2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2014)

Application of Network Analysis on Healthcare

Fei Wang*†, Uma Srinivasan†, Shahadat Uddin‡ and Sanjay Chawla*
*School of IT, University of Sydney, Sydney, Australia
†Capital Markets Cooperative Research Centre, Sydney, Australia
‡Complex Systems Research Centre, University of Sydney, Sydney, Australia
fwan7956@uni.sydney.edu.au umas@cmcrc.com shahadat.uddin@sydney.edu.au sanjay.chawla@sydney.edu.au

***Abstract*—The healthcare sector holds large amounts of semantically rich electronic data generated and used by different sections of the health care community. Data analytic techniques such as data mining and predictive modelling are being used to gain new insights into health care costs, performance and quality of care. In this context, social network analysis (SNA) has the unique ability to play a new role in exploring the context and situations that lead to efficient and effective healthcare. In this paper we describe a specific context of private healthcare in Australia and describe our SNA based approach (applied to health insurance claims) to understand the nature of collaboration among doctors treating hospital inpatients and explore the impact of collaboration on cost and quality of care. In particular, we use network analysis to (a) design collaboration models among surgeons, anaesthetists and assistants who work together while treating patients admitted for specific types of treatments (b) identify and extract specific types of network topologies that indicate the way doctors collaborate while treating patients and (c) analyse the impact of these topologies on cost and quality of care provided to those patients.**

Keywords: Health care, Social network

I. Introduction

Social network analysis is commonly used to study relationships between individuals and communities as they interact with each other. Analysing Facebook connections is one such classic example. The textbook by Easley and Kleinberg [1] offers deep insight into the complexity of a connected world. More interesting and novel applications of network theory are reported in specialised domains [2], [3]. In the healthcare domain, social network analysis has been used in different settings, for example to study collaboration among healthcare professionals in specific healthcare environments, to understand the impact of team structure on quality of care [4], [5], [6]. In this paper we describe our approach of applying social network analysis in the domain of health insurance claims. In particular, we use data from health insurance claims to design network-based models of collaboration among medical providers and analyse the impact of social networks and their underlying network structures, to discover provider communities and analyse the topology of the emerging community structure (of surgeons, anaesthetists and assistant surgeons) on treatment outcomes for patients who undergo specific category of surgeries, for example knee surgeries.

Previous work on health insurance analytics using data mining and predictive modelling [7] gives a good understanding of the semantics and the data available in a private health insurance (PHI) claim, and the claiming patterns of hospitals and medical providers. As an organisation, we provide health insurance business intelligence and analytics services for the majority of the Australian PHI sector. Within the context of Australian PHI there are two types of claims. A medical claim is sent by a doctor - also referred to as a provider - who performs a service to treat a patient who is a member of a particular private health insurer. The medical claim has information about the provider, the member, the hospital where the patient was treated, the details of the treatment and the cost of the services provided. A hospital claim is sent by a hospital's billing department and includes details of treatment, theatre charges, accommodation charges, prosthetics charges and charges for other services provided. Leveraging on that understanding, we have started using social network analysis techniques to model provider relationships, and analyse the impact of provider community structures on healthcare costs and quality of care.

We present two types of networks to explore collaboration among medical providers: (**i**) collaboration networks (CN) designed to capture the collaboration among surgeons, anaesthetists and assistant surgeons (**ii**) surgeon centric collaboration networks (SCCN) which explore an individual surgeon's connections.

In terms of the network representations used in this paper, a node in the network represents a (medical) provider such as surgeon, anaesthetist, assistant surgeon; the node size indicates the total amount charged by that provider; the thickness of the edge (or tie strength) connecting two nodes represents the number of common hospital admissions between the two providers. In this paper, an admission refers to a single episode of admitted patient care. The time interval between the date of admission and the date of discharge represents the length of stay for that admission.

In addition to size of nodes and tie strength, other network measures - closeness centrality and betweenness centrality that are related to the position of the node in the network, and centralization measures that indicate how central its most central node is in relation to how central other nodes are, can provide interesting insights about the influence of the node in the overall communication control capacity and the network. For example, the larger nodes with a more influential position in the network have the capacity to provide additional meaning within the context of the graph.

In the context of healthcare, the questions we are trying to answer are:

- Is there a team structure that emerges as providers work together on a number of shared admissions?
- What is the impact of an individual surgeon's network on cost and quality of care of the surgeries performed?
- What types network structures have positive or negative impact on cost and quality of care?

Our experiments indicate that betweenness centralisation in the SCCN network is the variable that has significant positive influence on Length of Stay (LoS), Complication rate and Medical cost. This gives an indication that nodes with high betweenness centrality are likely to be in more demand.

Our theoretical analysis combined with empirical investigation over a large dataset also suggest that surgeons who collaborate with more number of teams appear to have a lower average LoS.

The rest of the paper is organized as follows: Section II presents a brief review of collaboration models explored in the context of healthcare domain. Section III describes our research methodology to explore collaborations among providers using PHI claims data. Section IV presents an analysis of our findings. And finally Section V presents some conclusions and future work.

II. Collaboration in health care

A. Literature review

In the healthcare sector, collaboration among healthcare professionals has been studied from several perspectives. Cunningham et al. (2012) [8] have conducted an orderly review of 26 studies of professionals' network structures and analysed factors connected with network effectiveness and sustainability specifically in relation to the quality of care and patient safety. They noticed that cohesive and collaborative health professional networks can contribute to improving quality and safety of care. A classic study, led by Knaus and his team, identifies a significant relationship between the degree of nurse-physician collaboration and patient mortality in intensive care units (Knaus et al., 1986) [9]. They study treatment and outcome in $5,030$ intensive care unit patients and find that hospitals where nurse-physician collaboration is present report a lower mortality rate compared to the predicted number of patient deaths. Conversely, hospitals that are noted for poor communication among healthcare professionals exceed their predicted number of patient deaths. In a two group quasi-experiment on $1,207$ general medicine patients ($n = 581$ in the experimental group who received care from a specially designed care management plan that facilitated higher collaboration among hospital staff and $n = 626$ in the control group who received the usual care), Cowan et al. (2006)[10] notice average hospital length of stay, total hospitalization cost and hospital readmission rate are significantly lower for patients in the experimental group than the control group (5 versus 6 days, $p < .0001$) which contributes a 'backfill profit' of USD1,591 per patient to hospitals. Sommers et al. (2000) [11] examined the impact of an interdisciplinary and collaborative practice intervention involving a primary care physician, a nurse and a social worker for community-dwelling seniors with chronic illnesses. They conducted a controlled cohort study of 543 patients in 18 private office practices of primary care physicians. The intervention group received care from their primary care physician working with a registered nurse and a social worker, while the control group received care as usual from primary care physicians. They noticed that the intervention group produced better results in relation to readmission rates and average office visits to all physicians. Moreover, the patients in the intervention group also reported an increase in social activities compared with the control group. There are other studies emphasizing collaboration for effective patient outcome across professional boundaries within hospitals. By analysing data collected from 105 interviews (with 40 physician, 32 case managers, 23 physician office staff, 8 administrators and 2 case assistants), Netting and Williams (1996) [12] argue that there is a growing need to collaborate and communicate across professional lines rather than make assumptions about who can do what for better patient outcomes, professional satisfaction and hospital performance. There are other studies that analyse networked collaboration among healthcare specialists to explore different aspects of professional behaviour and quality patient care. For example, Fattore et al. (2009) [13] evaluate the effects of GP network organisation on their prescribing behavior and (Meltzer et al., 2010) [14] develop a selection criteria of group members in order to improve the effectiveness of team-based approach to patient care. Other studies include physician-pharmacist collaboration (Hunt et al., 2008) [15], physician-patient collaboration (Arbuthnott & Sharpe, 2009) [16], hospital-physician collaboration (Burns & Muller, 2008) [17], and inter-professional and interdisciplinary collaboration (Gaboury et al., 2009) [18].

Our study of surgeon collaboration presented in this paper offers a unique perspective as it combines theoretical analysis with empirical investigations of a PHI large dataset. The design of the collaboration model presented in this paper is influenced by the requirements of domain experts who wish to understand the nature of team structures that have an impact on cost as well as quality of care provided to patients for specific types of treatments. In order to keep it simple, in this paper we have used only knee procedures as the exemplar treatment group. We have designed similar models for other orthopaedic procedures as well other treatment groups such as cardiology and cardio-thoracic procedures. More details of the application scenario is explained in the following section.

B. Application scenario

The hospital and medical claims processed by an insurer contain data that specify the type of service provided during an admission, the length of stay for that admission, and the cost of that service. The service is specified as a Medicare Benefit Schedule (MBS) code [19], as stipulated by the Australian Government. The hospitals also send additional data related to an admission, once the patient is discharged. For any given hospital admission, we deal with three sets of data:

1) Medical claims - these show the provider-ID i.e. who performed a service, and the service is indicated by the MBS code for the specific type of treatment performed while the patient was in hospital;
2) Hospital claims - these are sent by the billing department of the hospital, and include MBS codes, accommodation cost, prosthetic costs, laboratory and radiology costs; and
3) Hospital discharge data that consolidates the patient's clinical care during that particular admission, and includes details such as length of stay, whether this was an unplanned readmission, and additional diagnosis codes that indicate complications or infections that occurred during that admission. Therefore the discharge data provides us with valuable information pertaining to quality of care.

We use data from all three sources to design our network models. The network graph presented in this paper represents the collaboration among three specific types of medical providers; the surgeons, the anaesthetists and assistant surgeons, as they perform knee-related surgical procedures.

III. Surgeon Collaboration network Design

In this section, we first report on the graphical models designed to capture the collaboration among medical providers. In addition we also explore a Surgeon-Centric Collaboration Network (SCCN) which explores an individual surgeon's connections. Finally, we provide network concepts that are related to our work.

A. *Design of Collaboration Network (CN)*

The objective of our initial design is to investigate the quality of care provided by a specific provider or a group of providers who collaborate while performing knee surgeries. The PHI domain experts are interested in understanding the impact of collaboration among the three types of providers: surgeons, anaesthetists, assistant surgeons (also refer to assistants in the rest of the paper). This leads us to design a tripartite graph in which the nodes correspond to the three types of providers.

The data we have for any private health insurer include the three types of datasets specified in Section II-B. Therefore, the information represented in the three sets of data offers us content-rich health information about each admission episode. The admissions are categorized by the treatment codes as specified in the MBS coding taxonomy. For example, knee surgeries are coded in the following hierarchy: 'Therapeutic → Surgical Operation → Orthopaedic → Knee'. The nodes represent the providers, and the edges as the number of common admissions shared by the two providers. We then associate the node size with the total medical charges of the corresponding provider. The three different types of providers are shown in three distinct colors: red for surgeons, blue for anaesthetists and light blue for assistants. A thicker edge indicates a higher number of shared admissions. Figure 1 shows an example of a collaboration graph.

Just by glancing at the graph, one can immediately identify the 'big' providers, i.e. providers with high medical charges, as well as highly connected providers. We can also see isolated cluster of providers. Often such isolated clusters indicate providers working in a specific geographic region. This network graph offers a powerful visualization to study collaboration among providers.

The primary focus of our investigation is to study the impact of the collaboration network structure on quality of care. To do this we consider all possible network features.

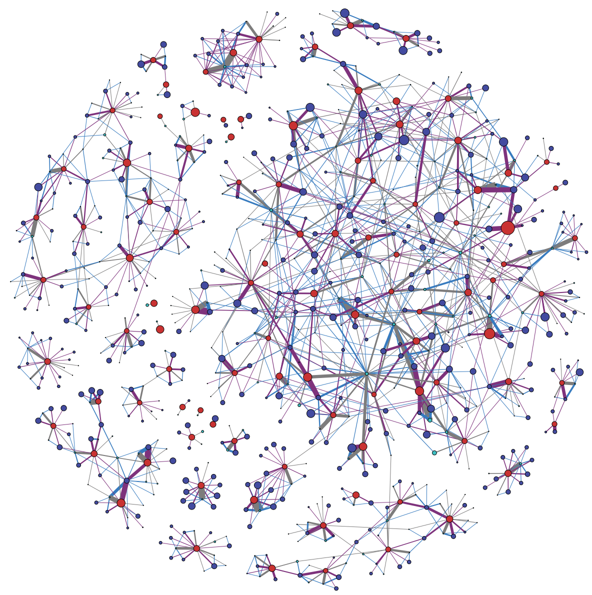

Fig. 1: The tripartite graph represent the collaboration between three types of providers: surgeons (red), anaesthetists (blue) and assistants (light blue). The edge thickness is modeled as the number of collaborating claims by two types of providers. The size of the node is modeled as the medical charge of the provider.

B. *Design of Surgeon Centric Collaboration Network (SCCN)*

Since the focus is on surgeons, we investigate a specific surgeon node in the CN and build a lower level Surgeon-Centric Collaboration Network (SCCN). The SCCN is a network of a specific surgeon. It shows how a specific surgeon collaborates with the assistants and anaesthetists, and the hospital(s) in which they work together while performing knee surgeries. The individual surgeon node is not shown in the SCCN as all admissions (which are modeled as the edges) relate to a particular surgeon. Therefore, we only model two types of edges, one is the edge between assistants and hospitals, the other is between anesthetists and hospitals. Since it's a surgeon centric network, we have not shown the links between assistants and anesthetists. However such links are shown in the CN graph. The SCCN network also shows the hospitals where the surgeon performs knee surgeries. The hospital node is represented symbolically in the form of a building. The size of the building indicates the total medical cost. Edge thickness is modeled as the number of admissions of the specific surgeon with an anaesthetic or assistant in that hospital. Two SCCN graphs are shown in Figure 2. The graph on the left shows a surgeon who only works in one hospital and collaborates with nine anaesthetists or assistants. The graph on the right shows a surgeon who works in two hospitals and collaborates with anaesthetists or assistants who also work in those hospitals.

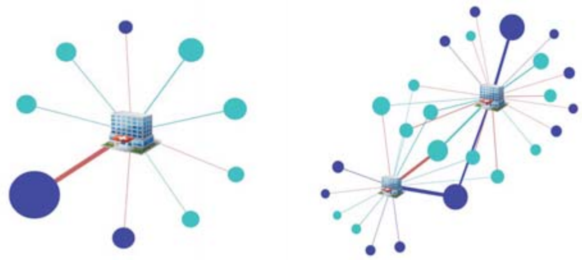

Fig. 2: Two SCCN graphs with hospital represented by a building icon.

C. Network concepts

1) Degree Centralisation: Before explaining degree centralisation, degree centrality should be defined. Degree centrality is one of the basic measures of network centrality. It is the proportion of nodes that are adjacent to a particular actor in a network. Degree centrality highlights the node with the most links to other actors in a network, and can be defined by the following equation for the actor (or node) i in a network having N actors (Wasserman and Faust 2003) [20]:

$$C_D^{'} = \frac{d(n_i)}{N-1}$$

The subscript D for 'degree' and $d(n_i)$ indicates the number of actors with whom actor i is connected. The maximum value for $C_D^{'}$ is 1 when actor i is linked with all other actors in the network. The set of degree centralities, which represents the collection of degree indices of N actors in a network, can be summarised by the following equation to measure network degree centralisation (Freeman et al. 1979) [21]:

$$C_D = \frac{\sum_{i=1}^{N}[C_D(n^*) - C_D(n_i)]}{[(N-1)][(N-2)]}$$

Where, $\{C_D(n_i)\}$ are the degree indices of N actors and $C_D(n^*)$ is the largest observed value in the degree indices. For a network, degree centralisation (i.e. the index C_D) reaches its maximum value of 1 when one actor chooses all other $(N-1)$ actors and the other actors interact only with this one (i.e. the situation in a star graph as illustrated in Figure 3). This index

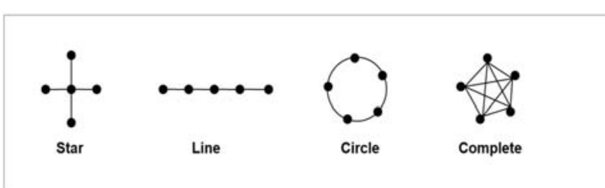

Fig. 3: Illustration of a star, a line, a circle and a complete graph.

(i.e. C_D) attains its minimum value of 0 when all degrees are equal (i.e. the situation in a circle graph as illustrated in Figure 3). Thus, C_D indicates varying amounts of centralisation of degree compared to both star and circle graphs.

2) Closeness Centralisation: Closeness centrality needs to be defined before explaining closeness centralisation. Closeness centrality, another view of actor centrality based on closeness or distance, focuses on how 'close' an actor is to all the other actors in a network (Freeman et al. 1979) [21]. The idea is that an actor is central if it can quickly interact with all other actors in a network. In the context of a communication relation, such actors need not rely on other actors for the relaying of information. For an individual actor, it can be represented as a function of shortest distances between that actor and all other remaining actors in the network. The following equation represents the closeness centrality for a node i in a network having N actors (Freeman et al. 1979; Wasserman and Faust 2003) [21], [20]:

$$C_C^{'}(n_i) = \frac{N-1}{\sum_{j=1}^{N} d(n_i, n_j)}$$

Where, the subscript C for 'closeness', $d(n_i, n_j)$ is the number of lines in the shortest path between actor i and actor j, and the sum is taken over all $i \neq j$. A higher value of $C_C^{'}(n_i)$ indicates that actor i is closer to other actors of the network, and will be 1 when actor i has direct links with all other actors of the network. The set of closeness centralities, which represents the collection of closeness indices of N actors in a network, can be summarised by the following equation to measure network closeness centralisation (Freeman et al. 1979) [21]:

$$C_C = \frac{\sum_{i=1}^{N}[C_C^{'}(n^*) - C_C^{'}(n_i)]}{[N-1][N-2]/[2N-3]}$$

Where, $\{C_C^{'}(n_i)\}$ are the closeness indices of N actors and $C_C^{'}(n_i)$ is the largest observed value in closeness indices. For a network, closeness centralisation (i.e. the index C_C) reaches its maximum value of unity when one actor chooses all other $(N-1)$ actors and the other actors have shortest distances (i.e. geodesics) of length 2 to the remaining $(N-2)$ actors (i.e. the situation in a star graph as illustrated in Figure 3). This index (i.e. C_C) can attain its minimum value of 0 when the lengths of shortest distances (i.e. geodesics) are all equal (i.e. the situation in a complete graph and circle graph as illustrated in Figure 3). Thus, indicates varying amounts of centralisation of closeness compared to star, circle and complete graph.

3) Betweenness Centralisation: Betweenness centrality will be defined first before explaining betweenness centralisation. Betweenness centrality is obtained by determining how often a particular node is found on the shortest path between any pair of actors (or nodes) in the network. It views an actor as being in a favoured position to the extent that the actor falls on the shortest paths between other pairs of actors in the network. That is, nodes that occur on many shortest paths between other pairs of nodes have higher betweenness centrality than those that do not (Freeman 1978) [22]. Therefore, it can be regarded as a measure of strategic advantage and information control. In a network of size n, the betweenness centrality for an actor (or node) i can be represented by the following equation (Wasserman and Faust 2003) [20]:

$$C'_B(n_i) = \frac{\sum_{j<k} \frac{g_{ij}(n_i)}{g_{jk}}}{[(N-1)(N-2)]/2}$$

Where, $i \neq j \neq k$; $g_{jk}(n_i)$ represents the number of shortest paths linking the two actors that contain actor i; and g_{jk} is the number of shortest paths linking actor j and k. From the set of betweenness centralities of N actors in a network betweenness centralisation can be defined by the following equation:

$$C_B = \frac{\sum_{i=1}^{N} [C_B(n^*) - C_B(n_i)]}{N-1}$$

Where, $\{C'_B(n_i)\}$ are the betweenness indices of N actors and $C_B(n^*)$ is the largest observed value in betweenness indices. For a network, betweenness centralisation (i.e. the index C_B) reaches its maximum value of unity when one actor chooses all other $(N-1)$ actors and the other actors have shortest distances (i.e. geodesics) of length 2 to the remaining $(N-2)$ actors (i.e. the situation in a star graph as illustrated in Figure 3). This index (i.e. C_B) can attain its minimum value of 0 when all actors have exactly the same actor betweenness index (i.e. the situation in a line graph as illustrated in Figure 3). Thus, C_B indicates varying amounts of centralisation of betweenness compared to both star and line graph.

4) Density: Density measures how connected a graph is. For example, if a graph G has N nodes, V edges. Then the density D_G of the graph G is calculated as :

$$D_G = \frac{2 * V}{N(N-1)}$$

D_G reaches maximum value as 1 when the graph is fully connected, and reaches minimum value as 0 when there is no edge.

IV. Data analysis

This section describes the experimental analysis staring with the data preparation, selection of network variables , selection of quality of care parameters, the regression model and finally an empirical investigation to compare the theoretical results within the context of the large PHI data corpus.

A. Data preparation

1) Selection of admission-related variables : In terms of non-network variables, we have identified four admission related features, which are shown in the top section of Table I. The admission data shows all the medical providers who are involved in treating a patient during that admission. We consider four types of providers that includes: anaesthetists, assistants, pathologists and imaging providers. Typically a surgery has one principal surgeon and assistant and anaesthetist who work with the surgeon during the surgery. Specifically, we consider the number of distinct providers a surgeon collaborates with while performing a knee surgery. For the data analysis, we consider the percentage of distinct providers who collaborate with the surgeon rather than the absolute number of providers. The percentage is calculated with the denominator as the sum of distinct number of the four types of providers collaborating with the surgeon in knee procedure.

2) Selection of network variables: For network features, we first consider CN graph as depicted in Figure 1. We have three types of nodes in the graph, out of which, around 500 nodes are 'surgeon' nodes. Amongst several possible network variables, we have specifically selected five network features as shown in Table I. We have chosen these variables as they have the potential to offer insights into the collaboration patterns among providers.

Clustering-coefficient: The local clustering coefficient of a vertex (node) in a graph quantifies how close its neighbors are to being a clique (complete graph). In our context, this represents the strength of the surgeon's network.

Number of triangles: The global clustering coefficient is based on triplets of nodes. A triplet consists of three nodes that are connected by either two (open triplet) or three (closed triplet) undirected ties. A triangle consists of three closed triplets, one centred on each of the nodes. In our context, a triangle shows the three types of providers i.e. surgeons, anaesthetists and assistants working together while performing knee surgeries. For example, in Figure 4, we depict a surgeon node with its surrounding assistants and anaesthetists. This is a small subset extracted from the CN graph shown in Figure 1. The surgeon node, which is represented as the red colour, has three triangles connected to it. This indicates that the surgeon performs knee surgeries frequently with three pairs of assistants and anaesthetists. The three centrality measures shown in Table I are explained in detail in Section III-C. The features from the CN graph are all node-level features.

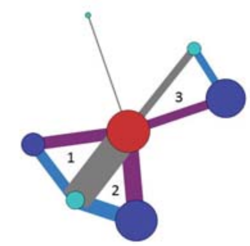

Fig. 4: Surgeon node with Number of triangle as 3

Next we consider network-level features. For each surgeon, we have one SCCN graph. We will consider the four network measures for the SCCN graph which are shown in the bottom section of Table I.

3) Selection of Quality of Care parameters : As for the quality of care, we have chosen three parameters:

LoS - the average length of stay for all admissions of the surgeon. This information is available in hospital discharge data as explained in Section II-B.

Medical cost - the average medical charge for all the knee related admissions teated by the surgeon.

Complication rate - calculated as the percentage of admissions with complications out of all the admissions of a surgeon.

4) Data cleansing and transformation: Our dataset for knee surgeries includes a total of 59, 256 admissions performed by 870 surgeons. However, in order to make robust conclusions,

TABLE I: The table shows all the features we have extracted from the data and the network.

	Non-network features
1	%. of distinct anaesthetists
2	%. of distinct assistants
3	%. of distinct pathologists
4	%. of distinct imaging providers
	Network features of CN
1	Clustering coefficient
2	Number of triangles
3	Degree centrality
4	Closeness centrality
5	Betweenness centrality
	Network features of SCCN
1	Degree centralisation
2	Closeness centralisation
3	Betweenness centralisation
4	Density

we only considered surgeons who had more than ten claims. We further looked at the distribution of the surgeons according to each variable shown in Table I and removed surgeons who appeared as outliers. Our analysis was carried out on a set of 559 surgeons. For all the variables in Table I, we applied z-score standardization. Thus each variable had a mean of zero and a standard deviation of one. In the simple regression analysis, that we will report on, this allow us to interpret the constant and the "slope" term appropriately.

B. Simple linear regression

The quality of care parameters introduced in Section IV-A3 are the dependent variables in all the regression experiments. Since the independent variables are semantically distinct in the healthcare domain, they have been dealt with independently. Hence an individual linear model has been constructed for each variable. Although most of the linear models have a low R^2 value, our focus are the β values, which are significant.

1) Non network features: Table II explores the impact of all admission-related features on the dependent variables (i.e. LoS, and Medical cost). We can see that a higher percentage of anaesthetist and assistant indicates a lower LoS and Medical cost, while a higher percentage of pathologists and imaging providers indicates a higher LoS and Medical cost. This is intuitive since admissions with more imaging may be more severe situations and thus incur longer LoS and higher Medical cost.

2) Network features of SCCN: In Table III, we observe that betweenness centralisation is the only variable that has significant positive influence on LoS, Complication rate and Medical cost. This can be interpreted as follows: From the perspective of a SCCN structure, a high betweenness centralisation indicates that the structure of the corresponding SCCN follows a star-like or centralized structure since betweenness centralisation reaches its highest value of 1 for a star network. A star-like or centralized network has few actors with higher betweenness centrality values and the rest actors have very low betweenness centrality values. In this type of network, only a small number of actors play major collaboration and communication role (Wasserman and Faust 2003). That indicates there is a presence of "network hubs" in this type of network. On the other side, if network actors have almost equal level of network connectivity (as like a line graph) then betweenness centralisation will be small and in such networks there does not present any "network hub". Therefore, SCCN, where participating actors have almost equal level of network connectivity, will produce lower LoS, Complication rate and Medical cost. In the context of health care domain, this offers an interesting insight. In their corresponding hospitals, healthcare managers or administrators could encourage a practice culture where each member will have equal level of network connectivity.

3) Network features of CN: Table IV explores the impact of the network position of the individual specialist in the complete network (CN) on independent variables (i.e. LoS, Complication rate). We can observe that in the case of both LoS and Complication rate, the variable 'Number of triangles' has a negative correlation. That is, when a surgeon works with a large number of distinct groups, LoS and Complication rate are lower.

Intuitively, we have two assumptions with respect to the variable 'Number of triangles': (**i**) Surgeons who work with large number of distinct assistants or anaesthetists could be involved in more complicated surgeries and thus resulting longer Los and higher complication rate. (**ii**) Surgeons who consistently work with only a few distinct assistants or anaesthetists have a lower number of triangles. For these cases, our analysis shows a higher LoS. Our conjecture is that this limits external influence of other providers on the surgeon. The converse case where the number of triangles is higher clearly shows lower LoS. Thus, to figure out which assumption is true we investigated the different categories of knee surgeries and their impact on LoS in Section IV-C.

C. Treatment analysis

Table V shows the distribution of the different types of knee surgeries performed in the dataset used for analysis in this paper. The dataset used includes about $59,256$ knee surgeries performed by 559 surgeons over a period of 2 years. As per the MBS descriptions, there are four broad categories of knee surgeries with varying degrees of complexity. Accordingly, the average length of stay for each category of knee surgery varies. Column 3 of the table also shows the distribution of the four categories of knee surgery. We conducted an empirical investigation to analyze the performance of teams of surgeons indicated by the No, of triangles as shown in Table IV. Table VI shows two groups of providers: Group A and group B. Group A represents the 200 surgeons having the least Number of triangles, and group B represent 200 surgeons with largest Number of triangles. The purpose of this analysis is to compare the Average LoS for each category of knee surgery for the two groups of providers.

TABLE II: Table explores the impact of all non-network attributes on quality of cares (i.e. LoS, Medical cost)

Model	Dependent Variable	Independent Variable	R^2 value	β	Constant	Sig.
1	LoS	**%. of distinct anaesthetists**	**0.098**	**-0.438**	**3.506**	**0**
2		**%. of distinct assistants**	**0.023**	**-0.214**	**3.506**	**0**
3		%. of distinct pathologists	0.003	0.074	3.506	0.211
4		**%. of distinct imaging providers**	**0.179**	**0.592**	**3.506**	**0**
5	Medical cost	**%. of distinct anaesthetists**	**0.042**	**-58.127**	**1016.063**	**0**
6		**%. of distinct assistants**	**0.011**	**-29.481**	**1016.063**	**0.015**
7		%. of distinct pathologists	0.002	14.200	1016.063	0.240
8		**%. of distinct imaging providers**	**0.074**	**77.484**	**1016.063**	**0**

TABLE III: The table explores the impact of the network structure around a specialist (based on SCCN) on quality of cares (i.e. LoS, Complication rate and Medical cost).

Model	Dependent Variable	Independent Variable	R^2 value	β	Constant	Sig.
1	LoS	**Degree centralization**	**0.014**	**0.164**	**3.506**	**0.005**
2		**Closeness centralization**	**0.023**	**0.212**	**3.506**	**0**
3		**Betweenness centralization**	**0.033**	**0.253**	**3.506**	**0**
4		Density	0	0.024	3.506	0.681
5	Complication rate	Degree centralization	0.002	0.002	0.047	0.343
6		Closeness centralization	0.001	0.001	0.047	0.496
7		**Betweenness centralization**	**0.014**	**0.006**	**0.047**	**0.005**
8		Density	0.001	-0.001	0.047	0.494
9	Medical cost	Degree centralization	0	-1.684	1016.063	0.889
10		**Closeness centralization**	**0.013**	**32.698**	**1016.063**	**0.007**
11		**Betweenness centralization**	**0.011**	**29.638**	**1016.063**	**0.014**
12		**Density**	**0.009**	**-27.014**	**1016.063**	**0.025**

TABLE IV: The table shows the impact of network position of individual specialist in the complete network (CN) on quality of cares (i.e. LoS, Complication rate).

Model	Dependent Variable	Independent Variable	R^2 value	β	Constant	Sig.
1	LoS	Clustering coefficient	0.001	0.052	3.506	0.384
2		**Number of triangles**	**0.005**	**-0.101**	**3.506**	**0.089**
3		Degree centrality	0.003	-0.080	3.506	0.179
4		Closeness centrality	0.004	-0.085	3.506	0.149
5		Betweeness centrality	0	-0.016	3.506	0.788
6	Complication rate	Clustering coefficient	0.002	-0.002	0.047	0.277
7		**Number of triangles**	**0.007**	**-0.004**	**0.047**	**0.048**
8		Degree centrality	0.002	-0.002	0.047	0.298
9		Closeness centrality	0.002	0.002	0.047	0.350
10		Betweeness centrality	0	0.001	0.047	0.712

TABLE V: Average LoS and percentage of admissions of the four knee categories.

Treatment type	Average LoS	Admissions %
Knee arthroscopy	1.24	**58**
Knee Revision	4.40	1
Knee Reconstruction	2.15	9
Knee Replacement	7.66	28

Next we compare the Average LoS for each category of knee surgery for the two groups in Table VI with the Average LoS of the complete dataset summarised in Table V. We can observe that in all four categories of knee surgeries, group B consistently has a lower Average LoS compared to group A, as well as the whole dataset shown in Table V. The empirical investigation implies that surgeons who work with a higher number of teams appear to have a lower length of stay. One could intuit that there is social learning that comes into play. However, further investigation is required to confirm

TABLE VI: We can observe that, in terms of all the four treatment types, group B consistently has a lower Average LoS compared to group A and also the whole dataset as shown in Table V.

Treatment type	**Average LoS**		**Admissions %**	
Group	A	**B**	A	B
Knee arthroscopy	1.31	**1.21**	58	57
Knee Revision	7.69	**4.01**	1	1
Knee Reconstruction	2.38	**2.07**	7	10
Knee Replacement	7.67	**7.61**	31	28

the intuitive analysis. Since the category distribution for group A and B are almost identical. This makes assumption (**i**) in Section IV-B3 invalid. Thus assumption (**ii**): Lower Number of triangles limits external influence of other providers on the surgeon, is a possible explanation.

V. Summary and future directions

In this paper we have investigated the impact of network structure on the performance of surgeon teams with respect to efficiency metrics including Medical costs, Length of Stay (LoS) and Complication rate. Our data set was obtained from Australian PHI data and consists of both medical and hospital claims. To reduce the impact of confounding variables, we focused our analysis on "knee surgeries." Our results provide a strong indication that network features like degree, betweenness and closeness centralization and number of triangles have a statistically significant impact on efficiency metrics. In particular, for surgeon centric networks, betweenness centralization is significant for all three metrics: Length of Stay, Complication rate and Medical cost. This observation can potentially be used by health care providers to reorganize surgical teams and improve the overall efficiency of health care delivery. For future work we would like to investigate the question of "social learning," i.e., medical providers who are connected to each other, over time, develop similar treatment delivery patterns.

VI. Acknowledgments

The authors of this paper acknowledge the financial support of the Capital Markets CRC.

References

[1] D. Easley and J. Kleinberg, "Networks, crowds, and markets: Reasoning about a highly connected world," 2012.

[2] T. B. Achacoso and W. S. Yamamoto, *AY's Neuroanatomy of C. elegans for Computation*. CRC Press, 1991.

[3] L. A. Adamic and N. Glance, "The political blogosphere and the 2004 us election: divided they blog," in *Proceedings of the 3rd international workshop on Link discovery*, pp. 36–43, ACM, 2005.

[4] S. Uddin, L. Hossain, J. Hamra, and A. Alam, "A study of physician collaborations through social network and exponential random graph," *BMC health services research*, vol. 13, no. 1, pp. 234–247, 2013.

[5] B. E. Landon, N. L. Keating, M. L. Barnett, J.-P. Onnela, S. Paul, A. J. OMalley, T. Keegan, and N. A. Christakis, "Variation in patient-sharing networks of physicians across the united states," *JAMA*, vol. 308, no. 3, pp. 265–273, 2012.

[6] M. L. Barnett, N. A. Christakis, A. J. OMalley, J.-P. Onnela, N. L. Keating, and B. E. Landon, "Physician patient-sharing networks and the cost and intensity of care in us hospitals," *Medical care*, vol. 50, no. 2, pp. 152–160, 2012.

[7] U. Srinivasan and B. Arunasalam, "Leveraging big data analytics to reduce healthcare costs," IEEE, 2013.

[8] F. C. Cunningham, G. Ranmuthugala, J. Plumb, A. Georgiou, J. I. Westbrook, and J. Braithwaite, "Health professional networks as a vector for improving healthcare quality and safety: a systematic review," *BMJ quality & safety*, vol. 21, no. 3, pp. 239–249, 2012.

[9] W. A. Knaus, E. A. DRAPER, D. P. Wagner, and J. E. Zimmerman, "An evaluation of outcome from intensive care in major medical centers," *Annals of Internal Medicine*, vol. 104, no. 3, pp. 410–418, 1986.

[10] M. J. Cowan, M. Shapiro, R. D. Hays, A. Afifi, S. Vazirani, C. R. Ward, and S. L. Ettner, "The effect of a multidisciplinary hospitalist/physician and advanced practice nurse collaboration on hospital costs," *Journal of Nursing Administration*, vol. 36, no. 2, pp. 79–85, 2006.

[11] L. S. Sommers, K. I. Marton, J. C. Barbaccia, and J. Randolph, "Physician, nurse, and social worker collaboration in primary care for chronically ill seniors," *Archives of Internal medicine*, vol. 160, no. 12, pp. 1825–1833, 2000.

[12] F. E. Netting and F. G. Williams, "Case manager-physician collaboration: Implications for professional identity, roles, and relationships," *Health & Social Work*, vol. 21, no. 3, pp. 216–224, 1996.

[13] G. Fattore, F. Frosini, D. Salvatore, and V. Tozzi, "Social network analysis in primary care: The impact of interactions on prescribing behaviour," *Health Policy*, vol. 92, no. 2, pp. 141–148, 2009.

[14] D. Meltzer, J. Chung, P. Khalili, E. Marlow, V. Arora, G. Schumock, and R. Burt, "Exploring the use of social network methods in designing healthcare quality improvement teams," *Social science & medicine*, vol. 71, no. 6, pp. 1119–1130, 2010.

[15] J. S. Hunt, J. Siemienczuk, G. Pape, Y. Rozenfeld, J. MacKay, B. H. LeBlanc, and D. Touchette, "A randomized controlled trial of team-based care: impact of physician-pharmacist collaboration on uncontrolled hypertension," *Journal of general internal medicine*, vol. 23, no. 12, pp. 1966–1972, 2008.

[16] A. Arbuthnott and D. Sharpe, "The effect of physician–patient collaboration on patient adherence in non-psychiatric medicine," *Patient education and counseling*, vol. 77, no. 1, pp. 60–67, 2009.

[17] L. R. Burns and R. W. Muller, "Hospital-physician collaboration: Landscape of economic integration and impact on clinical integration," *Milbank Quarterly*, vol. 86, no. 3, pp. 375–434, 2008.

[18] I. Gaboury, M. Bujold, H. Boon, and D. Moher, "Interprofessional collaboration within canadian integrative healthcare clinics: Key components," *Social Science & Medicine*, vol. 69, no. 5, pp. 707–715, 2009.

[19] "MBS online." http://www.mbsonline.gov.au/.

[20] W. Stanley and K. Faust, "Social network analysis: methods and applications," *Cambridge: Cambridge University*, 2003.

[21] L. C. Freeman, D. Roeder, and R. R. Mulholland, "Centrality in social networks: Ii. experimental results," *Social networks*, vol. 2, no. 2, pp. 119–141, 1980.

[22] L. C. Freeman, "Centrality in social networks conceptual clarification," *Social networks*, vol. 1, no. 3, pp. 215–239, 1979.