss, 2011; Ragin, 2008a). To some extent, survey items that are measured on a Likert scale have built-in membership scores.

As mentioned earlier, we opted to use average excess readmission ratio and TPS as our outcomes. For both measures of quality, we calculated both the national average and the standard deviation. For the first measure of quality of care using the average excess readmission ratio, we set up a 'low average excess readmission ratio' set, because the lower the ratio the better the quality. A national excess readmission ratio average was calculated by taking the mean of the rate for over 3500 hospitals located across the US as the industry standard and the base value to evaluate the membership scores. We also calculated the standard deviation. The cut-off point for full membership for this set was then set as the result of the national average excess readmission ratio minus 1 SD (0.92); the anchor for the crossover point was 0.99, the national average excess readmission ratio; and the cut-off point for full non-membership was set at the value of the national average excess readmission ratio plus 1 SD (1.10).

For the second measure of quality, we established a 'high TPS' set because, as with most performance measures, the higher the score the higher the quality. Two domains, patient experience of care and clinical process of care, were used to assess hospital performance. A performance score and an improvement score were calculated for each measure, and a domain score calculated for each of the two domains. The TPS was calculated using the weighted domain scores. The Clinical Process of Care domain score was weighted as 70% of the TPS, and the Patient Experience of Care domain was weighted as 30% of the TPS. Using the same statistical measures, the cut-off point for full membership in the high TPS set was 53.14 (national TPS plus 1SD), 40.48 was the score for the crossover point (national TPS), and 27.82 (national TPS minus 1 SD) was the fully-not-in-the-set point.

The configuration conditions selected for this study were: the six BDA capabilities; two complementary organizational resources (i.e. evidence-based decision-making culture and data governance); and two organizational capabilities (i.e. planned dynamic capabilities and improvisational capabilities). All the items for the variables except the BDA capabilities were extracted from the literature and measured using validated scales. As this study used a 7-point Likert scale for the construct survey, we set up the high-level membership sets using 6 as the fully-in-the-set cut-off point, 4 as the crossover point, and 2 as the fully-not-in-the-set point.

*Truth table analysis.* After calibration, sets can be subjected to fuzzy truth table analysis to examine the relationship between the configuration conditions and the outcome. Scholars have recommended testing the conditions that might be necessary to achieve the desired outcome before analysing sufficiency (Legewie, 2013), where a 'necessary' condition is defined as a condition such that the outcome would not have occurred in its absence. After the necessary conditions analysis, we then ran the truth table algorithm, choosing the outcome and conditions, and applying the standard analysis procedure on fsQCA. Frequency and consistency cut-off points were specified in this step. Here, the minimum acceptable frequency of cases for solutions was set at 1 and the lowest acceptable consistency cut-off at 0.75, which meets the recommended minimum threshold of 0.75 (Ragin, 2008a). This process clarifies any relationships between combinations of potentially causal or descriptive characteristics and the outcome of interest. The output of a fuzzy-set truth table analysis consists of one or more combinations of characteristics associated with an outcome. We present the results in the next section.

## Results of fsQCA analysis

This section presents the configurations that resulted from the fsQCA analysis of low average excess readmission ratio and high total performance score, shown in Tables 5 and 6, respectively. The configurations are expressed using the notation system established by Ragin and Fiss (2008). The data in Tables 5 and 6 reveal that all of the consistency scores for configurations are above the suggested cut-off value of 0.75 (Legewie, 2013), which suggests that the models based on these configurations are adequately specified. Fuzzy-set qualitative comparative analysis also yields an overall solution coverage and solution consistency. Overall solution coverage measures the proportion of memberships in the outcome that is explained by the complete solution, while overall solution consistency roughly corresponds to the degree to which these configurations consistently result in high quality. This means that the five solutions listed in Table 4 consistently explain 83.2% of the low average excess readmission ratio, while the four solutions listed in Table 6 consistently explain 75.7% of high TPS. Overall solution coverage indicates the extent to which these configurations cover high quality of care (Ragin, 2008a). In a fuzzy-set relationship, this states the percentage of the membership of the outcome set that can be captured by the configurations. Here, the complete solution can capture 56.9% of the low average excess readmission ratio and 55.4% of high TPS.

Table 5 shows that, among the five solutions considered, Solution O1S1 has the highest unique coverage score (0.159), indicating that acquiring analytical and data interpretation capabilities from BDA systems with the support of three other BDA capabilities and improvisational capabilities will

*Table 5. Configurations for low average excess readmission ratio (n = 34 cases)*

| | Solution | | | | |
|---|---|---|---|---|---|
| | O1S1 | O1S2 | O1S3 | O1S4 | O1S5 |
| *BDA capabilities* | | | | | |
| Data integration capability | • | ⊗ | • | • | • |
| Analytical capability | ⬤ | ⬤ | ⬤ | ⬤ | ⬤ |
| Data interpretation capability | ⬤ | ⬤ | ⬤ | ⬤ | ⬤ |
| Predictive capability | • | • | • | • | • |
| Analytics personnel's technical skills | • | ⊗ | • | • | • |
| Analytics personnel's business skills | • | ⊗ | ⊗ | ⊗ | • |
| *Complementary organizational resources* | | | | | |
| Evidence-based decision-making culture | | ⊗ | ⊗ | ⊗ | • |
| Data governance | ⊗ | ⊗ | ⊗ | • | • |
| *Organizational capabilities* | | | | | |
| Dynamic capability | ⊗ | ⊗ | • | ⊗ | • |
| Improvisational capability | • | ⊗ | ⊗ | ⊗ | • |
| Consistency | 0.803 | 0.967 | 0.827 | 0.897 | 0.921 |
| Raw coverage | 0.387 | 0.153 | 0.212 | 0.225 | 0.241 |
| Unique coverage | 0.159 | 0.036 | 0.022 | 0.032 | 0.053 |
| Overall solution consistency | 0.832 | | | | |
| Overall solution coverage | 0.569 | | | | |

*Note:* Black circles (•) indicate the presence of a causal condition, and circles with 'x' (⊗) indicate absence of a causal condition; big circles = core conditions; small circles = peripheral conditions; Blank spaces indicate 'don't care'.

*Table 6. Configurations for high TPS (n = 29 cases)*

| | Solution | | | |
|---|---|---|---|---|
| | O2S1 | O2S2 | O2S3 | O2S4 |
| *BDA capabilities* | | | | |
| Data integration capability | • | ⊗ | • | • |
| Analytical capability | • | ⊗ | • | • |
| Data interpretation capability | ⬤ | ⬤ | ⬤ | ⬤ |
| Predictive capability | • | • | • | ⬤ |
| Analytics personnel's technical skills | • | ⊗ | • | • |
| Analytics personnel's business skills | ⬤ | ⊗ | ⊗ | • |
| *Complementary organizational resources* | | | | |
| Evidence-based decision-making culture | | ⊗ | ⊗ | • |
| Data governance | ⊗ | ⊗ | ⊗ | • |
| *Organizational capabilities* | | | | |
| Dynamic capability | ⊗ | ⬤ | • | • |
| Improvisational capability | • | ⊗ | ⊗ | • |
| Consistency | 0.779 | 0.922 | 0.724 | 0.919 |
| Raw coverage | 0.421 | 0.164 | 0.209 | 0.269 |
| Unique coverage | 0.195 | 0.041 | 0.015 | 0.059 |
| Overall solution consistency | 0.757 | | | |
| Overall solution coverage | 0.554 | | | |

enable healthcare organizations to reduce their average excess readmission ratio in terms of their clinical processes. The necessary condition analysis for this outcome revealed that analytical capability and data interpretation capability are necessary conditions, with consistency scores of 0.901 and 0.979, respectively. This implies that, for a healthcare organization to have a low readmission rate, they almost always have high analytical capability and high data interpretation capability. Five different configurations resulted in low AERRs excess readmission ratios, meaning that five different paths could lead to this outcome. All these solutions shared the same two necessary conditions (i.e. data analytics and data interpretation), but these were accompanied by various other combinations of elements. All the four BDA capabilities were either core factors or contributors in all solutions except for data integration capability, which was absent in Solution O1S1. The two complementary organizational resources (evidence-based decision-making culture and data governance) only contributed to Solutions O1S4 and O1S5. Solution O1S5 appears especially hard to achieve, because it contains all the causal elements; however, it covers 5% of our cases uniquely, which means that there are healthcare organizations in the US that are capable of achieving a high level of quality of care by building all their BDA capabilities, including complementary organizational resources, dynamic and improvisational capabilities.

Table 6 shows four potential paths that healthcare organizations could follow to achieve high total performance. The main similarity among these four configurations is the presence of a high level of data interpretation capability from the BDA systems, a necessary condition for this outcome, with a score of 0.9. Solution O2S1 uniquely explains 19.5% of the variances of high TPS, indicating that a high total performance can be achieved by a high level of data interpretation capability and the cultivation of the analytics personnel's business skills. However, Solution O2S2 differs considerably from Solution O2S1 in that most elements are absent. Here, only two core elements and one support element are needed for a healthcare organization to achieve a high TPS, namely high levels of data interpretation capability from BDA systems and dynamic capabilities, with a supportive role for predictive capability from the BDA system. Solution O2S3 shows that a healthcare organization with high levels of BDA system capabilities, analytics personnel with technical skills and dynamic capabilities for operation can still achieve a high TPS even without high levels of analytics personnel with business skills, an evidence-based decision-making culture, good data governance or improvisational capabilities. Solution O2S4 is identical to Solution O1S5. In this configuration, a healthcare organization has high levels of all the elements considered in this study. It represents an ideal situation that is far from easy to achieve, as evidenced by its unique coverage of 5.9%; only two organizations achieved this high level of TPS, largely due to their high levels of BDA system data interpretation capability and predictive capability, supported by the BDA system's data integration capability and predictive capability, high levels of both types of analytics personnel's skills (technical and business), and high levels of dynamic and improvisational capabilities.

## Discussion

Our findings reveal that most solutions achieving a high level of quality of care have a high level of analytical and data interpretation capabilities combined with data integration capability, predictive capability and analytics personnel's technical skills. In other words, when a healthcare organization lacks a high level of organizational capabilities (dynamic and improvisational capabilities) and organizational resources, the combination of BDA capabilities can still give it a low readmission rate. This finding reaffirms the results of studies by Akter *et al.* (2016) in which the use of BDA can directly improve firm performance. As previously noted, data analytical and interpretation represent the most important components of BDA system for healthcare organizations that encompass the abilities to analyse large amounts of clinical data to understand the past and current state for specific target variables and to generate clinical summary in real time or near real time for role-based decision-making. Indeed, our results agree with those reported by Wang *et al.* (2018b), who indicated that these two BDA capabilities play vital roles in improving the meaningful use of EHR practices and the efficiency of evidence-based medicine practices and meaning, which in turn facilitates quality of care in healthcare.

Surprisingly, the evidence-based decision-making culture is not present in most of the solutions. Unlike the findings reported by previous studies (e.g. Popovič *et al.*, 2012; Ross, Beath and Quaadgras, 2013), our fsQCA result shows that evidence-based decision-making culture is absent in most of the solutions considered, being included only in Solution O1S5 and O2S4. A possible explanation for this result is that, in a healthcare organization such as a clinic, physicians treating patients tend to rely on their professional experience in making decisions, rather than on a system output that they may not be familiar with or been trained to use (Watson, 2014).

It is worth noting that the importance of dynamic and improvisational capabilities is highlighted in some solutions, particularly in developing dynamic capability to improve patient satisfaction (O2S2, O2S3 and O2S4). This discovery confirms the findings of several studies (e.g. Côrte-Real, Oliveira and Ruivo, 2017; Wamba *et al.*, 2017; Wang and Byrd, 2017), which report that dynamic capability plays a key role in leading BDA success. Although organizational capabilities have been shown to be significant predictors of business value creation in a number of different contexts (Pavlou and El Sawy, 2006, 2010), these organizational capabilities are either hard to build or require more long-term planning, so a short-term effect is hard to uncover.

Last but most importantly, one specific capability that is facilitated by BDA systems, data interpretation capability, is the common core causal element of both the desired outcomes considered here. As previously noted, data interpretation capability can generate meaningful clinical summaries in real time or near real time and present them in an easily interpreted format using visual dashboards/systems to yield sharable information and knowledge, such as historical reports, executive summaries, drill-down queries, statistical analyses and time series comparisons to different decision-makers. As suggested by Wang and Byrd (2017), the ready availability of this information assists healthcare analysts to recognize emerging healthcare issues, such as medical errors, potential patient safety issues or inappropriate medication use, enabling them to alert medical professionals and patients so that prompt remedial action can be taken. As incorrect interpretation of the clinical reports generated could lead to serious errors of judgment and questionable decisions, it is important for healthcare organizations to develop interpretation by providing analytical training courses to those employees who will play a critical support role in the new information-rich work environment in the earlier stages of BDA adoption.

### *Theoretical implications*

This study is a preliminary attempt to apply configuration logic and the fsQCA approach to understanding how healthcare performance can be triggered by BDA. Our findings extend current understanding about big data in terms of 4Vs (volume, variety, velocity and veracity) and contribute to the management literature in three ways. First, this study represents a response to an important question raised by Schryen (2013): how do IT resources, IT capabilities and organizational capabilities jointly create business value? Extending beyond traditional interpretations by the RBT, researchers have stressed the particular interrelationships between IT-related elements and organizational elements (organizational resources and capabilities) in the IT business value generation process (Kohli and Grover, 2008; Mikalef and Pateli, 2017; Nevo and Wade, 2010). While the existing literature on big data predominantly suggests that each BDA element solely leads to organizational performance, this study, which applies fsQCA, provides further evidence to support their view by confirming that BDA implementation does indeed depend on the joint effects of BDA capability, complementary organizational resources and organizational capabilities.

Second, a major debate in the field of management concerning the value of dynamic capability for firm performance has been going on (Easterby-Smith, Lyles and Peteraf, 2009) since Eisenhardt and Martin (2000) asserted that the value for competitive advantage lies in the resource configurations that they create, not in the capabilities themselves. Kohli and Grover (2008) respond to this debate by arguing that firm performance may be the result of particular combinations of input elements, such as IT resources and organizational resources. El Sawy *et al.* (2010) support this view by confirming that a strategic advantage can be built by the holistic confluence among environmental turbulence, dynamic capabilities and IT systems. We clarify this debate by providing further evidence showing that healthcare performance can be improved in hospitals with dynamic capabilities, in conjunction with the support from the effective use of their analytical and data interpretation capabilities as well as organizational resources such as data governance. This implies that dynamic capabilities cannot by themselves be a source of a sustainable competitive advantage; rather, it should be developed though the synergistic effect of BDA capabilities and other organizational resources.

Finally, organizational capabilities such as dynamic capability and improvisational capability typically play an enabler or a mediator role in linking IT to business value (Pavlou and El Sawy, 2006, 2010; Wu and Hu, 2012). Extending the theoretical perspective from a strategic alignment between IT and business to co-evolution, previous studies have suggested that the key to successful implementation of health information technologies (HIT) is to carefully orchestrate the complex and dynamic interactions between organizational capabilities and HIT throughout the business process (Agarwal *et al.*, 2010; Goh, Gao and Agarwal, 2011; Novak *et al.*, 2012). Although these studies have mentioned the systemic notion of co-evolution among individual elements for information systems success, examining the effect of co-evolution with conventional correlation-based linear methods (e.g. two-way correlations, testing moderator/mediator effect) does not support the holistic view required to capture the non-linear interdependent interactions among these elements. To the best of our knowledge, this is among the first study that examines the complex interactions among BDA and the organizational capabilities driving healthcare performance.

*Managerial implications*

From a practical perspective, our study advances an understanding of the 'black box' between BDA and firm performance by exploring the complex causality among BDA capabilities, complementary organizational resources and organizational capabilities. Our findings not only reveal the synergy effect of BDA capabilities and BDA human resources (which here refers to the technical skills of the organization's analytics personnel) in achieving improved readmission rates and patient satisfaction, but also show that BDA cannot achieve this in isolation from other elements, as organizational resources and capabilities play a supportive role. These fsQCA results provide the 'secret recipes' needed to achieve healthcare performance by considering the presence or absence of the various 'ingredients'. These secret recipes could be the useful solutions for healthcare practitioners, leveraging BDA to improve healthcare performance. By comparing the similarities and differences between multiple equifinal configurations, we extract patterns that produce the desired level of healthcare quality in terms of improved readmission rates and patient satisfaction. Based on the patterns identified, healthcare organization managers can adopt solutions specifically tailored to their own characteristics or situations to achieve high healthcare quality and avoid the expensive pitfalls of misplaced BDA investments.

In practice, most organizations continue to struggle to make progress on their BDA initiatives because implementing a BDA system can be an expensive and risky undertaking (Watson, 2014). It typically costs a big data project approximately $9.3m to build and maintain a Hadoop system over a five-year period (Winter, Gilbert and Davis, 2013). Our fsQCA results offer a limited set of useful configurational solutions to achieve high quality of care so that it enables healthcare organizations to develop a clear path to BDA success.

## Limitations and future research

While we believe that the fsQCA method can contribute to our research, this method suffers from a number of limitations. First and foremost, fsQCA depends on prior knowledge or an extensive literature on the subject to select appropriate conditions and outcomes and reduce the number of configurations to a manageable level (Liu *et al.*, 2017). The configurations are sensitive to the range of conditions included – adding or removing conditions could result in very different solutions. Although the selection of the conditions in our analysis was built on the business value of IT generation framework provided by Melville, Kraemer and Gurbaxani (2004) and was informed by a comprehensive review of the extant literature on BDA, the conditions we chose came mainly from exploratory studies or case studies, with little support from empirical evidence. As a result, one or more care quality drivers could have been overlooked or overestimated. To address this concern, a more rigorous study should be conducted to identify what constitutes stable conditions, for example, by incorporating a mixed-method research design, such as a qualitative Delphi approach, and content analysis to provide a stronger basis for condition selection.

Second, there are limitations and disadvantages related to our data set. A major limitation is the small sample size for our matchup data set. Although fsQCA is sensitive to case selection (Liu *et al.*, 2017), it does allow for the analysis of small to medium numbers of cases (e.g. 10 to 50) that traditional regression-based methods may not be able to solve (Ragin, 2008b). We also sought to address real-world issues by using actual measures, such as the average excess readmission ratio from the CMS database, for assessing healthcare performance rather than scaled self-reporting performance. Adopting this approach enabled us to interpret more accurately the implications of each configuration. As fsQCA requires larger samples (50+ cases) to reduce contradictions, healthcare researchers may want to consider a second analysis using fsQCA data from the HIMSS and CMS databases, which contain a large number of cases that could be used to develop robust IT business value explanatory models.

## Conclusion

In summary, rather than simply exploring the direct effect of BDA on healthcare performance through a traditional linear causal analysis – as tends to be the case in existing big data research – we have focused on examining the systemic, equifinal and discontinuous interactions among BDA elements and other organizational elements. Applying configuration theory and fsQCA in this study has allowed us to discover not only single drivers, but also sets of conditions that determine the quality of care triggered by BDA in healthcare. These findings from fsQCA advance our understanding of how BDA-enabled IT capabilities combine with other organizational elements to achieve business value in healthcare. Most importantly, we offer evidence that different solutions leading to the same healthcare performance due to the effective use of IT, and other organizational elements do indeed exist. This demonstrates that fsQCA is a useful and appropriate tool for assessing the business value of BDA that can offer new insights to improve understanding of the factors contributing to the business value of BDA. As the use of fsQCA is still in its infancy in most business domains, more substantive discussions of the possibilities opened up by this new approach are needed if we are to reap the full benefit of applying fsQCA to investigations of the impact of new technology on firm performance.

## Appendix A: Skill sets for analytical personnel

| Studies | Analytical personnel's technical skills | Analytical personnel's business skills |
|---|---|---|
| Chiang and Stohr (2012) | Analytical skills (e.g. data-mining, deviational analysis and anomaly detection, geospatial and temporal analysis)<br>IT skills (e.g. relational databases, data warehouse, Hadoop, MapReduce, unstructured data management) | Business knowledge and communication skills |
| Wixom *et al.* (2014) | SQL and Query skills<br>Basic analytics<br>Data management<br>Data integration<br>Reporting (OLAP) skills<br>Research methods<br>Visualization<br>Advanced analytics<br>Data and text mining<br>Programming<br>No SQL skills | Communication skills<br>Business requirement<br>Business knowledge<br>Emerging topics |
| Mamonov, Misra and Jain (2014) | Applied statistics<br>Technical skills<br>Analytical software | Soft skills (e.g. communication and presentation, teamwork) |
| Wilder and Ozgur (2015) | Solid foundation in computer science and mathematics<br>Understand how data is managed | Identify and exploit business opportunities, frame business problems and interpret the results |
| Cegielski and Jones-Farmer (2016) | Technical skills (e.g. ability to integrate analyses from multiple sources into a business solution, ability to use data visualization/graphical tools to interpret data, and ability to frame a business problem or question analytically) | Business skills (e.g. independent learner, organizational skills, industry-specific knowledge) |

## Appendix B: The definitions of key constructs used in the current study

| Constructs | Definitions |
|---|---|
| BDA capability | The ability to acquire, store, process and analyse 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 (Wang and Hajli, 2017). |
| Data integration capability | The ability to transform different types of data into a data format that can be read by the data analysis platform (Wang and Byrd, 2017). |
| Analytical capability | The ability to drive decisions and actions through the extensive use of data and different analytical techniques based on the specific mechanisms used for analytics, thus addressing the various needs of users and other stakeholders (Ghosh and Scott, 2011). |
| Predictive capability | The process of using a set of sophisticated tools to develop models and estimations of what the environment will do in the future (Wessler, 2013, p. 21). |
| Data interpretation capability | The ability to produce a healthcare matrix and reports that evaluate patient care and service and identify areas for improvement (defined by current study). |
| Analytical personnel | The members of an organization who have an analytical mindset and help derive value from BDA. |
| BDA-enabled complementary organizational resources | A specific type of organizational resource with the aid of BDA that tend to be tacit, idiosyncratic and deeply embedded in the organization. |
| Evidence-based decision-making culture | An organizational culture of embracing evidence-based management and embedding evidence-based decision-making in the core values and processes of the organization (Davenport, Harris and Morison, 2010). |
| Data governance | Built on IT governance, aims to formulate data rules and policies and provide a vision and guidelines relating to privacy, security, life cycle and ownership of data by aligning the objectives of multiple functions (Kooper, Maes and Lindgreen, 2011; LaValle *et al.*, 2011). |
| Organizational capability | The ability to adapt to ongoing changes in the business processes and functional activities of the firm (Luo, Fan and Zhang, 2012). |
| Dynamic capability | The ability to sense and shape opportunities and threats, seize market opportunities and maintain competitiveness (Barreto, 2010; Teece, 2007). |
| Improvisational capability | An organization's learned ability to respond to unexpected environmental turbulence quickly, simultaneously forming and executing novel solutions by reconfiguring available resources (Pavlou and El Sawy, 2010). |

## Appendix C: Measurement and items

| *BDA capabilities* | |
|---|---|
| Data integration capability (Wang and Byrd, 2017) | Integrate seamlessly clinical data across multiple departments in real time or near real time |
| | Track medical events based on the rules that have been built on hospital claims |
| | Search clinical databases for all data related to patients |
| Analytical capability (Wang and Byrd, 2017) | Analyse large amounts of clinical data to understand the past and current state for specific target variables |
| | Explore the causes of medical events from clinical data |
| | Support real-time processing of multiple clinical data streams |
| Predictive capability (Wang *et al.*, 2017) | Discover patterns among specific variables of interest across departments |
| | Analyse data from different sources and use the results to predict future trends |
| | Provide actionable insights from clinical data in a format readily understood by healthcare providers |
| Data interpretation capability (Wang and Byrd, 2017) | Generate clinical summary in real time or near real time and present in visual dashboards |
| | Provide outputs for role-based decision-making |
| *Analytics personnel skills* | |
| Technical skills (Cegielski and Jones-Farmer, 2016) | Ability to integrate analyses from multiple sources into a business solution |
| | Ability to use data visualization/graphical tools to interpret data |
| | Ability to frame a business problem or question analytically |
| | Ability to solve pre-framed business problems or questions analytically |

(*Continued*)

| *BDA capabilities* | |
|---|---|
| Business skills (Cegielski and Jones-Farmer, 2016) | Ability to be an independent learner<br>Organizational skills<br>Healthcare knowledge |
| *Complementary organizational resources* | |
| Evidence-based decision-making culture (Popovič *et al.*, 2012) | Our hospital usually uses evidence-based insights for the creation of new healthcare service<br>Our hospital is open to new ideas and approaches that challenge current or future projects on the basis of new insights<br>Our hospital allows the incorporation of available information within any decision-making process. |
| Data governance (Khatri and Brown, 2010) | Data principle (clarifying the role of data as an asset)<br>Data quality (establishing the requirements of intended use of data)<br>Metadata (establishing the semantics of data so that it is interpretable by the users)<br>Data access (specifying access requirements of data)<br>Data life cycle (determining the definition, production, retention and retirement of data) |
| *Organizational capabilities* | |
| Planned dynamic capabilities (Pavlou and El Sawy, 2010) | Our hospital frequently generates, disseminates and responds to market intelligence about customer needs<br>Our hospital has adequate routines to acquire, assimilate, transform and exploit existing resources to generate new knowledge<br>Our hospital is effective in managing dependencies among resources and tasks to synchronize activities<br>Our hospital effectively integrates disparate employees' inputs through heedful contribution, representation, and interrelation into our group |
| Improvisational capabilities (Pavlou and El Sawy, 2010) | Our hospital is successful in figuring out our actions as we go along<br>Our hospital effectively improvises when carrying out our activities<br>Our hospital could spontaneously readjust our activities according to competitive environments |

*Measurement for quality of care*

As Agarwal *et al.*'s (2010) health information technology impact framework suggests, we operationalize healthcare performance by using quality of care. Quality of care is a key component of the business value expected from healthcare information technologies (Bardhan and Thouin, 2013). To assess the quality of care, we took advantage of the recently released Hospital Compare Data database to gather data from the Hospital Readmissions Reduction Program (HRRP) and the Hospital Value-Based Purchasing (HVBP) Program. A hospital's excess readmission ratio is a measure of that hospital's readmission performance compared to the national average for a comparable set of patients with the same conditions. While there are a variety of quality outcome measures that could be considered, we chose excess readmission ratio, as this most accurately reflects the total process of care received. Hospitals can provide a better quality of care if the risk of being readmitted for the same diagnosis in the future is reduced (Bardhan, Oh, Zheng and Kirksey, 2015). The average excess readmission ratio (AERR) was calculated using the following formulae; the higher the ratio the worse the quality of care.

(1) Excess readmission ratio = risk-adjusted predicted readmissions/risk-adjusted expected readmissions.
(2) Average excess readmission ratio = (Excess Readmission Ratio for Pneumonia + Excess Readmission Ratio for heart failure + Excess Readmission Ratio for acute myocardial infarction + Excess Readmission Ratio for total hip/knee arthroplasty + Excess Readmission Ratio for Chronic Obstructive Pulmonary Disease)/5

Another way to measure the quality of care is in terms of the patient satisfaction data provided by the Hospital Value-Based Purchasing (HVBP) programme from CMS. This programme is part of CMS' long-standing effort to link Medicare's payment system to quality. The programme implements value-based purchasing for the payment system that accounts for the largest share of Medicare spending. Hospitals are paid for inpatient acute care services based on the quality of care, not just the quantity of the services they provide. From this data, two domains can be used to assess hospital performance: (1) Patient experience of care and (2) Clinical process of care. The patient experience of care domain is comprised of the Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) Survey measures. The Clinical Process of Care domain is comprised of selected Inpatient Quality Reporting (IQR) Programme's Process of Care measures from the Acute Myocardial Infarction (AMI), Healthcare Associated Infections (HAI), Heart Failure (HF), Pneumonia (PN), and Surgical Care Improvement Project (SCIP) measure sets. A performance score and an improvement score are calculated for each measure, after which a domain score is calculated for each of the two domains. The Total Performance Score (TPS) is calculated using the weighted domain scores. The Clinical Process of Care domain score is weighted as 70% of the TPS, and the Patient Experience of Care domain is weighted as 30% of the TPS.

## Appendix D: Item loadings and cross loadings

| | DIA | ANA | INT | PRE | TS | BS | CUL | DG | DYN | IM |
|---|---|---|---|---|---|---|---|---|---|---|
| DIA1 | 0.801 | −0.096 | 0.091 | 0.020 | 0.167 | −0.118 | 0.057 | 0.033 | 0.273 | 0.027 |
| DIA2 | 0.782 | 0.180 | 0.173 | −0.088 | 0.268 | 0.072 | 0.089 | −0.070 | 0.193 | −0.104 |
| DIA3 | 0.886 | 0.028 | −0.010 | 0.097 | 0.124 | 0.060 | 0.086 | −0.064 | 0.036 | −0.009 |
| ANA1 | 0.077 | 0.873 | 0.103 | 0.094 | −0.004 | −0.047 | 0.101 | −0.078 | −0.046 | 0.258 |
| ANA2 | 0.021 | 0.913 | −0.024 | 0.118 | −0.089 | 0.080 | 0.032 | −0.107 | −0.038 | 0.121 |
| ANA3 | −0.023 | 0.802 | 0.135 | 0.094 | −0.245 | −0.053 | 0.034 | −0.025 | 0.105 | 0.202 |
| INT1 | 0.045 | 0.092 | 0.907 | 0.074 | −0.113 | 0.069 | 0.047 | −0.152 | 0.164 | −0.004 |
| INT2 | 0.181 | 0.085 | 0.843 | 0.168 | −0.205 | 0.100 | 0.073 | −0.129 | −0.041 | 0.170 |
| PRE1 | 0.098 | 0.072 | 0.150 | 0.867 | 0.099 | 0.090 | 0.093 | −0.090 | −0.005 | 0.067 |
| PRE2 | −0.020 | 0.184 | −0.004 | 0.832 | 0.269 | −0.019 | −0.009 | 0.034 | 0.014 | 0.059 |
| PRE3 | −0.017 | 0.048 | 0.073 | 0.836 | 0.144 | 0.196 | −0.072 | −0.048 | 0.005 | 0.053 |
| TS1 | 0.065 | −0.082 | −0.123 | 0.192 | 0.833 | 0.094 | −0.057 | −0.010 | 0.179 | 0.075 |
| TS2 | 0.127 | −0.100 | −0.035 | 0.060 | 0.865 | 0.089 | −0.037 | 0.090 | 0.162 | −0.008 |
| TS3 | 0.188 | −0.131 | −0.132 | 0.201 | 0.792 | 0.262 | −0.016 | −0.026 | 0.112 | −0.016 |
| TS4 | 0.208 | −0.054 | −0.051 | 0.149 | 0.804 | −0.011 | −0.061 | .081 | 0.091 | −0.229 |
| BS1 | 0.013 | 0.101 | 0.266 | 0.060 | 0.156 | 0.731 | 0.271 | 0.058 | −0.142 | −0.018 |
| BS2 | 0.015 | −0.102 | −0.042 | −0.008 | 0.133 | 0.853 | −0.069 | −0.016 | 0.027 | 0.056 |
| BS3 | −0.015 | 0.031 | 0.032 | 0.246 | 0.063 | 0.828 | 0.069 | 0.155 | −0.060 | 0.022 |
| CUL1 | 0.128 | −0.006 | −0.015 | −0.151 | −0.074 | −0.040 | 0.922 | 0.003 | −0.063 | 0.030 |
| CUL2 | 0.054 | 0.154 | 0.084 | 0.057 | −0.045 | 0.122 | 0.833 | −0.156 | −0.096 | 0.077 |
| CUL3 | 0.038 | 0.014 | 0.060 | 0.130 | −0.029 | 0.111 | 0.778 | −0.292 | −0.124 | 0.206 |
| DG1 | −0.157 | −0.256 | −0.115 | −0.028 | 0.172 | 0.040 | −0.012 | 0.812 | 0.087 | 0.052 |
| DG2 | 0.031 | −0.267 | −0.281 | −0.026 | 0.056 | 0.030 | 0.002 | 0.765 | −0.003 | −0.069 |
| DG3 | −0.062 | 0.039 | −0.014 | 0.077 | −0.039 | 0.037 | −0.206 | 0.804 | 0.044 | 0.059 |
| DG4 | 0.034 | 0.069 | 0.028 | −0.124 | 0.086 | 0.110 | −0.014 | 0.650 | −0.228 | 0.041 |
| DG5 | 0.029 | 0.053 | −0.009 | −0.019 | −0.104 | −0.025 | −0.155 | 0.806 | −0.009 | −0.180 |
| DYN1 | 0.091 | −0.011 | −0.093 | −0.054 | 0.161 | −0.034 | −0.282 | 0.032 | 0.768 | −0.114 |
| DYN2 | 0.189 | 0.000 | 0.023 | 0.034 | 0.073 | −0.070 | −0.041 | −0.058 | 0.883 | 0.054 |
| DYN3 | 0.128 | −0.090 | 0.035 | 0.084 | 0.110 | −0.089 | −0.046 | 0.008 | 0.915 | 0.067 |
| DYN4 | 0.042 | 0.189 | 0.274 | −0.100 | 0.282 | 0.093 | 0.043 | −0.118 | 0.715 | −0.211 |
| IM1 | 0.020 | 0.245 | 0.228 | 0.072 | −0.067 | 0.024 | 0.174 | −0.001 | −0.077 | 0.832 |
| IM2 | −0.015 | 0.152 | −0.008 | −0.074 | 0.055 | 0.046 | 0.046 | −0.079 | 0.045 | 0.898 |
| IM3 | −0.070 | 0.174 | −0.039 | 0.217 | −0.142 | 0.004 | 0.079 | 0.011 | −0.078 | 0.827 |

*Note:* DIA = data integration capability; ANA = analytical capability; INT = data interpretation capability; PRE = predictive capability; TS = personnel's technical skills; BS = personnel's business skills; CUL = evidence-based decision-making culture; DG = data governance; DYN = planned dynamic capabilities; IM = improvisational capabilities. Bold numbers indicate item loadings on the assigned constructs.

## References


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

Akter, S., S. F. Wamba, A. Gunasekaran, R. Dubey and S. J. Childe (2016). 'How to improve firm performance using big data analytics capability and business strategy alignment?', *International Journal of Production Economics*, **182**, pp. 113–131.

Armstrong, J. S. and T. S. Overton (1977). 'Estimating nonresponse bias in mail surveys', *Journal of Marketing Research*, **14**, pp. 396–402.

Bagozzi, R. P., Y. Yi and L. W. Phillips (1991). 'Assessing construct validity in organisational research', *Administrative Science Quarterly*, **36**, pp. 421–458.

Bardhan, I. R. and M. F. Thouin (2013). 'Health information technology and its impact on the quality and cost of healthcare delivery', *Decision Support Systems*, **55**, pp. 438–449.

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

Barreto, I. (2010). 'Dynamic capabilities: a review of past research and an agenda for the future', *Journal of Management*, **36**, pp. 256–280.

Bergh, D. D. and E. N. K. Lim (2008). 'Learning how to restructure: absorptive capacity and improvisational views of restructuring actions and performance', *Strategic Management Journal*, **29**, pp. 593–61

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

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

Cegielski, C. G. and L. A. Jones-Farmer (2016). 'Knowledge, skills, and abilities for entry-level business analytics positions: a multi-method study', *Decision Sciences Journal of Innovative Education*, **14**, pp. 91–118.

Chiang, R. H., P. Goes and E. A. Stohr (2012). 'Business intelligence and analytics education, and program development: a unique opportunity for the information systems discipline', *ACM Transactions on Management Information Systems (TMIS)*, **3**, Article 12.

Chin, W. W. (1998). The partial least squares approach to structural equation modeling. In G. A. Marcoulides (ed.), *Modern Methods for Business Research*, pp. 295–336. Mahwah, NJ: Lawrence Erlbaum Associates.

Côrte-Real, N., T. Oliveira and P. Ruivo (2017). 'Assessing business value of big data analytics in European firms', *Journal of Business Research*, **70**, pp. 379–390

Davenport, T. H. (2006). 'Competing on analytics', *Harvard Business Review*, **84**, pp. 98–107.

Davenport, T. H. and J. G. Harris. (2007). *Competing on Analytics: The New Science of Winning*. Boston, MA: Harvard Business School Press.

Davenport, T. H. and D. J. Patil (2012). 'Data scientist: the sexiest job of the 21st century', *Harvard Business Review*, https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century.

Davenport, T. H., J. G. Harris and R. Morison (2010). *Analytics at Work: Smarter Decisions, Better Results*. Boston, MA: Harvard Business School Press.

Demchenko, Y., P. Grosso, C. D. Laat and P. Membrey (2013). *Addressing Big Data Issues in Scientific Data Infrastructure*. San Diego, CA: IEEE.

Easterby-Smith, M., M. A. Lyles and M. A. Peteraf (2009). 'Dynamic capabilities: current debates and future directions', *British Journal of Management*, **20**, pp. S1–S8.

Eisenhardt, K. M. and J. A. Martin (2000). 'Dynamic capabilities: what are they?', *Strategic Management Journal*, **21**, pp. 1105–1121.

El Sawy, O. A. and P. A. Pavlou (2008). 'IT-enabled business capabilities for turbulent environments', *MIS Quarterly Executive*, **7**, pp. 139–150.

El Sawy, O. A., A. Malhotra, Y. Park and P. A. Pavlou (2010). 'Seeking the configurations of digital ecodynamics: it takes three to tango', *Information Systems Research*, **21**, pp. 835–848.

Fichman, R. G. (2004). 'Going beyond the dominant paradigm for information technology innovation research: emerging concepts and methods', *Journal of the Association for Information Systems*, **5**, pp. 314–355.

Fink, L., N. Yogev and A. Even (2017). 'Business intelligence and organisational learning: an empirical investigation of value creation processes', *Information & Management*, **54**, pp. 38–56.

Fiss, P. C. (2007). 'A set-theoretic approach to organisational configurations', *Academy of Management Review*, **32**, pp. 1180–1198.

Fiss, P. C. (2011). 'Building better causal theories: a fuzzy set approach to typologies in organisation research', *Academy of Management Journal*, **54**, pp. 393–420.

Fiss, P. C., B. Cambré and A. Marx (2013). *Configurational Theory and Methods in Organisational Research*, Vol. **38**. Bingley: Emerald.

Fornell, C. and D. F. Larcker (1981). 'Evaluating structural equation models with unobservable variables and measurement error', *Journal of Marketing Research*, **18**, pp. 39–50.

Galbraith, J. R. (1974). 'Organisation design: an information processing view', *Interfaces*, **4**, pp. 28–36.

Gandomi, A. and M. Haider (2015). 'Beyond the hype: big data concepts, methods, and analytics', *International Journal of Information Management*, **35**, pp. 137–144.

Gefen, D., D. W. Straub and M. C. Boudreau (2000). 'Structural equation modeling techniques and regression: guidelines for research practice', *Communications of the Association for Information Systems*, **4**, pp. 1–78.

Ghosh, B. and J. E. Scott (2011). 'Antecedents and catalysts for developing a healthcare analytic capability', *Communications of the Association for Information Systems*, **29**, pp. 395–409.

Goh, J. M., G. Gao and R. Agarwal (2011). 'Evolving work routines: adaptive routinization of information technology in healthcare', *Information Systems Research*, **22**, pp. 565–585.

Grant, R. M. (1996). 'Toward a knowledge-based theory of the firm', *Strategic Management Journal*, **17**, pp. 109–122.

Gupta, M. and J. F. George (2016). 'Toward the development of a big data analytics capability', *Information & Management*, **53**, pp. 1049–1064.

Hadida, A. L. and W. Tarvainen (2014). 'Organisational improvisation: a consolidating review and framework', *International Journal of Management Reviews*, **17**, pp. 437–459.

Hus, B. S. and E. Griese (2018). 'Making better use of health care data', *Harvard Business Review*, https://hbr.org/2018/03/making-better-use-of-health-care-data.

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

Kallinikos, J. and N. Tempini (2014). 'Patient data as medical facts: social media practices as a foundation for medical knowledge creation', *Information Systems Research*, **25**, pp. 817–833.

Kettinger, W., C. Zhang and D. A. Marchand (2011). 'CIO and business executive leadership approaches to establishing company-wide information orientation', *MIS Quarterly Executive*, **10**, pp. 157–174.

Khatri, V. and C. V. Brown (2010). 'Designing data governance', *Communications of the ACM*, **53**, pp. 148–152.

Kiron, D. and R. Shockley (2011). 'Creating business value with analytics', *MIT Sloan Management Review*, **53**, pp. 57–63.

Kiron, D., R. B. Ferguson and P. K. Prentice (2013). 'From value to vision: reimagining the possible with data analytics', *MIT Sloan Management Review*, **54**, pp. 1–19.

Kiron, D., P. K. Prentice and R. B. Ferguson (2012). 'Innovating with analytics', *MIT Sloan Management Review*, **54**, pp. 47–51.

Kohli, R. and V. Grover (2008). 'Business value of IT: an essay on expanding research directions to keep up with the times', *Journal of the Association for Information Systems*, **9**, pp. 22–39.

Kooper, M. N., R. Maes and E. R. Lindgreen (2011). 'On the governance of information: Introducing a new concept of governance to support the management of information', *International Journal of Information Management*, **31**, pp. 195–200.

Kowalczyk, D. W. I. M. and P. Buxmann (2014). 'Big data and information processing in organisational decision processes', *Business & Information Systems Engineering*, **6**, pp. 267–278.

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

Lee, Y., S. Madnick, R. Wang, F. Wang and H. Zhang (2014). 'A cubic framework for the chief data officer: succeeding in a world of big data', *MIS Quarterly Executive*, **13**, pp. 1–13.

Legewie, N. (2013). 'An introduction to applied data analysis with qualitative comparative analysis', *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*, **14**, Article 15.

Liu, Y., J. Mezei, V. Kostakos and H. Li (2017). 'Applying configurational analysis to IS behavioural research: a methodological alternative for modelling combinatorial complexities', *Information Systems Journal*, **27**, pp. 59–89.

Loshin, D. (2010). *Master Data Management*. Burlington, MA: Morgan Kaufmann.

Luo, J., M. Fan and H. Zhang (2012). 'Information technology and organisational capabilities: a longitudinal study of the apparel industry', *Decision Support Systems*, **53**, pp. 186–194.

Mamonov, S., R. Misra and R. Jain (2015). 'Business analytics in practice and in education: a competency based perspective', *Information Systems Education Journal*, **13**, pp. 4–13.

McAfee, A. and E. Brynjolfsson (2012). 'Big data: the management revolution', *Harvard Business Review*, **90**, pp. 59–68.

Melville, N., K. Kraemer and V. Gurbaxani (2004). 'Review – Information technology and organisational performance: an integrative model of IT business value', *MIS Quarterly*, **28**, pp. 283–322.

Mikalef, P. and A. Pateli (2017). 'Information technology-enabled dynamic capabilities and their indirect effect on competitive performance: findings from PLS-SEM and fsQCA', *Journal of Business Research*, **70**, pp. 1–16.

Misangyi, V. F., T. Greckhamer, S. Furnari, P. C. Fiss, D. Crilly and R. Aguilera (2017). 'Embracing causal complexity: the emergence of a neo-configurational perspective', *Journal of Management*, **43**, pp. 255–282.

Moorman, C. and A. S. Miner (1998). 'Organisational improvisation and organisational memory', *Academy of Management Review*, **23**, pp. 698–723.

Nevo, S. and M. R. Wade (2010). 'The formation and value of IT-enabled resources: antecedents and consequences of synergistic relationships', *MIS Quarterly*, **34**, pp. 163–183.

Novak, L., J. Brooks, C. Gadd, S. Anders and N. Lorenzi (2012). 'Mediating the intersections of organisational routines during the introduction of a health IT system', *European Journal of Information Systems*, **21**, pp. 552–569.

Pavlou, P. A. and O. A. El Sawy (2006). 'From IT leveraging competence to competitive advantage in turbulent environments: the case of new product development', *Information Systems Research*, **17**, pp. 198–227.

Pavlou, P. A. and O. A. El Sawy (2010). 'The "third hand": IT-enabled competitive advantage in turbulence through improvisational capabilities', *Information Systems Research*, **21**, pp. 443–471.

Podsakoff, P. M. and D. W. Organ (1986). 'Self-reports in organisational research: problems and prospects', *Journal of Management*, **12**, pp. 531–544.

Podsakoff, P. M., S. B. MacKenzie, J. Y. Lee and N. P. Podsakoff (2003). 'Common method biases in behavioral research: a critical review of the literature and recommended remedies', *Journal of Applied Psychology*, **88**, pp. 879–903.

Popovič, A., R. Hackney, P. S. Coelho and J. Jaklič (2012). 'Towards business intelligence systems success: effects of maturity and culture on analytical decision making', *Decision Support Systems*, **54**, pp. 729–739.

Popovič, A., R. Hackney, R. Tassabehji and M. Castelli (2018). 'The impact of big data analytics on firms' high value business performance', *Information Systems Frontiers*, **20**, pp. 209–222.

Raghupathi, W. and V. Raghupathi (2014). 'Big data analytics in healthcare: promise and potential', *Health Information Science and Systems*, **2**, pp. 1–10.

Ragin, C. C. (2000). *Fuzzy-Set Social Science*. Chicago, IL: University of Chicago Press.

Ragin, C. C. (2008a). *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago, IL: University of Chicago Press.

Ragin, C. C. (2008b). *User's Guide to Fuzzy-Set/Qualitative Comparative Analysis 2.0*. Tucson, AZ: Department of Sociology, University of Arizona.

Ragin, C. C. and P. C. Fiss (2008). Net effects analysis versus configurational analysis: an empirical demonstration. In C. C. Ragin (ed.), *Redesigning Social Inquiry: Fuzzy Sets and Beyond*, pp. 190–212. Chicago, IL: University of Chicago Press.

Rai, A., R. Patnayakuni and N. Seth (2006). 'Firm performance impacts of digitally enabled supply chain integration capabilities', *MIS Quarterly*, **30**, pp. 225–246.

Ravasi, D. and M. Schultz (2006). 'Responding to organisational identity threats: exploring the role of organisational culture', *Academy of Management Journal*, **49**, pp. 433–458.

Roberts, N. and V. Grover (2012). 'Leveraging information technology infrastructure to facilitate a firm's customer agility and competitive activity: an empirical investigation', *Journal of Management Information Systems*, **28**, pp. 231–270.

Roski, J., G. Bo-Linn and T. Andrews (2014). 'Creating value in health care through big data: opportunities and policy implications', *Health Affairs*, **33**, pp. 1115–1122.

Ross, J. W., C. M. Beath and A. Quaadgras (2013). 'You may not need big data after all', *Harvard Business Review*, **91**, pp. 90–98.

Samorani, M. and L. R. LaGanga (2015). 'Outpatient appointment scheduling given individual day-dependent no-show predictions', *European Journal of Operational Research*, **240**, pp. 245–257.

Schryen, G. (2013). 'Revisiting IS business value research: what we already know, what we still need to know, and how we can get there', *European Journal of Information Systems*, **22**, pp. 139–169.

Seddon, P. B., D. Constantinidis and H. Dod (2012). 'How does business analytics contribute to business value?' *Proceeding of the 33rd International Conference on Information Systems*, Orlando, FL.

Shah, N. D. and J. Pathak (2014). 'Why health care may finally be ready for Big Data'. Available at https://hbr.org/2014/12/why-health-care-may-finally-be-ready-for-big-data. [accessed 12 January 2015].

Sharma, R., S. Mithas and A. Kankanhalli (2014). 'Transforming decision-making processes: a research agenda for understanding the impact of business analytics on organisations', *European Journal of Information Systems*, **23**, pp. 433–441.

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

Srinivasan, R. and M. Swink (2018). 'An investigation of visibility and flexibility as complements to supply chain analytics: an organisational information processing theory perspective', *Production and Operations Management*, **27**, pp. 1849–1867.

Tanriverdi, H. (2006). 'Performance effects of information technology synergies in multibusiness firms', *MIS Quarterly*, **30**, pp. 57–77.

Teece, D. (2007). 'Explicating dynamic capabilities: the nature and microfoundations of (sustainable) enterprise performance', *Strategic Management Journal*, **28**, pp. 1319–1350.

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

Trkman, P., K. McCormack, M. P. V. De Oliveira and M. B. Ladeira (2010). 'The impact of business analytics on supply chain performance', *Decision Support Systems*, **49**, pp. 318–327.

Wamba, S. F., A. Gunasekaran, S. Akter, S. J. F. Ren, R. Dubey and S. J. Childe (2017). 'Big data analytics and firm performance: effects of dynamic capabilities', *Journal of Business Research*, **70**, pp. 356–365.

Wang, Y. and N. Hajli (2017). 'Exploring the path to big data analytics success in healthcare', *Journal of Business Research*, **70**, pp. 287–299.

Wang, Y. and T. A. Byrd (2017). 'Business analytics-enabled decision making effectiveness through knowledge absorptive capacity in health care', *Journal of Knowledge Management*, **21**, pp. 517–539

Wang, Y., L. Kung and T. A. Byrd (2018a). 'Big data analytics: understanding its capabilities and potential benefits for healthcare organisations', *Technological Forecasting and Social Change*, **126**, pp. 3–13

Wang, Y., L. Kung, W. Y. C. Wang and C. G. Cegielski (2018b). 'An integrated big data analytics-enabled transformation model: application to health care', *Information & Management*, **55**, pp. 64–79.

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

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

Wessler, M. (2013). *Big Data Analytics for Dummies*. Hoboken, NJ: John Wiley & Sons, Inc.

Wilder, C. R. and C. O. Ozgur (2015). 'Business analytics curriculum for undergraduate majors', *INFORMS Transactions on Education*, **15**, pp. 180–187.

Winter, R., R. Gilbert and J. R. Davis (2013). *Big Data: What Does It Really Cost*? Cambridge, MA: Winter Corporation.

Wixom, B., T. Ariyachandra, D. E. Douglas, M. Goul, B. Gupta, L. S. Iyer and O. Turetken (2014). 'The current state of business intelligence in academia: the arrival of big data', *Communications of the Association for Information Systems*, **34**, Article 1.

Wixom, B., B. Yen and M. Relich (2013). 'Maximizing value from business analytics', *MIS Quarterly Executive*, **12**, pp. 111–123.

Wu, I. L. and Y. P. Hu (2012). 'Examining knowledge management enabled performance for hospital professionals: a dynamic capability view and the mediating role of process capability', *Journal of the Association for Information Systems*, **13**, pp. 976–999.

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

Zammuto, R. F., T. L. Griffith, A. Majchrzak, D. J. Dougherty and S. Faraj (2007). 'Information technology and the changing fabric of organisation', *Organization Science*, **18**, pp. 749–762.

Yichuan Wang is a Senior Lecturer (Associate Professor) in Digital Marketing at the Sheffield University Management School, University of Sheffield, United Kingdom. He received his PhD in Business & Information Systems from the Raymond J. Harbert College of Business, Auburn University, United States His research interests include big data analytics and social media marketing. His research has been published in journals such as British Journal of Management, Annals of Tourism Research, Industrial Marketing Management, Journal of Business Research, Information & Management, IEEE Transactions on Engineering Management, International Journal of Production Economics, and Technological Forecasting and Social Change.

LeeAnn Kung is an Assistant professor of Information Systems at Georgia Southern University. She earned her PhD of Business (MIS) from Auburn University in 2015. She also earned a MS of Information Science from University at Albany, SUNY and a M.Ed. from National-Louis University. Her research interests include organizational capabilities, business value of information systems and innovations, and big data. She has published in top IS and business journals such as the *Journal of Information Technology*, *Industrial Marketing Management*, *Information and Management*, *Technological Forecasting and Social Change* and *Journal of Enterprise Information Management*.

Suraksha Gupta is the Professor of Socially Responsible Marketing at Newcastle University London. She holds a PhD from Brunel University in UK, an MBA from Institute of Management Technology in India and a Bachelor of Commerce from University of Delhi, India. Her research publications have appeared in highly reputed journals such as *Journal of World Business*, *Journal of Business Research*, *Industrial Marketing Management*, *European Journal of Marketing*, *Technological Forecasting and Social Change* and *Thunderbird International Business Review*.

Sena Ozdemir is a Senior Lecturer in Marketing at Coventry University, Business School, United Kingdom. Her research focuses on the areas of new product development including strategic new product alliances, global product development, the deployment of Big Data Analytics and digital technologies for innovation, inter- firm cross-functional new product development team integration, and social innovation. She is currently serving as the editorial board member for Journal of Business-to-Business Marketing. Her publications have appeared in journals such as Industrial Marketing Management, Journal of Business Research, Journal of Business-to-Business Marketing, and Innovation: Organization & Management, and so forth.