Burkey, Mark L.; Bhadury, Joyendu; Eiselt, Horst A.; Toyoglu, H.

**Article**
The impact of hospital closures on geographical access: Evidence from four southeastern states of the United States

*Operations Research Perspectives*

**Provided in Cooperation with:**
Elsevier


This Version is available at:
http://hdl.handle.net/10419/178275

**Terms of use:**
Documents in EconStor may be saved and copied for your personal and scholarly purposes. You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

http://creativecommons.org/licenses/by-nc-nd/4.0/
The impact of hospital closures on geographical access: Evidence from four southeastern states of the United States

M.L. Burkeya, J. Bhaduryb, H.A. Eiselt°C, H. Toyoguldd

a School of Business and Economics, North Carolina A&T State University, Greensboro, NC, USA
b School of Business Administration and Economics, The College at Brockport, State University of New York, 350 Campus Drive, Brockport, New York 14420-2931, USA
c Faculty of Business Administration, University of New Brunswick, Fredericton, New Brunswick, Canada
d Sabre Airline Solutions, Southlake, TX, USA

A R T I C L E   I N F O
Article history:
Received 23 August 2016
Revised 16 February 2017
Accepted 23 March 2017
Available online 1 April 2017

Keywords:
Health care
Access to health care
Proximity
Hospital closures
Location problems
Facility planning

A B S T R A C T
This paper examines the effects of hospital closures on geographical access by potential patients, using data from four southeastern U.S. states. Using optimization models designed to minimize the adverse effects of hospital closures, extensive computations are performed and the results are discussed. The effects of the closures on the rural areas is also investigated. Finally, the paper determines which hospitals are most likely among those to be closed assuming that up to 10% of the existing hospitals in each of the four states were to be shut down. The overall conclusion of the empirical findings is that while differences exist among the states, efficiency, coverage, and equality measures for geographical access do not suffer significantly if only a few hospitals are closed in each state, provided these closures are done optimally to minimize impact. Further, for efficiency objectives, decision makers can follow a sequential strategy for closures and still be guaranteed optimality. The paper also discusses the effects of hospital closures on equity and it examines whether or not rural areas are disproportionately affected by closures.

© 2017 The Authors. Published by Elsevier Ltd.
This is an open access article under the CC BY-NC-ND license.
(http://creativecommons.org/licenses/by-nc-nd/4.0/)

1. Introduction

Hospital closures or plans to close hospitals can be found almost daily in newspapers around the globe. References referring to this occurrence in the United States [1,2], United Kingdom [3], Germany [4], Canada [5] are but a few examples of this trend. It is important to note however that this is not a new trend. In fact, a report [6] from the United Kingdom regarding the years 1979–86 shows a large number of hospitals being closed in those years in the United Kingdom. While there are exceptions, a common observation about such closures is that the rural and underpopulated areas tend to be among those hit hardest.

It is helpful to try and understand some of the major forces behind many hospital closures. The first of these has to do with cutbacks in public spending since governments provide substantially for the expenditures for health care in most countries, including in the United States, where two government-run programs, Medicare and Medicaid, cover a substantial portion of the US population. This expense is significant and growing in most countries, accounting for between 9%–13% of GDP in most developed countries (with the United States being at outlier at 17%), see [7]. As of the summer of 2015, seventeen countries have a public debt in excess of 100% of their respective gross domestic products, see [8]. It is apparent that in order to keep the burden of debt manageable, it is necessary to make adjustments or save money. Aside from a generally unpopular strategy to raise taxes in case of public health care or increasing fees in case of private care, the main way is to cut services of some sort: eliminating or decreasing state subsidies, cut services or require co-pays, shutting down programs, and similar courses of action. One of the possibilities includes the closure of hospitals. Whether or not hospitals are private or public, ownership (and hospital size) appear to makes a difference when considering the closure of hospitals, as noted by Noh et al. [9] in the case of Korea. Furthermore, consolidation among private health care companies often leads to closure of hospitals [10]. The second important reason for the closures or relocations has to do with the inability to attract physicians to rural areas, see e.g., [11] and [12]. Nelson [13] provides some arguments that are designed to counter this trend. Our model is designed to have a “central planner,” who
has the overall viability of the system in mind, rather than that of a subsystem such as an individual hospital or group of hospitals as is often the case in privately run systems.

Another reason for hospital closures is the reduction of demand. For instance in the early 1960s, an appendectomy required an average post-surgery hospital stay of 6.3 days [14], while advances in technology (laparoscopic appendectomy in this instance) have reduced this requirement to 1–2 days [15]. Finally, the fourth reason to cut funding or close hospitals is a result of the inefficiency of hospitals, see, e.g., [16] in the case of Germany.

While the closure of hospitals has been a well-known phenomenon for a while now, its impact on health care consumers has been less studied in the academic literature. Samuels et al. [17] note that among the hospitals closed in 1989 in the United States, two thirds were in rural areas. Updates on the closures of rural hospitals are reported in [18–21] and [22].

However, these papers above also point out that most closed hospitals were in close proximity of other, open, hospitals. This leads to the conclusion that most closures had no discernible effect on the quality of health care. Rosenbach and Dayhoff [23] examined hospital closures in eleven states that appear to represent a reasonable cross-section of the United States. The closures did not cause any changes in the mortality rates, and patients in affected areas were often rerouted to urban teaching hospitals. Liu et al. [24] report the effects of the closure of very small hospitals in rural Saskatchewan. They report that the extensive closures have had no negative effect on the quality of health care and the well-being of the population. Lindrooth et al. [25] investigate the effects of urban hospital closure on the operations and costs of the remaining hospitals. Buchmueller et al. [26] investigated hospital closures in Los Angeles County and used distances/travel time as their main criterion. It was determined that most hospitals that were actually closed were located in more affluent areas. As a matter of fact, the authors noted some improvements in the quality of health care after the closures of underperforming hospitals. In general, it was observed that closures shifted health care to more efficient hospitals and doctors’ practices. Finally, Lindbom [27] uses a very different angle on hospital closures: how do voters punish the ruling parties, if they remove services voters believe they are entitled to? A common thread of assumption in all of these works cited above is that they assess the impact of hospital closures that have already occurred. To the best of our knowledge, there are no references regarding the impact of hospital closures on geographical access to customers, under the assumption that these closures are thoughtfully planned to minimize the overall adverse impact on customers. That is the point of departure of this paper.

In an effort to partially address the literature gap noted above, this paper focuses on the problem of hospital closures and determines the impact of these closures on geographical access using data from 4 states in the Southeastern US. Optimization modeling is used in deciding which hospitals to close, so that the overall adverse impact on customers is minimized. Our primary motivation for this work is that since they have been sited in the past based on needs at that point in time, hospital services at a particular location may no longer be in need, based on changing demographics, shifting settlement patterns, different types of treatment, and many others. In other words, the closure of hospitals can and should be rationalized by considering the effects of the closure on various constituents, provided the selection of hospitals to close is done using analytical modeling to minimize the impact of the closures.

In order to conduct our study, we will need to introduce and discuss a number of potential criteria to be discussed in this context, and, just as importantly, devise quantitative measures to operationalize their use.

As far as efficiency is concerned, we can distinguish between individual efficiency, which can be measured in terms of the average access time a potential patient has to a hospital, and collective efficiency, which measures the proportion of the population that is within a given distance or time from a hospital. Another issue that is somewhat more contentious and difficult to quantify is the concept of “equity.” Following the literature, we will measure equality by the Gini index [28] which is the normalized area between the Lorenz curve [29] and the line of complete equality. While the Gini index has long been used as a criterion as an objective (for surveys, see, e.g., [30] and [31]), while for a more recent formulation and optimization, see [32], problems have been reported (see [33] and [34]) regarding the measure’s tendency to choose highly inefficient solutions. Finally, another important consideration in hospital closures concerns the rural – urban divide. More specifically, it investigates whether or not the hospital closures inordinately affect the rural population; see, e.g., [2]. In order to do so, we use a well-accepted measurement of the degree of rurality provided by the United States Department of Agriculture [35]. It will be elaborated upon below.

The main research questions in this contribution are then as follows:

- how do the recommended hospital closures obtained using optimization schemes compare to actual closures that have occurred in the past?
- As hospital closures are bound to occur over time, are the optimized closures stable? In other words, we want to enquire whether or not short-term optimal decisions will also result in long-term optima.
- how do the recommended closures affect equality of service?
- do the recommended closures inordinately affect the rural or the urban population?

Before we proceed, it is important to discuss some of the fundamental assumptions of our study and the consequent caution that must be exercised in interpreting the results. The models we are using assume that there is a central planning authority, such as a Department of Health that decides how many and which hospitals to close. In the area under consideration in this paper, viz., four States in the Southeastern United States, there is a mix of publicly and privately owned hospitals, whose closures are decided in decentralized fashion. However, our assumption that closures were centrally organized and optimized provides a lower bound on the damage in terms of coverage and average access time or distance that is done to customers. This is the context in which our computational results should be interpreted. Furthermore, closures may not necessarily be viewed as recommendations, as closures may happen for reasons other than the ones we use. While the closures recommended by our model are based on patient-centered criteria, viz., coverage and average access time, actual (real) closures may be based on very different criteria. In case of private facilities, profitability is the main criterion. There is a connection to accessibility and coverage—if few people live within a reasonable distance from the hospital, it will not be profitable and it is likely to face closure. Similarly, public facilities may face closures for reasons such as

- Demographic changes that eliminate the need for a hospital at that site,
- Changing patient needs have resulted in a hospital having excess capacity, rendering it unprofitable,
- The state of the hospital, i.e., the structure may be dilapidated to a point, at which it is too expensive to renovate/retrofit given new regulations,
- The capacity of the present hospital may not be sufficient and the site may not allow for appropriate expansions.
• The value of the facility, e.g., the present hospital may be located on very valuable land that can be sold for profit and a new hospital may be built elsewhere.

Similarly, construction of new hospitals will depend on issues such as budget considerations, an issue that is determined by politicians maybe a year before the plan is set up, which typically involves political haggling, and other imponderables that are difficult, if not impossible, to analytically track in a location model. Nonetheless, a surprising result in [36] is that this unknowable location process results in a location pattern that is very similar to a process that uses coverage and access time/distance as criteria, thus lending credibility to the analytical modeling process we have used in this paper.

The remainder of the paper is organized as follows. Section 2 discusses the measures used for three different aspects of the hospital closure problem and describes mathematical models to determine the optimal choice of hospitals to close. Section 3 discusses the data to be used in our analysis, Section 4 presents the results of a computational study involving four American states, and Section 5 reviews the main results of this contribution and provides some thoughts for future research.

2. The mathematical models

The application of operations research techniques to hospital location problems dates back quite a few decades. In the mid-1960s, Hakimi [37] provided the foundations to modern location theory in his seminal papers by proving the “node property” for p-median problems on networks (i.e., formulations that locate p facilities so as to minimize the sum of distances between customers and their closest facilities). A formal description of the mathematical model of p-median problems was first provided by ReVelle and Swain [38]. A different stream of research was followed by Toregas et al. [39], who initially formulated the location set covering problem for the location of emergency facilities. With this formal ground work being laid, Dökmeci [40] was one of the early contributors, who used operations research techniques to locate hierarchical regional health facilities. His model is based on earlier work in [41], which uses squared Euclidean distances and ends up as a nonconvex optimization problem. Ye and Kim [42] are a prominent example for research that does not apply Euclidean distances as a convenient proxy for actual road distances. Instead the authors employ geographical information systems that allows them to use actual road distances in their healthcare location problem. Their models of choice are network-based location set covering and maximal covering formulations. As early as 1976, Berlin et al. [43] formulated a model that simultaneously locates hospitals and ambulances. Even at this early stage, Fries [44] already compiled 188 references that dealt with the optimization of health care delivery, albeit only ten dealing with hospital location. More recent case studies include those in [45] and [46]. A framework for modeling hospital location is provided in [47], while surveys about applications of operations research in health care are found in [48–50].

Following Burkey et al. in [36], this paper will examine three aspects of hospital closures. A patient-based criterion examines the aforementioned individual efficiency, i.e., the effect of a closure on the average time it takes a patient to reach the nearest hospital. An organizational criterion looks at collective efficiency, i.e., the closures from the hospital’s (or caregiver’s) point of view: it considers the number of potential patients that are covered, i.e., that are within a certain distance or time from their nearest hospital. Finally, we examine a political or governmental criterion that measures the “equity” (to be defined later) of the service that is delivered to the potential clients. Due to the potential of generating impractical solutions when optimizing equity, the criterion is typically used as a secondary measure. In this paper, we will examine the impact of our optimization on equity rather than optimizing it. While this paper uses some definitions, concepts and measures introduced in [36], the two papers are fundamentally different. Burkey et al. [36] compare the existing, historically determined, locations of hospitals with their optimal locations given they could all be relocated now. In contrast, this work uses the existing solution and determines how deletions can be handled best in the sense that the quality of service is maintained at the highest possible level.

Our work will first define appropriate measures for each of the aforementioned criteria and then examine the effects on all three criteria when examining hospital closures. In order to operationalize our discussion, we need to define measures for each of the three criteria. For the patient-based criterion, we choose the average road network travel time between a patient and his closest hospital, provided this does not violate the capacity constraint. This choice follows the discussion in [36]. It expresses the accessibility of hospital care to the individual. Optimizing the accessibility with a given number of open hospitals is typically done by mathematical models that resemble the p-median formulation. It can be described as follows.

Suppose that we have n points, at which potential patients (i.e., the general population) are located. The number of people, who reside at some point i, is used as a proxy for the potential demand at that point, it will be abbreviated as wi. The present existing and known locations of the p hospitals are also known. Denote by dij the distance between customer i and hospital j. The capacity of hospital j is kj, defined as the number of beds available in a hospital multiplied by the average number of people a hospital bed will serve. The fact that we may, at least in emergencies, increase the number of beds in the short run by putting up additional beds in hallways, etc., is ignored here. In our model, the parameter p denotes the number of hospitals that is presently open, k is the number of hospitals that will remain open, so that p – k symbolizes the number hospitals that will close. Furthermore, f with |J| = p is the set of existing hospitals, and N with |N| = n is the set of customer locations.

Define now the location variables yij = 1, if we close a hospital at site j, and 0, if we do not. Furthermore, define allocation variables xij that denote the proportion of customers at site i who patronize hospital j. We can then formulate the hospital closure problems for k = 1, 1, 2, ..., p–1 as follows:

\[ p_k : \text{Min} \ sum_{i \in N} \sum_{j \in f} w_i d_{ij} x_{ij} \]

s.t. \[ \sum_{j \in f} x_{ij} = 1 \forall i \] ... (1)

\[ \sum_{j \in f} y_{ij} = p - k \] ... (2)

\[ x_{ij} \leq 1 - y_{ij} \forall i = 1, ..., n; j = 1, ..., p \] ... (3)

\[ \sum_{i=1}^{n} w_i x_{ij} \leq k_j \forall j = 1, ..., p \forall j = 1, ..., p \] ... (4)

\[ y_{ij} = 0 \text{ or } 1 \forall j \] ... (5)

\[ x_{ij} \geq 0 \forall i, j \] ... (6)

In the above capacitated p-median problem (a problem related to clustering, which was proved to be NP-hard by Garey and Johnson in [51], and received an early treatment in [52], the objective
function minimizes the sum of distances between the potential patients and the hospitals they are assigned to. Constraints (2) ensure that each customer's demand is fully satisfied, and the single constraint (3) guarantees that we close exactly \( p - k \) hospitals. With \( p \) denoting the number of existing hospitals and \( k \) denoting the number of remaining hospitals, \( (p - k) \) is the number of hospitals to be closed; the model will choose which hospitals to close. Constraints (4) ensure that no customers are allocated to a hospital that is chosen to be closed, and constraints (5) guarantee that the capacities of the hospitals are respected. Finally, the specifications (6) restrict the location variables \( y_{ij} \) to be zero or one, while the constraints (7) require that the allocation variables \( x_{ij} \) be non-negative. Note that one of the assumptions of the model is that the entire demand is satisfied, i.e., each customer will receive the service requested. Clearly, this assumption is not satisfied in all regions of the world. However, studying the part of the United States, this assumption is legitimate.

Consider now the organizational criterion. Hospitals will be interested in providing the best possible care they can. Furthermore, other factors being equal, this is a goal that can be achieved by being in close proximity to as many potential patients as possible. As argued by many authors (for a summary, see, e.g., [36]), there is a “30 min” (or 60 min, depending on the incident) time slot, the “golden (half) hour.” Critically injured patients treated within that time have a good chance of survival, whereas after this time elapses, the survival probability drops sharply as per [53]. Given that, it would stand to reason to measure the quality of health care as the number of potential patients or the proportion of the population that is located within 30 min of the hospital. In other words, define \( D \) as the service standard, i.e., the maximal customer-facility distance, within which a customer can be assumed to receive acceptable service. We can then define \( N_i = \{ j : d_{ij} \leq D \} \) as the set of all hospitals that are within the prescribed service distance of the potential patient \( i \in I \), the set of all patient locations. Using again zero-one location variables \( y_{ij} \) and also binary coverage variables \( x_{ij} \), which assume a value of one, if customer \( i \)’s location is within the service distance of a hospital, and zero if this is not the case. The maximal covering location problem (proposed originally by Church and ReVelle [54]; for a recent account, see, [55]) can then be formulated as

\[
P_{\text{MCLP} - k} : \text{Maximize } Z = \sum_{i,l} w_{il} x_{il} \quad (8)
\]

\[
s.t. x_{il} \leq |N_i| - \sum_{j \in N_i} y_{ij} \quad \forall i \in I \quad (9)
\]

\[
\sum_{j \in N_i} y_{ij} = p - k \quad (10)
\]

\[
y_{ij} = 0 \text{ or } 1 \quad \forall j \in J \quad (11)
\]

\[
x_{ij} \geq 0 \quad \forall i \in I \quad (12)
\]

The objective (8) maximizes the number of patients covered by the surviving hospitals, the constraints (9) ensure that a patient site is only considered covered, if and only if there exists at least one open hospital within the service distance, and the single constraint (10) ensures that exactly \( p - k \) hospitals are actually closed.

It is worth noting that the maximal covering problem as we have proposed it above does not use capacities as coverage only requires that a hospital is sufficiently close to a potential patient. The reason is that with capacities in the maximal covering problem, we can no longer ensure full coverage, which is required for a meaningful comparison to the results of the \( p \)-median problem above. Our initial empirical tests confirmed this with the observation that with capacities in place, full coverage was infeasible for all the states with any more than a few hospital closures, thereby making it difficult to compare the solutions to the results of the \( p \)-median problem.

The third attribute concerns the impact of the optimized solutions on “equity.” Interestingly, the word is actually a misnomer, as “equity” (or fairness) involves not only equality, by which it is often measured, but also “need.” A thorough discussion and collection of “equity” objectives was first provided in [30] as well as [31]. Among those “balancing” objectives that are designed to measure equality are variance, mean absolute deviation, Lorenz curve, and the Gini index. The Lorenz curve plots the proportion of the population, ordered from the closest to the farthest from their respective nearest hospital, to the proportion of the total distance that they must travel, where “total distance” is defined as the sum of the distances traveled by all the consumers in accessing their most proximate hospital. The curve allows statements such as “20% of the population must travel 50% of the total distance to the nearest hospital.” The Gini index then measures the normalized area between the Lorenz curve and the 45° line. A Gini index close to one indicates total inequality, while a Gini index of zero indicates total equality. For a good survey of the concept, including its origins and uses, see, e.g., [56].

It should be noted that equality objectives in location analysis cannot stand by themselves as they may (and typically do) lead to nonsensical solutions, in which all customers (potential patients in our scenario) could gain (i.e., be located closer to their closest hospital), if the hospital(s) were to move—at the expense of equality. In other words, an optimizer who considers equality as the sole criterion would prefer a solution, in which all patients were, say, ten miles away from the nearest hospital as opposed to a solution that has one patient eleven miles from the hospital with all other patients being directly next to a hospital. This is why equality objectives are commonly used only in conjunction with efficiency objectives.

It is to be expected that as the number of closed hospitals, i.e., the value of \( p-k \), increases, the average access time, the coverage, and equity will all deteriorate. The question is, of course, how quickly these measures get worse if these closures are determined optimally so as to minimize their adverse impact on the customers; this is the main contribution of our work.

3. Collection and preparation of data

This section presents some of the results as they are applied to the four Southern States Virginia, Tennessee, North Carolina, and South Carolina. These states have been chosen as they are reasonably comparable as far as area, population density, and demographic features are concerned (see [36]).

In this work, we begin by looking at data on hospital locations in 2001 provided by the American Hospital Association (AHA). We use this data from 15 years ago in order that we might compare the closures suggested by our models with actual hospital closures over the intervening years. The number of existing hospital beds is shown in Fig 1. Previous research has found that the marketplace in these states tends to locate hospital facilities in a surprisingly efficient way, see [36]. Section 4 looks at actual closures to see how well they match up with the predictions made by our optimization models.

The data on potential patients come from the year 2000 US Census, to match the hospital location data as closely as possible. We use population at the block group level, which are “neighborhoods” with typically 1000 – 1500 people. For ease of computation, we assumed that the block group centroid carries all the weight of the block group it represents (being fully aware of the potential aggregation errors; see, e.g., [57]). The weight of the block represents the number of people living in that region. We also used as
input the locations of existing hospitals and their respective capacities. As far as hospitals are concerned, we restricted ourselves to General Medical and Surgical hospitals, as they have been defined by the American Hospital Association in their surveys [58].

A natural expression for the capacity of a hospital is its number of beds, which we then had to convert to the number of potential patients that can be served by the facility. In order to facilitate this conversion, we examined the data in [59], which lists the supply of hospital beds per 1000 people in different countries (for example, the United States and the United Kingdom supply about 3 hospital beds per 1000 people on average, or 1 bed per 333 people). Since this average might be expected to contain some excess capacity (to be used in emergencies, epidemics, or represent unneeded capacity), setting a slightly lower minimum standard seems reasonable. In this paper we allow 1 bed per 500 people. Finally, we needed to determine our measure of accessibility between all hospital-customer pairs. Instead of using distance (which is often used as a convenient proxy), we employed the estimated travel time over the road network required to travel from a patient’s location (by census block group) to a hospital. Following [36], we used Microsoft Map Point 2004 to determine the distances, and then assumed speeds of 65 mph, 60 mph, 50 mph, 35 mph, and 20 mph on interstate highways, limited access highways, other major roads, minor roads, and city streets, respectively.

For the \( p \)-median problems, we have applied the following procedure. We first computed \((p - k)\), the largest number of hospitals that can be closed without violating the capacity constraint. In other words, at least \( k \) hospitals are needed to serve all customers. Hence, we solved individual optimization problems with \( k = k, k+1, \ldots, p \), hospitals. For example, in North Carolina, of the 117 hospitals currently in operation, the largest 47 are sufficient to satisfy the demand, i.e., \( k = 47\). This was used as the starting point of our series of tests: we solved \( k \)-median problems with \( k = 47, 48, \ldots, 117 = p \), the latter being the number of existing hospitals in the state. That way, we were certain to always achieve 100% coverage of the demand. The smallest numbers of hospitals that guarantee total coverage in the other three states are 28 for South Carolina, 24 for Tennessee, and 55 for Virginia. Incidentally, the proportion of hospitals that are sufficient to serve the entire population of the state, vary significantly between the states: the number is 40% in North Carolina, 44% in South Carolina, 20% in Tennessee, and 66% in Virginia. A small proportion, such that in Tennessee, indicates that either the largest hospitals in the state have a very large capacity, or the hospitals in the state have a combined capacity that exceeds the demand by a wide margin. We also solved maximal covering problems with \( k = 1, \ldots, p \) hospitals.

4. Computational results and discussion

All computations were performed on a laptop computer with Intel (R) Core (TM) i5-4310 U 2.00 GHz CPU, 16 GB RAM, and Windows 7 operating system. We used CPLEX 10.1 as the solver and GAMS 22.3 as the modeling language. In the branch and bound procedure, we employed the strong branching option for selecting the branching variable (this option conserves memory for large and difficult integer programming problems, which was necessary for us to be able to solve the larger instances in our problems set) and best-estimate search for selecting the next node when backtracking (an option, which limits the number of nodes examined in a search). In our experience, those two options generally improve the performance of CPLEX over the default settings. Other than that, we used GAMS/CPLEX’s default settings for the other parameters. Additional tests revealed that using a newer version, viz., CPLEX 12.6 did not reduce the computing times appreciably. Even with CPLEX 10.1, all computation times were under 190 min each, which is very reasonable for a strategic problem with up to 14,000 variables (and capacity constraints that are well-known to destroy the “integer-friendly” structure of the problem) such as the one considered in this paper.

First consider the results from the \( p \)-median series. As far as computation times are concerned, \( p \)-median instances are solved to optimality in the following solution times. North Carolina: 8300 s \((k = 47)\) to 5 s \((k = 117)\) South Carolina: 580 s \((k = 28)\) to 1 s \((k = 64)\) Tennessee: 10,860 s \((k = 24)\) to 2 s \((k = 118)\) Virginia: 3000 s \((k = 55)\) to 5 s \((k = 83)\).

The results of our tests are shown in Fig 1, where we plot the number of open hospitals against the average traveling time for all four states. It is obvious that the curves are decreasing, albeit to different degrees. In order to shed additional light on the results,
we determine the deterioration of the solution in terms of the increase of access time for up to 12 hospital closures (which represents a range between the minimum of 10.2% in Tennessee and a maximum of 18.8% in South Carolina). An inherent policy assumption here is that decision makers will not go to dramatic lengths and close more hospitals than 20% of existing hospitals at least in the medium run. We note that such a decrease of roughly 10%–20% of all hospitals in a state has a comparatively small effect on access time: the access times will increase between 0.54% in Tennessee to 2.9% in Virginia. We attribute these differences to the relatively high density of hospitals in Tennessee, as compared to the fairly low density of hospitals in Virginia, where the loss of a single hospital can be expected to have a more pronounced effect. This becomes apparent when we compute the percentage increase in access time for each percentage loss of hospitals, which is 0.13 in North Carolina, 0.14 in South Carolina, 0.05 in Tennessee, and 0.20 in Virginia.

What is also interesting in our computations is the stability of the solutions, which is very advantageous in sequential decision making. The concept in question is the so-called inclusion property. The inclusion property, if it exists, simply indicates that the optimal locations of $p$ facilities are fully included in the solution of the same problem with $p + 1$ facilities. It is very well known, that the inclusion property does not hold for $p$-median problems in general. However, in our series of tests, we found that the inclusion property did hold for all runs and was violated only in a single run.

This inclusion property that holds in almost all cases as in our computational runs, has important implications for planners. Since it is typically not known in advance how many hospitals are to be closed in the near future, and it is most likely not politically feasible to suggest closing a large number and the closures will occur of a significant time frame, the inclusion property allows optimal short term closures to be part of a medium- to long-term optimal solution. As mentioned earlier, this is a property that has been observed in almost all computations in this study, but it may (and most likely will) not necessarily hold in all cases, so that sequential closures may not be optimal at all points in time.

Further about robustness, it should be mentioned that in all four states, the hospital chosen to be closed given that a single hospital was to be closed optimally, was included in the set of hospitals to be closed in all solutions with 1, 2, ..., 12 closures. In other words, it would never be wrong to close the hospital that was identified in the optimal single-facility closure.

Fig. 2 shows that the service level, measure by the average travel time, is fairly stable for closures of up to about a third of all existing hospitals. It then starts to increase, until about two thirds of the existing hospitals are closed, after which time the service really deteriorates.

Fig. 3 shows the values of the Gini index in all runs for the four states as a function of the number of closed hospitals. It can be remarked in general that larger numbers of open hospitals tend to provide more equitable solutions as hospitals can be positioned even in areas in which the population density is sparse. Note, however, that this is not always true, see, e.g., the results from North Carolina that get less equitable even at the right tail of the distribution, i.e., when almost all of its hospitals remain open. The very apparent inequities in South Carolina and Virginia for small numbers of open hospitals are results of the small number of existing hospitals in these two states.

Things are very different when we consider the solutions for the maximal covering problem. First of all, all instances of the maximal covering problem were solved to optimality in under 0.2 s. As far as the results are concerned, in none of the four states does the closure of up to twelve hospitals change the coverage of the population. As a matter of fact, the first time the coverage decreases (albeit minimally), is when 24.79% and 23.44% of the hospitals are closed in the Carolinas, and 32.2% and 31.33% in Tennessee and Virginia, respectively. The appropriate curves are shown
In Fig. 3, in other words, as long as hospitals are closed following an optimized plan, coverage is never an issue, unless more than one fifth or even one third of the hospitals are closed. On the other hand, the inclusion property is violated in the runs for all states. In two out of the four states, the hospital closed in the run that closes only a single hospital, is not closed in any of the remaining eleven runs for 2, 3, ..., 12 hospital closures. In other words, if a sequential planner were to apply the solution that optimally closes a single hospital, and later decides to close additional hospitals, the first solution will no longer be optimal. This does present serious problems when attempting to apply maximum coverage solutions over time. The reason is that in a country such as the United States, decision-making regarding hospital closures is decentralized and follows a sequential process. If short-term optimal solutions are not part of a long-term optimum, long-term optima can only be achieved by behaving non-optimally in the short run (which may present problems when “selling” such solutions to the public), or by applying short-term optima and thus missing the best possible long-term solutions.
The Gini indexes for the maximal covering problems are shown in Fig. 5. They show fairly low inequalities with no closed hospitals. Somewhat surprisingly, as a number of hospitals are closed, the degree of inequality decreases. This is a consequence of the fact that most of the initial closures are concentrated in and around urban areas. As more hospitals close, inequalities increase and finally skyrocket, especially in Virginia.

Another major concern we want to address in this work deals with the aforementioned rurality of closures. Authors such as [60] point out that profitability, usage, population change, and the types of services delivered are among the major determinants for the closure of rural hospitals.

In order to operationalize the concept of rurality, we use the classification scheme used by the United States Department of Agriculture [35]. The classification of the indexes is summarized in Table 1.

In order to apply the measure to our results, we consider now only the hospitals to be closed, where the number of closures ranges from 1 to 12 (as a higher number of closures cannot be expected, at least not in the short to medium run). For each of these solutions, we compute the average rurality index of the closed hospitals based on the classification in Table 1. The average rurality indexes of the closed hospitals are shown in Table 2.

As Table 2 shows, the average rurality is relatively stable for larger numbers of closures. There is no reason that the index should be increasing or decreasing monotonously (which in part can be explained by the granularity of integer solutions), but it has much to do with the specific patterns of hospitals that exist in individual states.

We observe that there is no specific “rural bias” as far as suggested hospital closures are concerned. While the average North Carolina closure is in a reasonably-sized town of at least 20,000 people, South Carolina’s closures occur mostly in larger towns and cities. This trend is even more pronounced in Tennessee, where the average closure occurs in cities with at least 250,000 people. The sole outlier is Virginia, where the average closed hospital will be found in a smaller town. An analysis of the details reveals that rural hospitals (which we define as facilities with rurality index of 7 or higher) appear in North Carolina with five or more closings, in South Carolina not at all, in Tennessee only with 11 or more closures, and in Virginia throughout. In other words, regardless how many hospitals are closed in Virginia, there is at least one rural hospital among them. The general low degree of rurality of the closed hospitals in the Carolinas and Tennessee is based on smaller urban hospitals that do not contribute much to accessibility and could be closed according to this criterion. The state of Virginia is different, though, in that it has much fewer hospitals per capita than the other states, and the population is highly concentrated in essentially three centers: Alexandria/Manassas (adjacent to Washington), Virginia Beach/Norfolk, and Richmond. The distribution of the population in Virginia (by county) is shown in Fig 6. The remainder of the state is quite rural, hospitals are small, and their closure will not affect very many people.

Consider now the average rurality indexes when closures follow the plan devised by the maximal covering objective. The results are summarized in Table 3.
Most of the runs confirm the results obtained earlier for the \( p \)-median solutions. The average rurality index with the max cover objective are fairly similar to those obtained with the \( p \)-median objective in the Carolinas. As a matter of fact, even quite a few of the actual hospitals, whose closure is recommended, are the same: for instance, in North Carolina, the single recommended closures are identical with both objectives, and four out of the twelve recommended closures are the same facilities. The overlap in South Carolina is significantly less.

It is also interesting to observe that the average rurality index in the max cover problems in Tennessee is higher with the max cover problems than with the \( p \)-median objective. However, the real surprise is Virginia: while the average closure under the \( p \)-median objective had a high degree of rurality, this is no longer the case when it comes to the recommendations based on the max cover objective. As a matter of fact, given a max cover objective, Virginia closes the most urban hospitals among all states. The low degree of rurality among the hospitals that are to be closed according to the maximum covering objective is based on the binary choice rule. In other words, we consider a potential patient covered, if he is within 30 min of the nearest hospital. This means that hospitals in very dense urban agglomerations can be closed without affecting coverage. On the other hand, closing a rural hospital will typically affect coverage in a larger area. Furthermore, South Carolina and Tennessee do not close any rural (again, index 7–9) hospital, while North Carolina and Virginia include rural hospitals in all of their plans starting with two closed hospitals.

Finally in this context, we compare the average rurality index under the two objective functions given a fixed number of hospital closures. The results are shown in Table 4.

The results in Table 4 indicate that in the case of the \( p \)-median objective, closures are getting more rural as more and more hospitals are closed. This is in contrast to the results of the max cover problem, for which the rurality index slowly decreases as more hospitals are closed. However, both effects are relatively minor.

It is also interesting to compare our optimized closures with those that have actually occurred in practice. It is apparent that the actual closures are quite different from the optimized closures since in reality, most hospitals are privately owned in the USA and therefore their closures are not coordinated by a centralized planning agency such as the state government. For instance, in North Carolina five hospitals were actually closed (with some others having moved, and having their focus redefined). Among those five, three of the hospitals were never selected for closure in any of our optimization runs in either \( p \)-median or max cover solutions. Of the remaining two hospitals that were actually closed in North Carolina, one was selected for closure by our optimization models in only one case of 10 or more closures, while the other was selected for closure in several scenarios we considered. This clearly underscores the significant differences between hospitals that were actually closed versus those that would have been if optimization models had been used to minimize the adverse impact on customers.
Similarly, South Carolina closed four hospitals (and opened five new ones), two of which were never closed in any of the scenarios in the p-median and the max cover problems. Tennessee, having a large number of hospitals to begin with, closed ten hospitals (but opened three new ones). Nine (eight) out of the ten hospitals were never closed in the p-median (max cover) problems, only one of the actually closed hospitals was frequently suggested for closure in optimization runs. Finally, the State of Virginia, somewhat underserved with hospitals in the past, only closed two hospitals (but opened five new hospitals) One of the two hospitals was never suggested for closure in any of the optimization runs, the other was only infrequently suggested to be closed. It is also interesting to note that the average rurality of the actual closures is about 4.4 in the Carolinas and Virginia, and 2.8 in Tennessee, while the average rurality of the optimized closures is 3.6 in case of the p-median objective and 2.7 for the max cover objective (see Table 4). In other words, the actual closures affected more rural hospitals than the optimization suggested.

This part of the analysis compares the locations of the hospitals that were actually closed in these four states with those that were selected by our model to be shut down. In order to do so, we performed a proximity analysis between these two sets. To begin with, note that there was a total of 21 hospitals closed between 2001 and 2016. Of these, the sets of hospitals proposed by the two models captured 9 of them. Table 5 presents a breakdown of these 9 hospitals, along with some descriptive statistics. We see here that these 9 hospitals were much smaller than the average hospital in the data.

Next, we analyzed the location of the 12 hospitals that actually closed, but were not selected by the models in terms of their distance to some of the other hospitals selected by the models. The distance of each of these actual hospitals to the closest proposed hospitals is shown in Table 6 below.

Table 6 shows that—barring one case in North Carolina (Hospital 3)—the distance between the hospitals that were closed and those selected by our models was less than 50 miles in each case. Further, the median distance between the hospitals that were actually closed and those predicted by one of our models was 24.37 miles, which is under 30 min driving time under the reasonable assumption that all hospitals in these states are accessible by primary roads.

5. Summary, conclusions, and future research

This paper discusses the centrally optimized closure of hospitals given two main criteria: the access to a hospital to any given individual, and the accessibility of health care to the population within a given amount of time under the assumption that the hospitals targeted for closure are selected optimally so as to mitigate adverse effects on the population. Our first task was to operationalize the concepts of individual access and general accessibility, so as to be able to optimize the criteria. We identified the p-median problem, which minimizes the average access time for each customer to his closest hospital, and the maximum covering problem, which maximizes the proportion of the population, which is within a given radius (chosen here as 30 min) within a hospital. Our main findings are that the average access time only shows a minor increase as between one and twelve hospitals are closed. Optimizing the coverage indicated that coverage does not change at all if up to twelve hospitals are closed optimally. It was also determined that while the equality of access decreased with an increasing number of closures, that decrease was relatively minor, at least until we got close to the smallest number of hospitals that are able to serve all customers.

Another important question addressed in this paper was the rurality of closures. Contrary to expectations, it was determined that, with the exception of Virginia, the p-median objective closures were throughout in fairly urban areas. Closures in Virginia occurred predominantly in rural areas, which could be explained by the strongly concentrated population distribution. Without exception, the pattern of suggested closures with the maximum covering objective closed hospitals in mostly urban areas.

A number of extensions could be thought of. One issue, which we did not address in this paper, concerns the limitations of the binary coverage measure. The “covered – not covered” dichotomy could be softened and overcome by using one of the gradual covering models suggested in the literature; see, e.g., [61,62], or [63]. Another possible extension involves congestion, which may change the patient-hospital allocations and lead to longer access times. In order to model congestion, we could either include aspects of queuing in the model or apply the concept of double coverage. Yet another possibility would include the use of traffic assignment models the transportation of patients to hospitals.

Acknowledgments

We would like to acknowledge the work of two anonymous referees, whose comments and criticisms have led to a number of clarifications in the paper that have contributed to its improvement. Their advice is much appreciated.

References


