applied sciences

MDPI

Article

A Visual Analytics Framework for Inter-Hospital Transfer Network of Stroke Patients

Kyuhan Kwak, Jinu Park and Hyunjoo Song *

School of Computer Science and Engineering, Soongsil University, Seoul 06978, Republic of Korea
* Correspondence: hsong@ssu.ac.kr; Tel.: +82-2-820-0676

Abstract: Effective inter-hospital coordination is crucial in improving the stroke treatment process and outcomes. The introduction of endovascular thrombectomy (EVT) further emphasized the importance of coordination. Although previous studies considered various clinical data besides stroke in terms of the network structure between hospitals, a majority of these studies performed only quantitative analyses instead of topological analyses. This study proposes a new framework (PatientFlow) for constructing a network based on stroke patient transfer data and performing exploratory analysis. The proposed framework can visualize the network structure among hospitals at the national level and analyze the detailed structure through dynamic queries. The hub-and-spoke structure for each cluster derived through community detection can be compared visually and analyzed quantitatively using network measures. Further, the relationship between regions can be analyzed by aggregating the transfer of patients by province. PatientFlow allows medical researchers to perform an exploratory analysis to understand the network at the national, provincial, and community levels with multiple coordinated views.

Keywords: inter-hospital coordination; community detection; hub-and-spoke structure; stroke patient transfer data; endovascular thrombectomy (EVT)

Citation: Kwak, K.; Park, J.; Song, H. A Visual Analytics Framework for Inter-Hospital Transfer Network of Stroke Patients. *Appl. Sci.* **2023**, *13*, 5241. https://doi.org/10.3390/app13095241

Academic Editors: Il Dong Yun and Walid Abdullah Al

Received: 14 February 2023
Revised: 16 April 2023
Accepted: 19 April 2023
Published: 22 April 2023

1. Introduction

Improving the process and outcomes of stroke treatment requires effective inter-hospital coordination [1]. The introduction of endovascular thrombectomy (EVT) highlighted the significance of coordination even further [2]. The effectiveness of EVT was proved in several studies conducted in 2015 [3,4]; however, it was challenging to introduce it in all hospitals. Since only some major hospitals can afford the environment, smaller hospitals transfer patients who need EVT treatment to a capable hospital. Consequently, given this interest in transferring stroke patients, recent research focused on how the network between hospitals is structured [1,5].

Despite the interest and importance, information on network structure for stroke treatment is unknown and still needs to be studied [6]. In the Republic of Korea, 70 medium catchment areas have been established for inter-hospital coordination, and they are managed along with 17 large catchment areas centered on tertiary hospitals [7]. However, there could be differences between the catchment area established by the government and the service area in which healthcare institutions occupy the market through mutual competition. Thus, there are ongoing discussions between policymakers of central and local governments about determining the catchment areas and inspecting whether the hospitals within the areas are coordinating with each other [7].

Although previous studies considered various clinical problems besides stroke in terms of the network structure between hospitals, most of these studies performed only quantitative analyses on the related clinical outcomes instead of topological analysis. For example, studies on hospital-acquired infections (HAIs) analyzed national hospital networks in the United States [8] and France [9]. The network structure was used to develop a method for rapidly detecting disease prevalence [8] or a method to construct an optimized network for HAIs from patient transfer data [9]. Although stroke-related network studies [1,5] used geospatial data for the analysis, they focused on changes in the stroke treatment outcomes following the introduction of EVT or the increase or decrease in the number of patients transferred between hospitals rather than deriving and analyzing a hub-and-spoke structure of communities.

Therefore, deriving clear quantitative criteria for well-formed clusters from existing studies was challenging. As a result, it was necessary to perform clustering on the given data and then visually explore the well/poorly formed cases based on domain knowledge by looking at the topology of each cluster. Based on the results of this exploratory analysis, one could examine the network metrics of well/poorly formed clusters and heuristically derive quantitative commonalities for each case. However, there are difficulties in dynamically analyzing partial data of interest (e.g., clusters showing similar network metric values) with the abovementioned approaches since they did not focus on performing exploratory analysis.

This study proposes an interactive framework, PatientFlow, for network analysis of stroke care hospitals. It introduces a community-detection algorithm for patient transfer data at the national level to create an environment that facilitates an exploratory analysis by reflecting the geospatial characteristics of hospitals. The contributions of this study are summarized as follows:

1. Introduction of an integrated visual analysis environment supporting data preprocessing, cleansing, and filtering, allowing users to apply them dynamically;
2. Introduction of three-level exploratory analysis of patient transfers networks (i.e., national, provincial, and community levels) with multiple coordinated views;
3. Introduction of a comparative analysis of commonly used network measures between communities in clinical research;
4. Evaluation of efficacy through case studies with clinical researchers and comparison with previous studies.

The rest of this paper is organized as follows. Section 2 reviews the cases of analyzing stroke networks and visually analyzing network data in the healthcare field. We summarize the data covered in this study, describe the data types and preprocessing procedures, and provide an abstraction of the network analysis tasks performed by medical researchers (Section 3). Then, we introduce the visual analysis framework and summarize how the visualization has been constructed to reflect the task analysis results (Section 4). The content and results of the case study conducted with domain experts are described, and this is followed by further discussions. Finally, we conclude with Section 6, recapitulating the main findings and introducing future work.

2. Related Work

The patient transfer was visually analyzed with node-link diagrams for various clinical purposes. Fernández-Gracia et al. [8] analyzed the spread of pathogens in the patient transfer network in the United States. The network structure constructed based on patient transfers for two years was plotted on a map with the number of transferred patients indicated by the color of the link. The analysis results were used to introduce a method for selecting sensor hospitals based on a network measure (i.e., in-degree) and for quickly detecting disease prevalence based on temporal, geographical, and topological properties. Nekkab et al. [9] divided the patient group (1) with HAI, (2) with suspected HAI, and (3) all patients, and they derived the network connection structure based on the patient transfer pattern for determining the optimal network for studying HAIs. The transfer pattern was also presented on a map, and the authors identified notable transfer patterns that could help find a critical hospital at the regional and county level in France. The nodes in the diagram were colored according to the community derived from the detection algorithm [10]. Dong et al. [11] used a community detection algorithm [12], which extends the previous approach [10] for large networks, to analyze the clustering patterns of hospitals in China. The authors neglected geographic information in analyzing the network and plotting the nodes and links. We presented a hospital network using both geographical layout and force-directed layout algorithms for analysis of geographical relationships and topological analysis within each cluster, respectively.

A study on patient transfers was conducted regarding treatments for stroke patients. Adeoye et al. [2] studied the geographic access of stroke patients in the United States. They argued that a more efficient stroke treatment system was required by analyzing country-level data on the accessibility to hospitals and the number of patients who received treatments from the hospitals. Accessibility to treatment facilities was visualized using a choropleth map. Zachrison et al. [7] argued that the study of networks could help improve stroke treatment systems, and they introduced examples of community-detection algorithms employed in other medical studies. It was followed by studies that grafted a network analysis on various regions. For example, Zachrison et al. [5] analyzed stroke patient transfer data in the northeastern United States and visualized the network using ArcGIS. Each network constructed based on patient transfer data at two time-points (2007 and 2011) was expressed as a node-link diagram on a map; hospitals that sent a significant number of patients and hospitals that received them were analyzed. Another network analysis of the stroke treatment system was conducted for California [1]. The two networks constructed based on patient transfer data (between 2013 and 2014; and 2016 and 2017) were visualized on a map to examine changes from the introduction of EVT; the rate of progression of the EVT treatment and the rate of patient transfer were analyzed. Meanwhile, a study on telestroke care analyzed the hub-and-spoke structures and investigated how the network changed over time [13]. Our study analyzed the community in the network based on patient transfer data using the Louvain algorithm for directed networks [14]. We provided tailored visualization without commercial tools by identifying tasks through collaboration with medical researchers.

Prior work explored methods to visualize networks based on the data size (i.e., number of nodes and links) and the purpose of visualization. Komarek et al. [15] classified the layout method of placing nodes and links on a screen into seven types: force-directed, hive plot, adjacency matrix, arc diagram, Sankey diagram, chord diagram, and pivot graph. They compared the number of nodes and links suitable for network visualization. Vehlow et al. [16] conducted a survey and analyzed previous studies that visualized the group structure of graphs. According to the criteria from prior work [15], the number of nodes (i.e., about 1000 hospitals) and the number of links (i.e., about 2000 patient transfers) in our study were in the range for a force-directed layout. Nonetheless, we placed the nodes reflecting their actual locations on a map as it was necessary to support the analysis considering the geographical location of hospitals. Further, grouping was visualized through superimposition among the methods summarized in a previous study [16] to demonstrate the group structure derived through community detection. The map was designed to reduce the scope of analysis through various interactions. There were many overlapping nodes with many groups, which cannot be distinguished by color when visualizing them on the map.

Andrienko et al. [17] surveyed visual analytic research of movement and transportation systems. Patient transfers in this study are origin-destination (OD) travel data without start and end times. According to the survey, one commonly aggregates OD data into matrices or flows for visual analysis. While visualizing flows as lines on a map has been widely used [18], the following research worked on resolving visual clutters for more significant numbers of flow. For instance, Wood et al. [19] proposed an OD map in which each cell represents an OD vector as in an OD matrix. It preserved the spatial layout by introducing a spatial treemap, and reduced visual clutters caused by intersecting links. Andrienko et al. [20] worked on long-term flow data and proposed OD flow aggregation by direction and distance ranges. They used the aggregated flow to cluster time intervals and presented the flow with diagram maps. Von Landesberger et al. [21] also adopted spatial simplification of locations to avoid visual clutter in analyzing mass mobility data. Vrotsou et al. [22] proposed direction-based filtering to address the problem. They introduced an interactive visualization technique called Flowcube, which presents flows as three-dimensional arcs, and offered direction-based filtering to reduce the clutter. Yang et al. [23] proposed an interactive visual analytics system for epidemic control and introduced mouse hover interaction to focus on a specific region instead of visualizing the whole network. We adopted both approaches in visualizing patient transfers. As spatial information was crucial for the target users, we presented flows as lines on a map without any additional visual encodings, which might hinder the perception of spatial information. However, we provided interactive filters to narrow the analysis scope and dedicated views for force-directed node-link diagrams where each node represents aggregated locations (i.e., province).

3. Analysis of Data and User Task

As this study aims to propose a visual analysis framework for medical researchers interested in stroke patients, we collaborated with two researchers (one neurologist and one clinical statistician) from a quaternary hospital. We met monthly for three years, and they participated in data and task analysis. We independently searched for relevant prior work and reached a consensus on relevance in the monthly meetings. They shared insights over the literature review, provided domain knowledge to aid in determining network metrics, and summarized their analytic requirements. The overall framework from data and user task analysis results is summarized in Figure 1, and the following section describes the results.

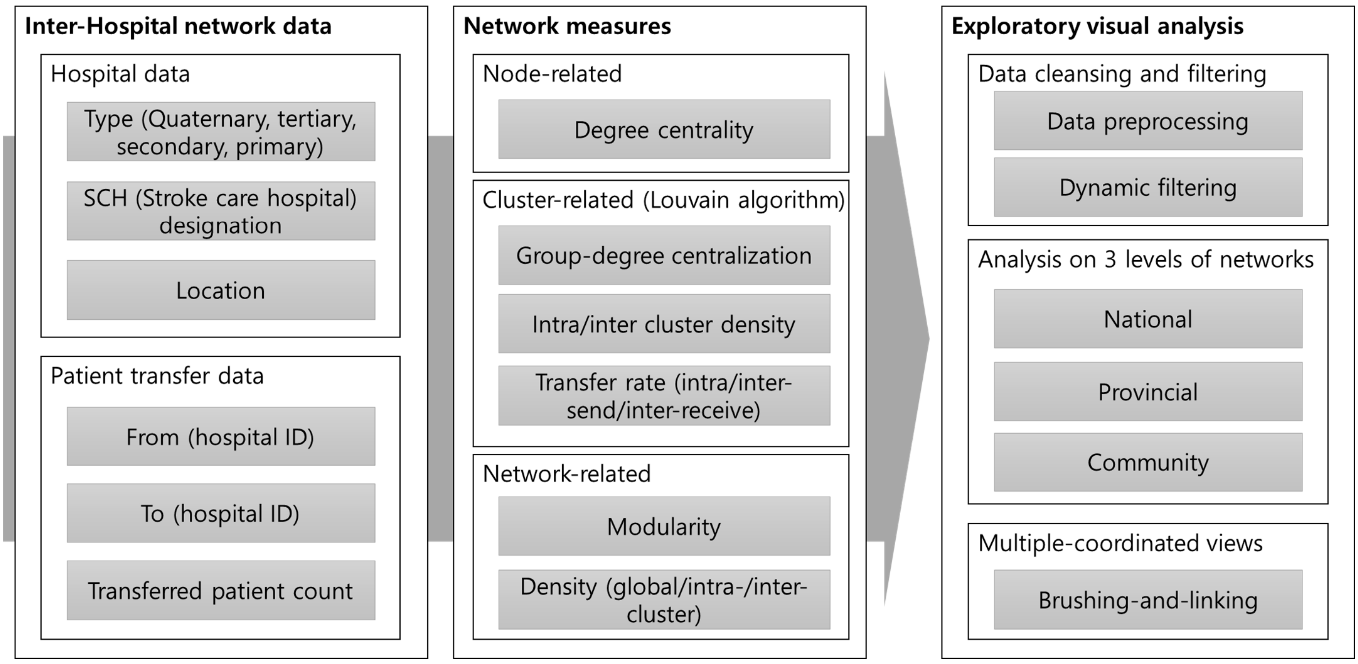


Figure 1. Exploratory visual analysis framework for stroke patients (PatientFlow).

3.1. Hospital and Patient Transfer Data

In the Republic of Korea, the HIRA (Health Insurance Review and Assessment service) accumulates and manages a list of hospitals registered as treatment institutions for specific diseases, as well as the data related to patient transfers between hospitals, their treatment processes, and outcomes. The data used by medical researchers who conducted the joint study consisted of the transfers of 19,113 stroke patients among 1009 hospitals from 1 July to 31 December 2016. The data used in this study are summarized in Table 1.

Table 1. Hospital and patient transfer data.

Category	Attribute	Description	Notes
Hospital	HID	Unique ID for hospital	-
	Type	Type of hospital	Quaternary hospital, tertiary hospital, secondary hospital, and primary hospital
	Address	Address of hospital	City, county, and district information
	Coordinate	Location of hospital	GPS coordinates calculated based on the government office
	SCH	Stroke care hospital designation status	-
Patient transfer	Type	Type of patient transfer	"Stay" or "transfer and stay"
	HID (from)	Hospital ID from which the patient is transferred	-
	HID (to)	Hospital ID to which the patient is transferred	-
	Patient count	Number of patients transferred	Total number of patients transferred within the aggregation period

Some data needed to be preprocessed before analysis, and parts of the provided data were anonymized to prevent the hospital's or personal information from being identified. For example, the detailed address information that can specify a hospital in the hospital list data is not provided, and only the regional administrative district (e.g., cities and provinces) and second-level administrative district (e.g., cities, counties, or districts) information are provided. The hospital ID is also randomly assigned to prevent the name of the hospital from being identified. Accordingly, the GPS coordinates of the government office (e.g., city hall, county office, or ward office) of the relevant administrative districts are used instead of the actual location of the hospital to display the location of the hospital on a map. In this case, random jittering is applied in the process of rendering on the screen to distinguish them on a map because multiple hospitals within the same administrative district can overlap. The distance between hospitals is calculated as a straight-line distance based on the GPS coordinates before jittering is applied.

Other data about the hospital include classification based on the size of the hospital (e.g., A = quaternary hospital, B = tertiary hospital, C = secondary hospital, and D = primary hospital). One can also classify the hospital based on the stroke care hospital (SCH) designation status. An SCH is a hospital equipped with healthcare professionals and an environment that can provide intensive treatment, including EVT, for stroke patients; the SCH designation status is considered important in the network analysis process in previous related studies. Based on the data received from the HIRA, all quaternary hospitals are designated as SCHs, and some tertiary hospitals are designated as SCHs. None of the secondary or primary hospitals are designated as SCHs. Accordingly, the interface is configured to facilitate the classification of B-type hospitals depending on their SCH designation status in the analysis process (Section 4.4).

We complied with restrictions on the confidentiality of data: patient transfer data only provides the sum of the number of patients transferred between hospitals to prevent the identification of individual patients. The data contains the ID number of each hospital and the number of patients transferred in and out. Among the types of patient transfer, those discharged after receiving treatment at the first hospital visited (stay), and those transferred to another hospital without receiving treatment at the first hospital visited and discharged after receiving treatment at the transferred hospital (transfer and stay) are examined. In practice, there are other cases where patients receive treatment at the first hospital and are discharged from the transferred hospital. The last case involves three hospitals where patients are transferred from the first hospital to the second hospital to receive treatment and then transferred to a third hospital to get discharged. As such cases account for about

0.43% of the total number of patients, this study only examined the two transfer types that accounted for the most (99.57%).

3.2. Data Preprocessing

In the visualization process, additional measures for the network are calculated. The calculated measures are mostly related to degree centrality and density, which were also used in a previous study [7]. However, while this study examined a nationwide network, the degree centrality of a hospital was not calculated for entire hospitals across the country. Instead, we first derive clusters through community detection, and then we measure the degree-centrality of a hospital within the belonging cluster. The community detection technique used in this study is the Louvain algorithm [12,14,24]. We used a variation of the algorithm [14] that reflects all characteristics of a directed (i.e., transfer direction) and weighted (i.e., the number of transferred patients) network. As a result, the measures for each node, each cluster, and the network are calculated (Table 2).

Table 2. Network measures for node, cluster, and entire network.

Category	Metric	Equation	Application	
Node	Degree centrality (in-degree)	$C_D(n_i) = \frac{n_i^{in}}{g-1}$	Find hub hospitals	(1) *
Cluster	Group-degree -centralization	$C_{GD} = \frac{\sum_{i=1}^{g}[C_D(n^*)-C_D(n_i)]}{max\sum_{i=1}^{g}[C_D(n^*)-C_D(n_i)]} = \frac{\sum_{i=1}^{g}[C_D(n^*)-C_D(n_i)]}{[(g-1)(g-1)]}$	Find highly centralized clusters	(2) †
	Intra-cluster density	$K_{intra} = \begin{cases} \frac{\lvert E_{ii}\rvert}{0.5\times g_i\times(g_i-1)}, & \text{where } g_i \geq 2 \\ 0, & otherwise \end{cases}$	Compare internal connectivity of clusters	(3) ‡
	Inter-cluster density	$K_{inter} = \frac{1}{l-1}\sum_{j=1,j\neq i}^{l}\frac{\lvert E_{ij}\rvert}{g_i\times g_j}$	Compare external dependency of clusters	(4) ¶
	Transfer rate (intra-cluster)	$T_{intra} = \frac{p_{intra}}{p_{intra}+p_{rcv}+p_{send}}$	Identify independent clusters	(5) §
	Transfer rate (inter-cluster receive)	$T_{rcv} = \frac{p_{rcv}}{p_{intra}+p_{rcv}+p_{send}}$	Identify influential clusters	(6) §
	Transfer rate (inter-cluster send)	$T_{send} = \frac{p_{send}}{p_{intra}+p_{rcv}+p_{send}}$	Identify highly dependent clusters	(7) §
Network	Modularity	$\frac{1}{m}\sum_{ij}\left[A_{ij} - \frac{n_i^{in}n_j^{out}}{m}\right]\delta\left(c_{n_i}, c_{n_j}\right)$, where $\delta(i,j) = \begin{cases} 1, i = j \\ 0, i \neq j \end{cases}$	Probe the strength of the division of a network	(8) ‖
	Global density	$\overline{K} = \frac{\lvert E\rvert}{0.5\times N(N-1)}$	Probe the overall density of a network	(9) #
	Mean intra-cluster density	$\overline{K}_{intra} = \frac{1}{l}\sum_{i=1}^{l}\frac{\lvert E_{ii}\rvert}{0.5\times g_i(g_i-1)}$	Compare the value with each intra-cluster density	(10)
	Mean inter-cluster density	$\overline{K}_{inter} = \frac{1}{0.5\times l(l-1)}\sum_{i=1}^{l}\sum_{j=i+1}^{l}\frac{\lvert E_{ij}\rvert}{g_i\times g_j}$	Compare the value with each inter-cluster density	(11)

* n_i = ith node in a cluster, n_i^{in} = in-degree of node n_i, n_i^{out} = out-degree of node n_i; † g = number of nodes in the cluster, $C_D(n^*)$ = largest value of $C_D(n_i)$ in the cluster, $max\sum_{i=1}^{g}[C_D(n^*) - C_D(n_i)]$ = maximum possible sum of differences in the cluster; ‡ $\lvert E_{ij}\rvert$ = number of undirected edges from cluster i to j, g_i = number of nodes in the ith cluster; ¶ l = number of clusters in the network, p_{intra} = number of patients transferred within a cluster; § p_{inter} = number of patients transferred from other clusters, p_{send} = number of patients transferred to other clusters; ‖ m = number of edges in the network, A_{ij} = weight between the nodes n_i and n_j, C_{n_i} = cluster to which node n_i belongs; # $\lvert E\rvert$ = number of undirected edges, N = number of nodes in the network.

The degree centrality of the ith node (n_i) in a cluster is the number of hospitals that sent patients to the node (i.e., in-degree) out of the total number of hospitals within the cluster, excluding the respective node $(g-1)$. Among three options (i.e., in-degree, out-degree, total degree) to calculate the degree centrality, the use of the in-degree in the calculation was determined after a review of network measures and confirmed by the collaborating medical researchers who also participated in user task analysis (Section 3.4). They wanted to focus on the role of the hub hospital in receiving and treating patients from nearby spoke hospitals. We tested various thresholds of degree centrality to determine a hub hospital. In this study, heuristically, a node with a value of 0.5 or higher is classified as a hub node. The collaborating researchers used our tool to analyze 93 clusters from Louvain clustering and found that 86 groups had a node located in the cluster's center with a degree centrality higher than 0.5. Three groups only had one node, and the remaining four had a central node with a centrality between 0.3 and 0.5. As a result, the collaborators determined the threshold as 0.5 in this research.

Measures calculated for each cluster include group degree centralization, intra-cluster density, and inter-cluster density, which are used for community analysis of networks in previous studies [25,26]. Regarding group-degree-centralization, the previous study calculated the value for the undirected network. Thus, we adjusted the denominator from $[(g-1)(g-2)]$ to $[(g-1)(g-1)]$ to accommodate the characteristics of the directed network (i.e., minimum in-degree is 0). Furthermore, three types of patient transfer rates were calculated for each cluster based on feedback from medical researchers. The proportion of patients transferred within each cluster, the proportion of patients transferred out to other clusters, and the proportion of patients transferred in from other clusters are calculated using the number of patients who visited the hospital in each cluster at least once (i.e., excluding stay patients) as the denominator.

Indices calculated for the entire network included modularity [14], which is calculated for community detection, and three densities (i.e., global, mean intra-cluster, and mean inter-cluster), which are used to evaluate community detection results in a previous study [26]. In this study, the adequacy of community detection is examined by determining if three density values satisfied $\overline{K}_{inter} < \overline{K} < \overline{K}_{intra}$.

3.3. Simulation Data Generation

The data handled in this study contain sensitive information even after thorough anonymization, and it was impossible to export them outside the designated institution for the study. Therefore, we prepared simulation data for development and verification. Based on the statistical representative values of the actual data, simulation data are generated with a similar number of hospitals per administrative district. For inter-hospital connections, sending and receiving hospitals were randomly selected based on the regional hospital distribution calculated from the number of hospitals located in each region. In addition, the number of transferred patients was randomly drawn from the patient count distribution of inter-hospital transfers (i.e., for randomly selected sending/receiving hospitals, we assigned randomly selected patient counts following the distribution of the actual data). These simulation data were used to validate feasibility and capability of the proposed visualization in the development process.

3.4. User Task Analysis

Through monthly meetings with the two collaborating medical researchers and a literature review, analyses performed in previous studies and additionally demanded analyses were identified. Two information visualization researchers and a neurologist searched for network analysis papers on stroke patients and cross-checked the results. Then, we summarized the task from the resulting list of articles. At the research meeting, we prepared statistics regarding network topology and discussed them along with clinical outcomes prepared by the medical researchers. While examining the proposed prototype of the analysis framework, we reviewed research tasks and collected feedback as they performed the task using the prototype. The tasks that influenced the design of the proposed framework are as follows:

3.4.1. Multi-Perspective Analysis of Topology (T1)

The medical researchers wanted to analyze the topology of how networks were formed. This can be divided into two categories: analyzing quantitative measures (e.g., centrality, density) that represent topology, and inspecting a node-link-based visualization for visual analysis of topology. The latter approach of analyzing through visualization was divided into a method of placing nodes on a map based on GPS coordinates, and a method of placing nodes using a force-directed layout, depending on whether geospatial information was used when placing nodes in a two-dimensional space. For collaborating researchers, all three analysis methods were used because the results that could be drawn varied depending on the three methods.

3.4.2. Analysis at Multiple Scales (T2)

The medical researchers tried to proceed with the analysis by setting the scope of the topology analysis in various ways. Some previous studies analyzed patient transfer data on a national scale [8,9], and others limited the analysis scope to local regions [1,5]. In addition, the analysis was based on administrative regions, other cases chose community as a scope of analysis. Zachrison et al. showed the possibility of network analysis using the community-detection algorithm [7], and Dong et al. analyzed communities in the inter-hospital network [11]. The literature review indicated that insights derived based on the scope of each analysis were expected to be different. Therefore, we decided to aid analysis with all three scales (i.e., national, regional, and community). Section 3.5 describes the relevance and characteristics of three network analysis levels.

3.4.3. Analysis of Hub-and-Spoke Structure (T3)

The medical researchers wanted to analyze the hub-and-spoke structure at the community level. Prior work confirmed that the designation as an SCH was followed by forming a network with nearby hospitals. Thus, one of the goals of the analysis was to identify how such network formation occurred in the Republic of Korea. In a previous study [13], hub hospitals were designated prior to an analysis based on the size of hospitals. However, the researchers wanted to detect communities solely from patient transfers and analyze the resulting hub-and-stroke structures. They wondered if there was a discrepancy between the intended and actual network structures. Thus, we classified hub hospitals by network measures regardless of the SCH designation status or hospital size. Moreover, we tried to visualize the hub-and-spoke structure to support comparative analyses of communities.

3.4.4. Analysis of Interprovincial Transfers (T4)

The medical researchers have attempted to analyze patient transfer patterns at the provincial level. Since the large catchment areas align with provincial areas in the Republic of Korea, investigating interprovincial patient transfers could indirectly examine whether the catchment area was well established. For example, it is commonly known that some regions in the Republic of Korea are highly dependent on specific hospitals. There were even cases where patients transferred to such hospitals from other provinces. The collaborating researchers wanted to confirm such known patterns with our framework and find other notable patterns from patient transfer data.

3.4.5. Exploratory Analysis of Patient Counts (T5)

The medical researchers wanted to analyze the proportion of patients by transfer types (i.e., transfer in/out and stay) in each hospital. However, the initial statistics examination with the researchers revealed that the proportion of patients who stayed without transfer was high (83.71%). Thus, the number of patients transferred in or out could be relatively small for visual comparison when we include the number of stayed patients. Still, excluding them from the analysis was not preferred by the researchers as the numbers could provide a cue about how many patients a hospital could accommodate. As a result, we could conclude that, while the researchers do not have a hypothesis, they still want to investigate patient distribution by transfer types with exploratory analysis. Moreover, they also showed interest in investigating the distribution over the whole nation, provinces, and communities.

3.5. Three Network Analysis Levels

The task analysis result indicated that the analyses of the three levels of the network were conducted in prior work and also wanted by the target users. This section summarizes the relevance and characteristics of the networks as follows. The national-level analysis reveals overall patient transfer trends and the geographical distribution of hospitals. One could analyze network measures over the country and explore the geographical relationship of hospitals of interest. The community-level analysis could complement the holistic analysis by depicting hub-and-spoke structures in each cluster. It can help users examine the role of individual hospitals in the community. The provincial level shows network characteristics regarding administrative divisions. It reveals a relationship between a predefined group of hospitals.

While prior work barely looked into the design space, the users can also closely coordinate the three levels of the network. For instance, exploring at the community level could draw groups of interest (e.g., groups with low centrality), and one can locate them in the national-level view to investigate the geographical relationship. One can also start from the provincial level to find communities of interest and drill down to exceptional hospitals. For example, the aggregated transfers between provinces could aid in finding an excessive amount of inter-province transfers. Then, one can browse communities within the sending/receiving regions to find communities where spoke hospitals are in remote administrative divisions.

4. PatientFlow: The Visual Analytics Framework

We designed PatientFlow for medical researchers who intend to analyze the network structure based on patient transfer data. Multiple-coordinated views enable an exploratory analysis based on the data and task analysis results. This section describes the overview, the views for the cluster-level and province-level analyses, and the dynamic queries for selecting or excluding data to be analyzed. PatientFlow is a web-based application that uses D3.js [27] for visualization and the JavaScript library of bubble sets [28] for expressing cluster coverage on a map. We used Graphology [29] for the Louvain clustering and calculating network metrics. Regarding the task analysis result, we denoted the resolved task (i.e., T1, T2, T3, and T4) for each framework component.

4.1. Overview of Patient Transfer Network

The overview area shows the distribution of hospitals, connectivity between hospitals, and quantitative measures depicting the network structure. We used two layout methods to place the hospitals and visualize links between them: force-directed and geographic layout (T1). Since the nodes in the force-directed layout push the others away, a user can observe the community detection result without visual clutters from overlapping nodes [30]; however, it is difficult to reflect the geographical location. Geographic layout might result in overlapped nodes if there are multiple closely located hospitals. Thus, we employed both methods so one can toggle between the options by clicking the icons in the top left corner of the topology view (Figure 2b). In the map, the node's shape indicates the type of the hospital (i.e., circle for SCH and square for non-SCH). It aids analysis of the geographical distribution by hospital types. In such cases, one could hide the links between the nodes to minimize occlusion. On the contrary, one could only leave links in the view and hide node markers to focus on connection density.

Figure 2. PatientFlow interface. (**a**) A list of interactive charts presents network characteristics. One can filter or highlight items by brushing or pointing. Patient transfer distances less than 150 km are selected, and one can confirm the applied filter with chips on top of the screen. (**b**) The topology view presents hospitals and patient transfers between them on a map. (**c**) The bar chart depicts the number of patients for each hospital in three categories (i.e., transfer-in, transfer-out, and stay). (**d**) A list of clusters yielded from a community detection algorithm is presented in three visualizations (node-link diagram, chord diagram, and bar chart). Each is designed to support analyzing topology, patient transfer patterns, and the number of patients for each hospital. The total number of clusters and selected clusters are shown on top of the view (e.g., 5 out of 95 clusters are selected) (**e**) The comparison view supports quantitative comparison between the selected clusters. One can select a network measure from the list on the left to make a comparison. (**f**) The province view presents patient transfers between provincial-level regions with a directed node-link diagram and a diverging bar chart. Depending on the transfer direction, the number of patients is colored in blue or orange. A table and a bar chart on the bottom show the hub hospitals within the region. One can click on a hospital in any chart to locate it in the topology view.

For each cluster derived from the community-detection results, we used bubble sets [28] to show the assumed coverage. (T4) It could show the geographical size of the cluster more clearly than simply coloring the node markers and could also aid in estimating the catchment area. However, the patient's address or point of departure when visiting the hospital was missing from our data. Thus, it does not represent accurate catchment areas as in [2], where Adeoye et al. assessed the accessibility of hospitals by region and visualized it. Still, it is possible to indirectly identify regions not participating in the treatment network for stroke patients. Moreover, any patient transfer between distant hospitals can be confirmed immediately through more giant bubbles spanning several areas.

The total number of patients that passed through each hospital was another aspect that collaborating researchers wanted to examine besides topology (T5). For each hospital, there are patients transferred to another hospital (transfer-out), transferred from other hospitals (transfer-in), and received treatment without being transferred (stay). The researchers wanted to examine the distribution of all three kinds of patients. Therefore, we included a view showing the number of patients per hospital with a stacked bar chart under the topology view (Figure 2c). The color encoding is kept consistent throughout the system, with blue, orange, and red indicating transfer-in, transfer-out, and stay, respectively.

During data analysis, we found far more stay cases than the other types (i.e., transfer-in or transfer-out). As a result, the number of stay cases hindered the visual examination of patient distribution for transferred patients. Thus, we placed a legend at the top right corner, which the user could toggle the visibility of each patient type by clicking an item in the legend. We also enabled grouping bars by provinces and changing the order of hospitals by the number of patients: ascending/descending order of stay, transfer-in, or transfer-out patients. One can also zoom in to narrow the analysis scope by brushing the chart. As there are more than 1000 hospitals, one might have to drill down several times by consecutive brushings. When a user brushes several times, buttons appear next to the title on top, indicating each drill-down step. One can click one of them to roll back to that state. If one finds a hospital of interest after brushing and browsing, one can locate it in another view by clicking on the bar. A user can change the mouse interaction mode (i.e., clicking or brushing) by toggling the button in the bottom right corner.

Task analysis results indicated that the researchers demanded multi-perspective analysis with quantitative measures as well as analysis on topology and geographical distribution (T1). We placed several interactive charts on the left side of the overview area to comply with this demand (Figure 2a). With a histogram of link distance, a user can analyze the distribution and filter links by brushing the chart. A boxplot of link distance for each province depicts the distribution and outliers (Figure 3). Regarding inter-province transfers, we count them based on the origin province. We assumed that the origin of long-distance transfer would be more interesting than the destination, as it might indicate a poor local network at the origin province. For instance, KAW province has a comparably significant deviation. When one clicks the outlying circle, the corresponding link becomes highlighted in red (Figure 3). One can also perform a similar examination at the cluster level with a brushable bar chart. As some clusters only have a couple of hospitals as members, visualizing with a boxplot was not applicable in some cases. Thus, we introduced a bar chart showing each cluster's longest link distance (Figure 4a). We also prepared other bar charts depicting network measures of hub nodes and clusters (i.e., measures in Table 2). For example, there is a bar chart for degree centrality (Figure 4b), where the color strip on the right shows the value range for the primary hub (red), secondary hub (orange and yellow), and spoke hospital. It could aid in selecting each class of hospital. We also included a bar chart (Figure 2a) to visualize the number of hospitals by size and SCH designation status (e.g., A-SCH, B-SCH, B, C, and D).

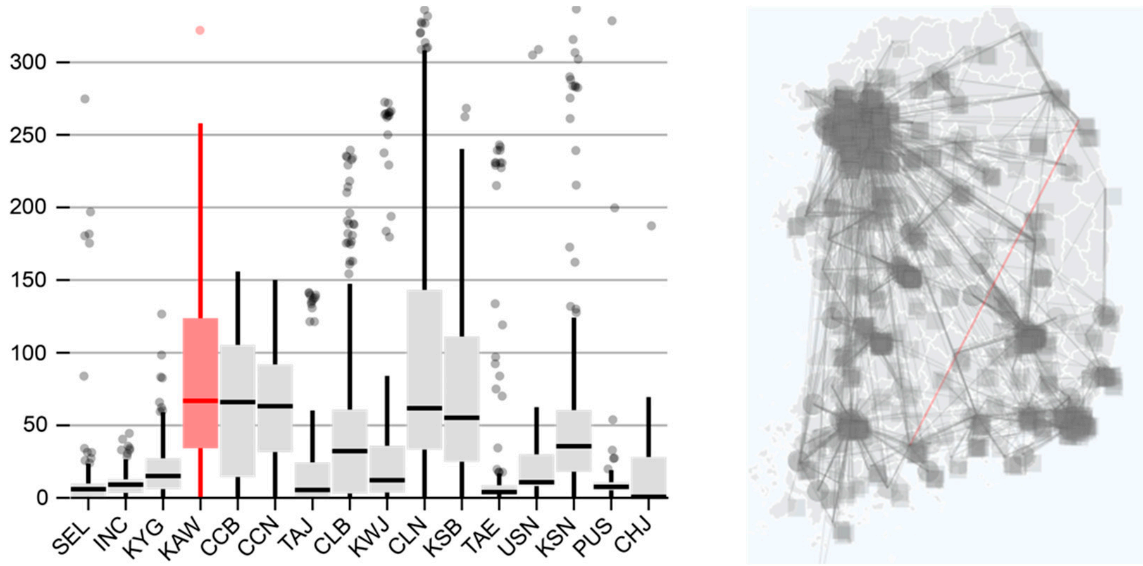
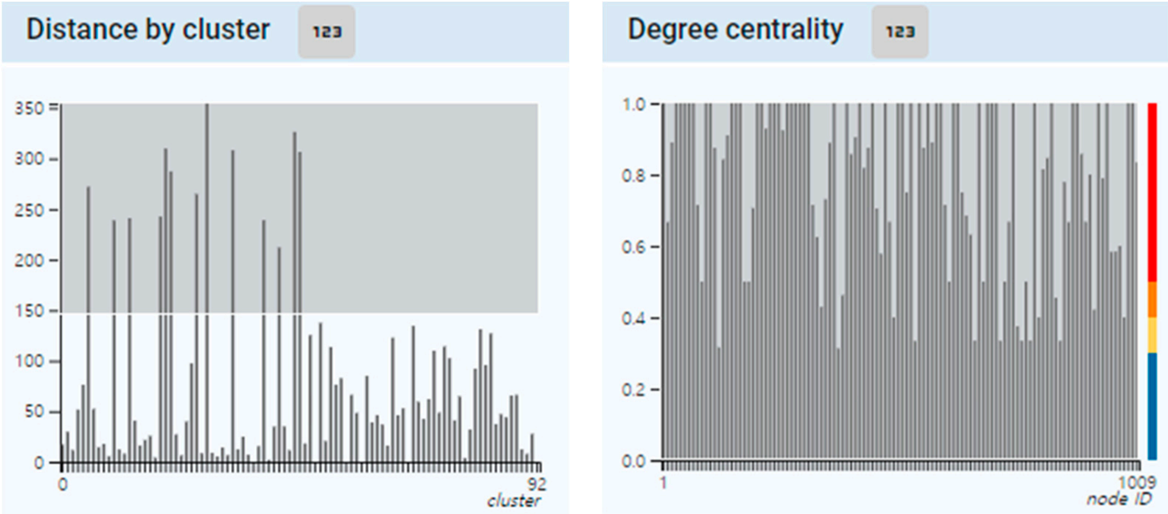


Figure 3. Box plot of the patient transfer distance aggregated by province. One can click the outlier dots in the box plot to highlight the transfer link in the map.



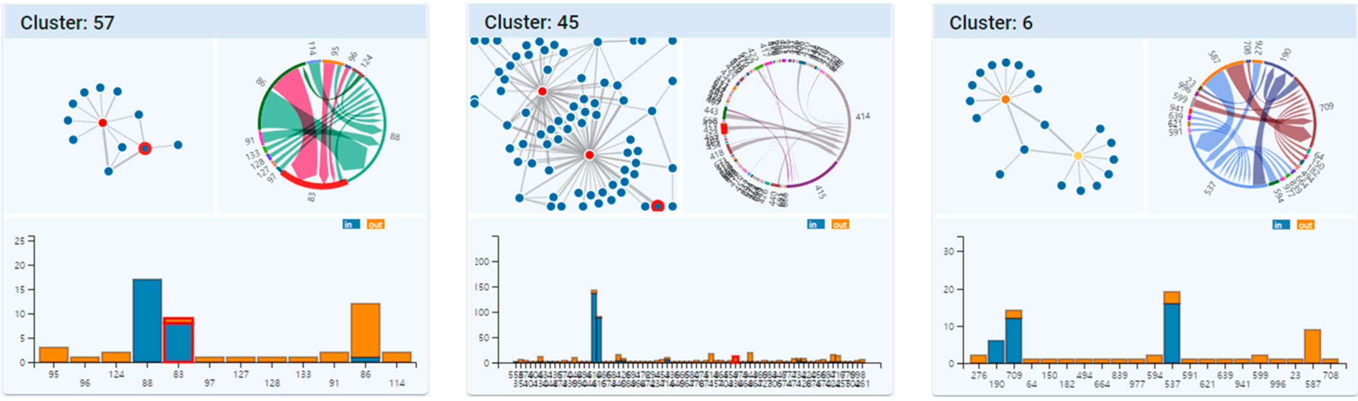
(**a**) (**b**)

Figure 4. Brushable bar charts for (**a**) the longest link within each cluster and (**b**) the degree centrality of each hospital. Darker gray background indicates the brushed area (i.e., selected range). The color strip on the right side of the degree centrality chart shows the value range for a primary hub (red), a secondary hub (orange and yellow), and a spoke hospital (blue).

4.2. Comparative Cluster Topology Analysis

Based on the task analysis results, we designed the cluster view with three types of visualizations for analyzing hub-and-spoke structures at the community level (T3). The node-link diagram at the top left uses a force-directed layout to depict the cluster's topology. We colored the node by degree centrality: red, orange, yellow, and blue correspond to values greater than 0.5, 0.4, 0.3, and 0, respectively (Figure 5). These values are determined heuristically during the meeting with the researchers. We classified a node with a degree centrality greater than 0.5 as a hub node and a node with a degree centrality less than 0.3 as a spoke node. Nodes with a centrality between 0.3 and 0.5 are assumed to be secondary hubs. We further divided the secondary hub range into two parts for detailed classification, using 0.4 as a threshold. As a result, one can distinguish a node receiving many patients from other hospitals within the cluster from a node receiving a relatively small number of patients while sending many patients to other hospitals.

The chord diagram on the upper right (T5) enables a more detailed examination of patient transfer counts than the node-link diagram (Figure 5). While the node-link diagram provides a better topology overview, one can hardly compare the number of patients transferred between the hospitals. On the contrary, the thickness of the arcs in the chord diagram corresponds to the patient count, and one can visually compare them. Each circumference segment represents the hospital with a distinct background color. The arcs in the diagram have identical colors to the target hospital, and the sharp arrowhead represents the transfer direction. Thus, one could quickly identify the hospital with a considerable number of visited patients and the hospital receiving the most significant number of patients. The chord diagram can display stay patients with self-links, but since most patients stayed at the same hospital, it hindered analyzing transfer patterns. As a remedy to a relatively small number of transferred patients, we put a button at the title bar to toggle whether to include stay patients or not. The bar chart at the bottom shows the exact number of patients and follows the toggled state.



(**a**) (**b**) (**c**)

Figure 5. Analysis of cluster-level patient transfers. One can click on circle markers in the node-link diagram, circumference segments in the chord diagram, or bars in the bar chart. The clicked element becomes highlighted with a red outline in all three charts: (**a**) hospital ID 83 and (**b**) hospital ID 434. (**c**) The background color of circle markers corresponds to degree-centrality: red, orange, yellow, and blue correspond to values greater than 0.5, 0.4, 0.3, and zero, respectively. One can locate secondary hubs colored orange and yellow.

Then, we arranged additional space for the comparative analysis of clusters (T1). One can browse a list of clusters by scrolling but examining 90 clusters derived from 1009 hospitals could be challenging. After several reviews on network measures with medical researchers, we selected group degree centralization as a sorting criterion: one could start the analysis from the most/least centralized cluster. In addition, one can click the title bar on top of each cluster to select it, and the selected clusters move to the top of the list for future reference. The selected clusters also appear in the comparison view located at the bottom for comparison of network measures, enabling quantitative analysis of network topology. A user can select a measure from the list on the left (Figure 2e), and a bar chart shows the corresponding value of selected clusters. Moreover, one can select multiple measures to combine them into a normalized stacked bar chart (Section 5.1). It could aid the comparison of multiple proportions where their sum equals one (e.g., three types of transfer rates in Table 2).

We aimed to encourage multi-perspective analysis by locating and analyzing a hospital of interest from another view (T1 and T2). Thus, we linked the charts within the cluster analysis area and the other charts from different perspectives. One can click on circle markers in the node-link diagram, circumference segments in the chord diagram, or bars in the bar chart. Then, the clicked hospital becomes highlighted in all three charts and other linked visualizations. For instance, when one clicks the circle marker in the node-link diagram, the markers representing the hospital are highlighted with a red outline in all three charts in the cluster view and the map in the topology view. Further, it automatically scrolls to the province it belongs to in the province view (Section 4.3) on the right side.

4.3. Provincial-Level Analysis

The task analysis revealed a demand for analysis at the provincial level in addition to the community level. Nekkab et al. [9] analyzed transfer patterns at the regional and county level, and we took a step forward by adopting multiple interactive coordinated views. We aggregated transfer patterns by administrative district units and listed hub hospitals in the area to support this analysis requirement (T4). The top two visualizations display patient transfers between provinces (Figure 6a). In the node-link diagram on the left, the province of interest lies in the center, and other provinces that sent/received patients to/from the province encircle it. Unlike the node-link diagram in the cluster view, two color-coded curved arrows (i.e., blue: transfer-in, orange: transfer-out) indicate the transfer directions. Since the number of entities could lead to visual clutter in the cluster view, we tried to minimize visual encodings such as directions. However, the total number of provincial-level divisions in the Republic of Korea is less than 20, and linked areas by patient transfers could be even fewer. Thus, we introduced bidirectional links and mapped the number of transferred patients to the thickness of the link. Although the bidirectional node-link diagram provides an overview while browsing multiple provinces, drawing the number of transferred patients could be challenging. An additional diverging bar chart on the right mitigates the problem. Moreover, one could compare the number of patients across multiple provinces with the chart. We used a consistent color scheme to display the transfer direction of patients (i.e., blue and orange).

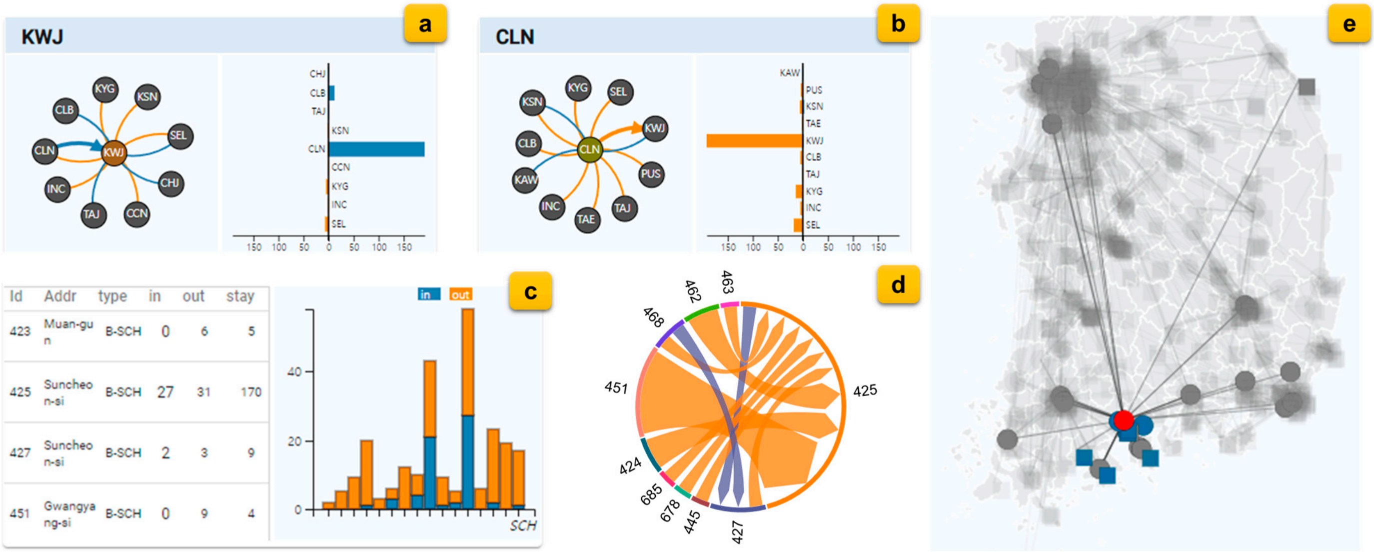


Figure 6. Analysis of provincial-level patient transfers. (**a**) The province of interest (i.e., KWJ) lies in the center, and other linked provinces encircle it. One can notice that KWJ receives many patients from CLN province, judging from a thick link on the left and the longest bar on the right. (**b**) As for CLN province, it mostly sends patients to others except for three provinces. (**c**) The table on the left shows the complete list of hubs in the area. One can identify the hub with the most significant number of patients in the stacked bar chart on the right (i.e., hospital ID 425). When one clicks the row or bar, the cluster with the hub becomes highlighted, and one can locate them in the cluster view. (**d**) The chord diagram shows that hospital ID 425 mostly receives patients from hospitals within the cluster but sends patients to hospitals outside the provinces judging from the bar chart in (**c**). (**e**) One can confirm it in the topology view by selecting the cluster and hovering a mouse on the hub. There are connected dark dray nodes outside the province.

There are two visualizations in the bottom row to display information on SCHs in the province (Figure 6c). The table on the left shows a complete list of SCHs. The columns contain a district, hospital type, and the number of patients who transferred in/out and stayed. Since tabular visualization is challenging to observe the overall trend, we put a stacked bar chart on the right that shows the three types of patient counts. We linked the two visualizations with the others (T2) and clicking a specific hospital in the table or the stacked bar chart highlights it in every linked visualization (e.g., topology view and cluster view). Clicking the column title sorts the items in the table and the stacked bar chart. As in the cluster view, we added a button on top to toggle whether to include stayed patients or not in the visualizations.

4.4. Dynamic Query and Community Detection

In PatientFlow, a user can filter hospitals and links in the interactive charts on the leftmost side of the screen (T1, T2, and T4). The task analysis result indicated that the collaborating researchers demanded exploratory analysis. Thus, we introduced several filters for links, hospitals, and clusters and placed chips at the top of the screen to check the currently applied filtering conditions. It could also aid in fine-tuning the conditions.

As a starter, one can filter links within a specific range by brushing a bar chart (Figure 2a). Zachrison et al. [1] assumed that patient transfers over an excessively long distance are uncommon because stroke is a time-sensitive disease. Therefore, prior work excluded the outliers after inspecting the distribution of travel distances. However, such a long-distance transfer could be meaningful for identifying defective communities. There are also no clear standard threshold values for filtering: users might have to apply various thresholds. With our tool, one can interactively adjust the range of transfer distances depending on the purpose of analysis. We added a brushable chart at the bottom of the histogram to provide an enlarged view of the selected range on top. Upon selection, the Louvain algorithm derives new communities with filtered links.

There are also filters for hospitals. Regarding the bar chart for hospital type, one can click on each bar to toggle selection. The bar for selected types becomess highlighted with red outlines. As the filtering affects every visualization in our tool, only the selected types of hospitals appear in other views. Thus, it could aid in analyzing the distribution of specific types of hospitals. For instance, if only the A-SCHs are selected, the A-SCHs and only the patient transfers between them become visible in the topology view. Users can click the apply button to update the community detection result with filtered hospitals (Figure 2a). Since the interaction includes or excludes hospitals in the network, it took relatively longer to come up with clustering results. Thus, we added a button on top to execute the clustering algorithm instead of updating the community in real time. It is also possible to filter hospitals by centrality. One can select only hub or spoke hospitals by brushing on the bar chart (Figure 4b). However, it does not derive new communities since it aims to find a hospital satisfying a specific condition without updating the clustered result.

Lastly, one can interactively filter clusters. One can find clusters with solid or weak hub-and-spoke structures by brushing a bar chart for group degree centralization. In the bar chart of inter-cluster density, users can identify highly interconnected clusters. Moreover, brushing the bar chart for the longest transfer distance within the cluster could reveal abnormally large clusters. We also considered deriving a convex hull or a concave hull based on the GPS coordinates of hospitals as an alternative. However, we took a conservative approach using the longest link distance and avoiding estimation, which might lead to a false perception of cluster coverage.

5. Evaluation

5.1. Case Study

We conducted a case study with two neurologists and a clinical statistician to evaluate the effectiveness of the proposed visualization tool. Since these researchers intended to derive clinical insights such as EVT outcome and mortality rate based on the findings through this framework, they attempted to use the tool in their own space. Moreover, due to the COVID-19 pandemic, we had to conduct the study online. Thus, we deployed the tool to the experts and encouraged to use it for a couple of weeks and then interviewed them for an hour.

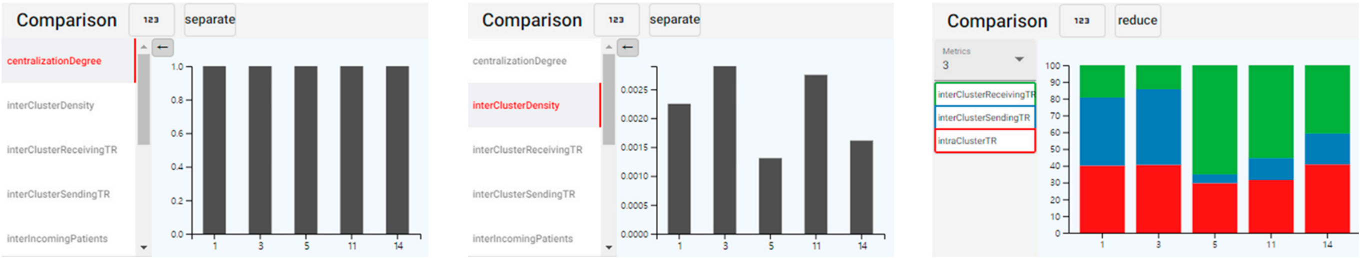
One of the common requests of the participants was verifying whether community detection was successful. They could validate the results by comparing the three density values in Table 2 and checking the modularity calculated from the Louvain algorithm. The validation result indicated that the communities were well derived. Then, they tried to examine the derived structure visually. At first, they used a node-link diagram, but thousands of nodes caused excessive visual clutters. Thus, they switched to a geographical layout to analyze distributions on a map and used the bubble sets to examine the community's coverage (Figure 2b). Since the bubbles were semi-transparent, the participants could notice that the east and southwest regions had a relatively small number of clusters. They found several outlying clusters with long-distance transfers by hovering a cursor on the bubbles. After a few filtering interactions in the distance bar chart (Figure 2b), they concluded that excluding links over 150 km of transfer distance would be appropriate for their dataset.

A priori, neurologists expected that patient transfer in a specific province would show different patterns compared to other regions. They tried to verify this widely known expectation. In the province view, they browsed patient transfers for each region and noticed singularities in some province pairs. Most provinces showed mutual patient transfers between them, even if the number of patients for each direction were somewhat biased toward one side. However, visual analysis with a node-link diagram and diverging bar chart identified some province pairs (e.g., CLN and KWJ) with dominant transfer directions (Figure 6a,b). In order to take a closer look, they clicked hospital ID 425 in the table (Figure 6c), which showed the highest number of patients among the SCHs in the CLN province. As they checked the belonging cluster, they found that the hospital received patients from the rest of the cluster, resulting in a degree centrality of 1.0. In the geographical map, they could find a reason for this discrepancy hospitals sent patients to the hospitals in neighboring provinces.

The case study participants observed the potential of applying a network analysis (i.e., analysis of community detection results and network measures) to the cases of hospitals in the Republic of Korea. The preliminary result from the case study led to a clinical research opportunity. In this section, we summarized additional use cases where PatientFlow could aid in finding notable patient transfer patterns. With limited access to medical data, we could not draw clinical implications from the following cases at the time of writing. Thus, we asked the case study participants to review them with their domain knowledge, and they regarded the cases as prominent patterns for future research.

In the bar chart of hospital types (Figure 2a), one could notice that all quaternary hospitals (i.e., A) are SCHs, about 75% of tertiary hospitals (i.e., B) are SCHs; and none of the secondary (i.e., C) and primary hospitals (i.e., D) are SCHs. One could assume that SCH is designated to lead the care of stroke patients among hospitals nearby. Thus, the user searched for uncommon hospitals with a degree centrality smaller than 0.5 in A-SCH hospitals. The user clicked the A-SCH bar in the bar chart of hospital types to filter out other types. Then, the user brushed the lower area in the degree centrality bar chart to examine hospitals with a degree centrality less than 0.5. Four hospitals in KAW, CCB, CCN, and CLN provinces matched the conditions. The user clicked the node in the geographical topology view, and he/she could notice that two of them were secondary hubs, but the rest were spokes in the cluster (highlighted with red outlines in Figure 5a,b). In the case of hospital ID 83, it showed a tendency to receive patients from the others within the cluster. However, the hub hospital in the cluster (i.e., hospital ID 88) received significantly more patients even though it was not an SCH (Figure 5a). Regarding hospital ID 434, it did not receive any patients from other hospitals. Instead, it sent most of the visited patients to one of the hub hospitals in the cluster, which is also A-SCH. Considering that the other 30 A-SCHs acted as a primary or a secondary hub, it might require further examination with additional clinical data.

During the case study, the participants used group degree centralization and the number of hubs in the cluster as criteria for classifying clusters. While such classification sufficed the demand, a user took a step forward by analyzing additional metrics for network topology. When a hospital receives patients from all the others in the cluster, group degree centralization becomes 1.0. Thus, the user selected five clusters (Figure 7a) with identical group degree centralization (i.e., 1.0) and star topology (i.e., cluster without inter-spoke links). Within the cluster, all five of them seemed similar to the user. Then, he/she selected inter-cluster density to investigate connectivity between clusters to find differences between the five clusters (Figure 7b). He/she wanted to investigate the cause and selected three transfer rates to analyze them in a stacked bar chart (Figure 7c). As a result, the user could identify two distinctive patterns: sending more patients to other clusters (e.g., cluster ID 1 and 3), and vice versa (e.g., cluster ID 5). When analyzing topology within a cluster, all five clusters would be similar. However, if the goal of the analysis is related to inter-cluster transfers, the user might need another perspective to investigate them properly.



(**a**) (**b**) (**c**)

Figure 7. Comparative cluster analysis view with multiple network measures. In this example, a user selected five clusters (i.e., cluster ID 1, 3, 5, 11, and 14). (**a**) The centralization degree for five clusters shows identical values. (**b**) However, inter-cluster densities vary among the five clusters. (**c**) Users can select multiple measures to analyze them with a stacked bar chart.

5.2. Comparative Analysis with Prior Work

A network analysis approach is in its early stage in clinical research after the appearance of EVT, and visualization emerged as an essential part of studying hospital coordination. Still, the prior work left a possibility to reach a higher level of maturity in contrast to the information visualization domain. To the best of our knowledge, this is the first paper to propose an interactive framework for the analysis of stroke care hospitals. Moreover, some research used commercial tools in the analysis process, which made quantitative comparison (e.g., task completion time, error rate) with our method difficult. Thus, we compared our work with previous research regarding data, analysis tasks, analyzed network levels, research focus, and visual representations (Table 3).

Prior work focused on analyzing patient transfer patterns (e.g., transfer volumes, the number of connections between hospitals, and the existence of community) and some related them to clinical variables (e.g., the number of performed treatments and clinical outcomes). Regarding the data, it ranged from a single state (or province) to a whole nation. The number of patients varied according to the target disease and collection period. The number of hospitals was primarily affected by the collection coverage. The research focus determined the analysis task. As a result, the previous studies partially conducted the analysis compared to ours. However, the case study participants generally agreed that the capability to perform additional tasks would benefit the prior work. Some studies investigated network changes over time, which this study did not cover; we only had access to patient transfer data within a single interval. We covered this limitation as a future research direction in the next section. Regarding the analyzed network levels, two previous works collected nationwide data, but only one examined the network in multiple scales. Even with the limited data coverage, some studies [1,5,11] tried to investigate networks at a community level.

Our study took a step forward from the prior work with multiple coordinated views. Most studies used node-link diagrams with geographic layouts to present spatial relationships between the hospitals and used only a few additional visualizations for analysis. Moreover, some studies [5,11,13] relied on external tools (e.g., ArcGIS, Gephi, and Google Maps) to show network structures. Thus, introducing and coordinating additional visualizations for multi-perspective analysis was challenging. PatientFlow can even apply data cleansing and filtering dynamically during a study. For instance, prior work [1] pointed out that link distances longer than 96 miles between hospitals can be an outlier since stroke requires immediate treatment. The user can interactively adjust the threshold and conduct a follow-up investigation immediately with PatientFlow.

Table 3. Comparison with prior work on hospital networks.

	Data (Country, Coverage and Period, Size, Target Disease)	Analysis Task [a]					Network Level [b]			Visualizations for Network Analysis
		T1	T2	T3	T4	T5	N	P	C	
Patient Flow	Republic of Korea Nationwide for 6-months 19,113 patients, 1009 hospitals Stroke	O	O	O	O	O	O	O	O	Node-link diagram (force-directed, geographic), Histogram, Chord diagram, Bar chart (diverging, stacked)
[8]	United States Nationwide for 2-years 12.5 M patients, 5667 hospitals Clostridium difficile	X	X	X	X	X	O	X	X	Node-link diagram (geographic)
[9]	France Nationwide for 1-year 21,279 patients [c], 1266 hospitals [c] Hospital-acquired infections	O	O	P	O	X	O	O	O	Parallel coordinates, Node-link diagram (geographic)
[11]	China Fujian province for 3-years 32,759 patients 1043 hospitals Hypertension	O	P	P	N	X	N	O	P	Node-link diagram (force-directed)
[5]	United States Eight states [d] for 5-years 154,631 patients, 394 hospitals Stroke	X	P	P	X	X	N	O	O	Node-link diagram (geographic)
[1]	United States California for 8-years 336,247 patients, 351 hospitals Stroke	X	P	X	N	X	N	O	P	Node-link diagram (geographic)
[13]	United States Three states [e] for 15-years 12,803 patients, 43 hospitals Stroke	O	N	O	X	X	N	X	O	Map, Line chart

[a] Refer to Section 3.4 for each task description. O = performed, X = not performed, P = partially performed, N = not applicable. [b] N = national, P = provincial, C = community. [c] HAI (Hospital-Acquired Infections)-specific network. [d] Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Rhode Island, and Vermont. [e] Massachusetts, New Hampshire, and Maine.

6. Conclusions

This study proposed a framework for deriving and analyzing a community from patient transfer data, geographical location, and hospital type (i.e., size and whether designated as SCH). The framework adopted the methodologies and suggested network measures in prior work [1,7]. Unlike cases with a clear hypothesis to be tested or the aim of making a predictive model for fixed variables, medical researchers performed an exploratory analysis to understand the topology at the time data were collected. We designed PatientFlow for a national-level overview and interactive analysis of inter-hospital coordination. Users can also examine patients' transfer patterns at the cluster and provincial-level administrative district units. To this end, PatientFlow supports visual and quantitative analyses with multiple-coordinated views.

Our approach fulfilled the requirement of the collaborating medical researchers. It led to clinical implications with additional medical data and inspired follow-up research. However, at the time of research, we could not access the clinical records of the medical research institution. Since the patient outcome at the hospital played an essential role in prior work [7], the researchers had to rely on a separate tool to analyze them with the findings from our tool. Regarding quantitative analyses, we are working on integrating the clinical records into our framework and adopting additional measures. For instance, additional network measures, such as eigen centrality or page rank, could be introduced to analyze the topology from another perspective. Moreover, additional data could introduce a subgroup analysis as in the prior work [9] on HAIs. We could divide patients into subgroups by gender, age, diabetes, or other meaningful factors for stroke treatment.

We are also working on adopting another community detection algorithm. The Louvain algorithm identified a total of 93 communities in this work. Even though we introduced interactivity (i.e., brushing-and-linking, zooming, and filtering) in our tool, communities located near the metropolitan area yielded visual clutter from occlusions. Jittering nodes resolved the problem to some extent, but overlapping bubble sets made it challenging to analyze the clusters in the topology view. Additional efforts on grouping adjacent clusters using district information and the current patient transfer information might help mitigate the occlusion problem. A visual comparison of multiple community detection results would also help determine the appropriate algorithm heuristically. Such an approach would lead to examining the changes in network structure over time, which has been a subject of interest in prior work [1,5]. It could also aid in comparing the derived communities to planned catchment areas for policy making.

One can use our research to analyze hospital networks for cases where interhospital transfers are inevitable. For instance, patient transfers are prevalent in the emergency department [31], but some studies reported challenges in arranging interhospital transfers [32,33]. The COVID-19 outbreak has also introduced the need for network analysis. Negative pressure rooms were proven effective for treatment [34], but the limited number of rooms in each hospital might cause patient transfers to other hospitals. In such cases, PatientFlow could aid clinical researchers in analyzing hospital networks and finding communities with a low level of cooperation.

Author Contributions: Conceptualization, K.K., J.P. and H.S.; methodology, K.K. and H.S.; software, K.K., J.P. and H.S.; validation, K.K., J.P. and H.S.; formal analysis, K.K. and H.S.; investigation, K.K. and H.S.; resources, H.S.; data curation, K.K., J.P. and H.S.; writing—original draft preparation, K.K. and H.S.; writing—review and editing, H.S.; visualization, K.K., J.P. and H.S.; supervision, H.S.; project administration, H.S.; funding acquisition, H.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was support by the Soongsil University Research Fund (New Professor Support Research) of 2020.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author with the permission of HIRA (Health Insurance Review and Assessment service). The data are not publicly available due to privacy.

Acknowledgments: We thank Hee-Joon Bae, Jihoon Kang, and Seong Eun Kim (Department of Neurology at Seoul National University Bundang Hospital) for their help and valuable comments on this research.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zachrison, K.S.; Hsia, R.Y.; Li, S.; Reeves, M.J.; Camargo, C.A.; Yan, Z.; Onnela, J.; Schwamm, L.H. Ischemic Stroke Systems of Care in California: Evolution in the Organization During the Mechanical Thrombectomy Era. *Stroke Vasc. Interv. Neurol.* **2022**, *2*, e000206. [CrossRef]
2. Adeoye, O.; Albright, K.C.; Carr, B.G.; Wolff, C.; Mullen, M.T.; Abruzzo, T.; Ringer, A.; Khatri, P.; Branas, C.; Kleindorfer, D. Geographic Access to Acute Stroke Care in the United States. *Stroke* **2014**, *45*, 3019–3024. [CrossRef] [PubMed]
3. Berkhemer, O.A.; Fransen, P.S.S.; Beumer, D.; van den Berg, L.A.; Lingsma, H.F.; Yoo, A.J.; Schonewille, W.J.; Vos, J.A.; Nederkoorn, P.J.; Wermer, M.J.H.; et al. A Randomized Trial of Intraarterial Treatment for Acute Ischemic Stroke. *N. Engl. J. Med.* **2015**, *372*, 11–20. [CrossRef]

4. Goyal, M.; Demchuk, A.M.; Menon, B.K.; Eesa, M.; Rempel, J.L.; Thornton, J.; Roy, D.; Jovin, T.G.; Willinsky, R.A.; Sapkota, B.L.; et al. Randomized Assessment of Rapid Endovascular Treatment of Ischemic Stroke. *N. Engl. J. Med.* **2015**, *372*, 1019–1030. [CrossRef] [PubMed]
5. Zachrison, K.S.; Onnela, J.-P.; Hernandez, A.; Reeves, M.J.; Camargo, C.A.; Cox, M.; Matsouaka, R.A.; Metlay, J.P.; Goldstein, J.N.; Schwamm, L.H. Ischemic Stroke Transfer Patterns in the Northeast United States. *J. Stroke Cerebrovasc. Dis.* **2019**, *28*, 295–304. [CrossRef] [PubMed]
6. Zachrison, K.S.; Dhand, A.; Schwamm, L.H.; Onnela, J.-P. A Network Approach to Stroke Systems of Care. *Circ Cardiovasc. Qual. Outcomes* **2019**, *12*, e005526. [CrossRef]
7. Lee, H.H. Significance and Limitations of the 2nd Basic Public Health Care Plan from the Perspective of Expanding Public Health Care. *Public Health Aff.* **2021**, *5*, e3. [CrossRef]
8. Fernández-Gracia, J.; Onnela, J.-P.; Barnett, M.L.; Eguíluz, V.M.; Christakis, N.A. Spread of Pathogens in the Patient Transfer Network of US Hospitals. In *Social, Cultural, and Behavioral Modeling*; Lee, D., Lin, Y.-R., Osgood, N., Thomson, R., Eds.; Springer International Publishing: Cham, Switzerland, 2017; Volume 10354, pp. 271–280.
9. Nekkab, N.; Astagneau, P.; Temime, L.; Crépey, P. Spread of Hospital-Acquired Infections: A Comparison of Healthcare Networks. *PLoS Comput. Biol.* **2017**, *13*, e1005666. [CrossRef]
10. Clauset, A. Finding Local Community Structure in Networks. *Phys. Rev. E* **2005**, *72*, 026132. [CrossRef]
11. Dong, X.; Tang, D.; Tang, C. Social Network Analysis in China's Hospital Healthcare. *Phys. A Stat. Mech. Appl.* **2021**, *565*, 125546. [CrossRef]
12. Blondel, V.D.; Guillaume, J.-L.; Lambiotte, R.; Lefebvre, E. Fast Unfolding of Communities in Large Networks. *J. Stat. Mech.* **2008**, *2008*, P10008. [CrossRef]
13. Sharma, R.; Zachrison, K.S.; Viswanathan, A.; Matiello, M.; Estrada, J.; Anderson, C.D.; Etherton, M.; Silverman, S.; Rost, N.S.; Feske, S.K.; et al. Trends in Telestroke Care Delivery: A 15-Year Experience of an Academic Hub and Its Network of Spokes. *Circ. Cardiovasc. Qual. Outcomes* **2020**, *13*, e005903. [CrossRef]
14. Dugué, N.; Perez, A. Direction Matters in Complex Networks: A Theoretical and Applied Study for Greedy Modularity Optimization. *Phys. A Stat. Mech. Appl.* **2022**, *603*, 127798. [CrossRef]
15. Komarek, A.; Pavlik, J.; Sobeslav, V. Network Visualization Survey. In *Computational Collective Intelligence*; Núñez, M., Nguyen, N.T., Camacho, D., Trawiński, B., Eds.; Springer International Publishing: Cham, Switzerland, 2015; Volume 9330, pp. 275–284.
16. Vehlow, C.; Beck, F.; Weiskopf, D. Visualizing Group Structures in Graphs: A Survey: Visualizing Group Structures in Graphs. *Comput. Graph. Forum* **2017**, *36*, 201–225. [CrossRef]
17. Andrienko, G.; Andrienko, N.; Chen, W.; Maciejewski, R.; Zhao, Y. Visual Analytics of Mobility and Transportation: State of the Art and Further Research Directions. *IEEE Trans. Intell. Transp. Syst.* **2017**, *18*, 2232–2249. [CrossRef]
18. Wood, J.; Slingsby, A.; Dykes, J. Visualizing the Dynamics of London's Bicycle-Hire Scheme. *Cartographica* **2011**, *46*, 239–251. [CrossRef]
19. Wood, J.; Dykes, J.; Slingsby, A. Visualisation of Origins, Destinations and Flows with OD Maps. *Cartogr. J.* **2010**, *47*, 117–129. [CrossRef]
20. Andrienko, G.; Andrienko, N.; Fuchs, G.; Wood, J. Revealing Patterns and Trends of Mass Mobility Through Spatial and Temporal Abstraction of Origin-Destination Movement Data. *IEEE Trans. Vis. Comput. Graph.* **2017**, *23*, 2120–2136. [CrossRef]
21. von Landesberger, T.; Brodkorb, F.; Roskosch, P.; Andrienko, N.; Andrienko, G.; Kerren, A. MobilityGraphs: Visual Analysis of Mass Mobility Dynamics via Spatio-Temporal Graphs and Clustering. *IEEE Trans. Vis. Comput. Graph.* **2016**, *22*, 11–20. [CrossRef]
22. Vrotsou, K.; Fuchs, G.; Andrienko, N.; Andrienko, G. An Interactive Approach for Exploration of Flows Through Direction-Based Filtering. *J. Geovisualization Spat. Anal.* **2017**, *1*, 1. [CrossRef]
23. Yang, C.; Zhang, Z.; Fan, Z.; Jiang, R.; Chen, Q.; Song, X.; Shibasaki, R. EpiMob: Interactive Visual Analytics of Citywide Human Mobility Restrictions for Epidemic Control. *IEEE Trans. Vis. Comput. Graph.* **2022**, *in press*. [CrossRef]
24. Newman, M.E.J. Modularity and Community Structure in Networks. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 8577–8582. [CrossRef] [PubMed]
25. Wasserman, S.; Faust, K. *Social Network Analysis: Methods and Applications*; Structural Analysis in the Social Sciences; Cambridge University Press: Cambridge, UK, 1994.
26. Miasnikof, P.; Shestopaloff, A.Y.; Bonner, A.J.; Lawryshyn, Y.; Pardalos, P.M. A Density-Based Statistical Analysis of Graph Clustering Algorithm Performance. *J. Complex Netw.* **2020**, *8*, cnaa012. [CrossRef]
27. Bostock, M.; Ogievetsky, V.; Heer, J. D^3 Data-Driven Documents. *IEEE Trans. Vis. Comput. Graph.* **2011**, *17*, 2301–2309. [CrossRef]
28. Collins, C.; Penn, G.; Carpendale, S. Bubble Sets: Revealing Set Relations with Isocontours over Existing Visualizations. *IEEE Trans. Visual. Comput. Graph.* **2009**, *15*, 1009–1016. [CrossRef] [PubMed]
29. Plique, G. Graphology, a Robust and Multipurpose Graph Object for JavaScript. *Zenodo* **2022**. [CrossRef]
30. Dkaki, T.; Mothe, J. An Energy-Based Model to Optimize Cluster Visualization. In Proceedings of the 2014 IEEE Eighth International Conference on Research Challenges in Information Science (RCIS), Marrakech, Morocco, 28–30 May 2014; pp. 1–11.
31. Dunn, M.J.G.; Gwinnutt, C.L.; Gray, A.J. Critical Care in the Emergency Department: Patient Transfer. *Emerg. Med. J.* **2007**, *24*, 40–44. [CrossRef]
32. Hains, I.M.; Marks, A.; Georgiou, A.; Westbrook, J.I. Non-Emergency Patient Transport: What Are the Quality and Safety Issues? A Systematic Review. *Int. J. Qual. Health Care* **2011**, *23*, 68–75. [CrossRef]

33. Craig, S.S. Challenges in Arranging Interhospital Transfers from a Small Regional Hospital: An Observational Study. *Emerg. Med. Australas.* **2005**, *17*, 124–131. [CrossRef]
34. Al-Benna, S. Negative Pressure Rooms and COVID-19. *J. Perioper. Pract.* **2021**, *31*, 18–23. [CrossRef]