

Optimization of Community Health Center Locations and Service Offerings with Statistical Need Estimation

Paul M. Griffin[§], Christina R. Scherrer[†], Julie L. Swann^{§*}

June 2007

Abstract

Community Health Centers (CHCs) provide family-oriented healthcare services for people living in rural and urban medically underserved communities; they are an important part of the government's plan to make healthcare more affordable. We develop an optimization model to determine the best location and number of new CHCs in a geographical network, as well as what services each CHC should offer at which level. We maximize the weighted demand coverage of the needy population subject to budget and capacity constraints, where costs are fixed and variable. We use statistical methods on national health databases to determine important predictors of healthcare need and disease weights, and we apply these methods to Census data to obtain county-based estimates of demand. Using several performance metrics such as the number of encounters, service of uninsured persons, and coverage of rural counties, we analyze the results of the system approach to location using the state of Georgia as a prototype and demonstrate that optimizing the overall network can result in improvements of 20% in several measures. We use our model to analyze policy questions such as how to serve the uninsured.

1 Introduction

1.1 Motivation

The delivery of health services in the United States is considered by many to be at a crisis point due to both the continued rising costs and the disparities in health status and receipt of preventative care between the poor and the wealthy. A primary concern is the large number of uninsured people in the country, which has risen above 45 million (Hawkins and Proser, 2004). People are more likely to forgo or postpone care if they have no health insurance (Cunningham and Kemper, 1998). This is particularly true for certain sectors of our population including the poor and the elderly,

*Corresponding author

[§]H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0205; Email: {pgriffin,jswann}@isye.gatech.edu

[†]Department of Industrial Engineering Technology, Southern Polytechnic State University, 1100 South Marietta Parkway, Marietta, GA 30060; Email: cscherre@spsu.edu

which in turn can lead to an increase in primary care expenses such as hospitalization for health conditions that could be avoided. A second concern is that even many Americans with health insurance still do not have access to health care. This may be due to low Medicaid reimbursement rates where some doctors are unwilling to take Medicaid patients or have quotas that have already been met (American Academy of Family Physicians, 2002; Berman et al., 2002; Cunningham and Nichols, 2005), or to the fact that many patients live in areas with a shortage of medical healthcare professionals (Bureau of Health Professions, 1992).

Community Health Centers (CHCs) provide family-oriented health care services for people living in rural and urban medically underserved communities. Their services include primary and preventive healthcare, outreach, and dental care, as well as links to welfare; Medicaid; Medicare; Women, Infants and Children (WIC); mental health and substance abuse treatment; and related services (Bureau of Primary Health Care, 2002). CHCs were first funded by the federal government as part of the war on poverty in the mid-1960s. They are now an important part of the US Government's plan to make healthcare more affordable (Office of the Press Secretary, 2004).

In 2003, CHCs served approximately 5 million of the uninsured (Taylor, 2004). CHCs are also often located in areas with few medical professionals and have already been effective in those areas at increasing access to primary care, reducing use of emergency rooms, and reducing preventable hospitalizations (see Blumenthal et al. (1995), Dievler and Giovannini (1998) and references therein). Current CHCs are also cost-effective. Studies have shown that health centers save the government as much as 30% per Medicaid user due to fewer hospital admissions (Falik et al., 1998).

CHCs are typically restricted to locations that are classified as having medical service needs, called medically underserved areas (MUA) (<http://bphc.hrsa.gov/chc/>). The criteria is based on a combination of the ratio of primary care physicians per 1,000 population, infant mortality numbers, the percentage of the population below poverty or above age 65 (Bureau of Primary Health Care, Division of Shortage Designation, 1995). In addition, there is a preference for the centers to be geographically distributed. These locating specifications may not be sufficient; note that currently over 90% of Georgia qualifies as being medically underserved (<http://bhpr.hrsa.gov/shortage/muadatadict.htm>). Research has found that the uninsured are not much more likely to live near a CHC (Cunningham and Hadley, 2004), and that some CHCs may be underutilized because they were not located well (Brecher and Forman, 1981).

Currently, decisions about CHC locations are made individually, without a systematic approach that considers the overall geographical network. Locations are chosen based on competitions of self-submitted proposals, which do not necessarily include the location that is best for the overall

system, or that would lead to a better geographic dispersion of resources (National Center for Primary Care). Our work provides the benefit of being able to consider the location and allocation decision as part of a larger network approach to decision making, and it also allows for consideration of healthcare delivery policy questions. Even after the significant budget expansion for CHCs within the current administration is completed, an improved process for location continues to be justified as CHCs have consistently been an important part of the healthcare safety net in the United States, and the policy discussion of access to healthcare (e.g., expanding CHCs) versus coverage of service (e.g., increasing Medicaid programs) will be ongoing.

For healthcare location problems, estimating demand is clearly an important consideration. To date there are not publicly available estimates of healthcare need for preventative care at sufficiently small geographical areas such as county or census tract. In determining locations of preventive healthcare facilities, Verter and LaPierre (2002) note that “further research is needed to understand the factors that attract people to participate preventive programs [*sic*]”. Another study that focuses on some issues with demand for healthcare facilities is Bennett (1981), where they estimate need using a large survey of members of the local population. While this may be effective for small studies, it is clearly not feasible for large systems. Our work also includes development of a process for estimating healthcare need within individual geographical areas from publicly available data.

In addition to data issues, one of the most difficult aspects of our healthcare problem is to characterize the objective. It is beneficial to cover as large a percentage of the state, geographically, as possible. However, it is also desirable that a high percentage of these covered people are in need of the services offered at the CHCs where they are served, especially in light of the research indicating that current CHCs may not be located to fully utilize their capacity. It is important to reach uninsured members of the communities, but it is also necessary to bring in a portion of patients who are able to provide revenue to the centers in return for their service. The latter is important to maintain financial viability of the centers, as studies have shown this is an issue with many CHCs, e.g., more than half of all CHCs reported operating deficits in 1997, 1998 and 1999 (McAlearney, 2002). Additionally, some types of services could be considered more important than others, for reasons such as their likelihood to prevent the spread of infectious diseases (HIV testing, for example) or their impact on key health benchmarks (such as prenatal care). In this research we develop quantitative methods to determine the contribution of disease to overall health and incorporate the uninsured as well as revenue from the insured to help achieve financial solvency.

The goal of this research is to determine the best location and number of new CHCs, as well as what services each CHC should offer, using publicly available data. We develop an optimization

model that maximizes the coverage of the weighted demand in the population, limited by the overall budget and the capacity levels of each facility and service. To populate the model, we use statistical methods on national health survey databases to determine predictors of some healthcare measures, and we use these predictors with Census data to estimate need within each county. We analyze the results of our model using first the state of Georgia as a prototype and then run a test case for the entire US, comparing our approach using several performance metrics to assess improvement. Our method demonstrates a way to use operations research to compare policies such as the comparison of the benefits of improved access versus increased insurance coverage. Our contribution includes adapting facility location type optimization models to the specific needs of this healthcare problem, integrating several public data sources with statistical methods to populate the model, and using the model to additionally analyze several policy questions.

1.2 Additional Literature

There is a rich body of literature on facility location models (see for example Daskin (1995)). A review of location models specifically applied to healthcare is found in Daskin and Dean (2004). Example applications have included locating hospitals (Sinuany-Stern et al., 1995), emergency medical services (Pirkul and Schilling, 1988), blood banks (Jacobs et al., 1996), and ambulances (Ball and Lin, 1993) among others. In many of the papers described, the focus is on developing heuristics that will solve the problems in a reasonable amount of time, and in some the emphasis is on adaptability to conditions changing in the system such as when siting ambulances. However, in many location problems in healthcare and with modern computers, run-time may not be an issue, as it was not for our problem. In our case, we focus on the modeling (objective and constraints) and on new procedures to estimate demand to populate the model.

The problem that we solve is a variant of the Maximal Covering Location Problem (MCLP), introduced in Church and ReVelle (1974), where locations are chosen to maximize the total coverage of a population. Papers that study capacitated MCLP problems include Current and Storbeck (1988) and Pirkul and Schilling (1991), which primarily focus on developing heuristics to solve the location problems. An MCLP healthcare problem was also studied in Verter and LaPierre (2002), who locate preventive healthcare facilities when the workload at each facility must meet a minimum level. In our case, we choose any number of facilities subject to a budget constraint, where the budget includes a fixed cost and variable portion for each of several services that may be open at a facility. In addition to selecting sites, we choose the specific set of services and capacities at a site, where the overall facility and individual services are limited by capacity constraints.

Another difference in our case is the form of the objective function, which is to maximize the improvement to perceived health, as measured by statistical procedures we develop. Other approaches to estimate the importance of a condition have been developed such as quality adjusted life years (QALYs) (see for example, Dolan et al. (2005) for a review). However, one problem with QALYs is that there can be correlation across multiple conditions, and hence overestimation of certain conditions. In addition, these estimates typically are not developed for different demographic groups, an important component in CHC location.

In our problem, we apply statistical methods to publicly available databases to estimate demand for a variety of health conditions. The method was used in Griffin et al. (2006a) to estimate dental caries in the population but not other diseases. Similar types of model-based estimations have also been performed by the Census Bureau to determine uninsured persons (U. S. Census Bureau, 2005). We also develop statistical links between health conditions and overall health, as described in Griffin et al. (2006c) for dental caries.

The rest of the paper is organized as follows. The optimization model is included in the next section, followed by an explanation of the data, including local estimation methods, in Section 2.2. Section 3 contains the results, and conclusions and future research are presented in Section 4.

2 The CHC Model

Our model determines both locations and service offerings for CHCs. We use a location optimization model that incorporates demand, costs, and facility size. Along with those services handled by a general practitioner, our model can choose to offer prenatal care and gynecology (OBGyn), dentistry, and mental health and substance abuse counseling (M/SA) by determining the staffing levels for those services. These correspond with types of services currently offered by CHCs (Bureau of Primary Health Care, 2002) but that are not necessarily offered at every CHC. There are also additional services that may not require specialized staffing.

In this section we first describe the overall objective. We then present the method we use for estimating local levels of demand from publicly available data. Third, we describe how we derive the cost values. Finally, we present the overall optimization model.

2.1 The Objective Function and Estimates of Coefficients

The overall goal of CHCs is to improve health promotion and access to medical services that improve the general health for the country, and in particular to underserved communities. Several

factors are important to achieve this goal, including: i) number of uninsured persons covered, ii) the prevalence of health conditions, and iii) the importance of health conditions to overall health. We focus on the importance of health conditions to drive the objective in our model, and the prevalence of condition is modeled in the demand constraint. We discuss results for the percentage of uninsured persons covered in the results.

In order to consider multiple related conditions, we use a statistical approach to estimate the importance of a health condition on overall health, controlling for other factors. The data we use is the 1999-2002 National Health and Nutrition Examination Survey (NHANES), which is publicly available at <http://www.cdc.gov/nchs/nhanes.htm>. NHANES consists of a survey questionnaire and physician examination of a sample of persons in the United States, along with weights to convert data to represent the overall population. We regress the presence or absence of a health condition with self reported general health (SRGH), controlling for sociodemographic factors such as race/ethnicity, age, and income. This allows us to explicitly consider any interactions and is therefore applicable to communities with differing sociodemographics (a requirement for CHC location decisions). It should be noted that SRGH has been found to be a good predictor of other measures such as morbidity and mortality (Idler and Benyamini, 1997; Knight et al., 2003; Kington et al., 1997).

Full details of the technique are given in Griffin et al. (2006c), but we summarize the application to CHCs here. In NHANES, we define SRGH at three levels: 1=excellent or very good, 2=good, and 3=fair or poor. In addition, for each of the conditions considered, each person (N=9461) in the survey was asked if they have been told by a physician that they have the condition. NHANES also contains full sociodemographic data including income, race/ethnicity, age, and gender. We use SAS-callable SUDAAN to perform a logistic regression with the final model obtained using Kleinbaum's hierarchical backward selection method. The reference group was non-Hispanic whites with income greater than 200% of the federal poverty level (FPL) since the remaining significant sociodemographic factors in the final model were race and income. The odds-ratios, corresponding weights, and prevalence values are given in Table 1 for the health conditions. All were significant at the 5% level.

The impact of each condition on SRGH is simply the inverse of the odds ratio, and these define the weights of our objective function. Note that we aggregated conditions based on which type of health professional would treat them. Our objective then is to maximize the number of annual weighted patient encounters, summed over the various locations and services. This is in essence equivalent to maximizing SRGH for the community over the set of services the CHCs offer.

The goals of the Department of Health and Human Services report, *Healthy People 2010* (U.S. Department of Health and Human Services, 2000) are to: i) increase quality and years of healthy life, and ii) eliminate health disparities. Focusing on the importance of health conditions using self reported general health directly aligns with this first goal. We address disparities by incorporating need (which is a function of sociodemographic conditions) into the objective, and in a companion policy paper (Griffin et al., 2006b), we analyze locating CHCs to improve access to primary care for those who do not already have it.

One limitation of this approach is that OBGyn services such as deliveries are not a “health condition”. We therefore treat OBGyn first as high importance (equal to the highest weight), due to prenatal care being an important service for CHCs to provide (<http://bphc.hrsa.gov/chc/programexpectations.htm>). We then compare results obtained treating OBGyn as average importance (equivalent to general service).

2.2 Data Estimation

The previous section showed how we obtain weights for different services, but for the optimization model, we still need estimates of the number of people with a condition, the number of encounters per service type, and costs. We describe our methods for obtaining these here.

2.2.1 Demand Estimation

While national data is publicly available for the prevalence of health conditions (NHANES, for example), there is very little data available for smaller regions. The Behavioral Risk Factor Surveillance System (BRFSS), publicly available at <http://www.cdc.gov/BRFSS/>, is one example that contains some data for states and metropolitan areas, but not at the county level. In addition, it is unlikely that we will have adequate resources in the near future to collect such data, especially if a clinical examination is required. We therefore use a statistical technique to perform local estimation based on the approach by Griffin et al. (2006a). The data is from NHANES, BRFSS, and the U.S. Census (<http://www.census.gov>), which are all publicly available.

As an example, to forecast prevalence of dental caries for each local region (a county or census tract), we took the sum of the percentage of the population of that region that fall in each US Census category (race/ethnicity, income, gender) times the prevalence of that category. In order to validate this method, we partition the NHANES data into 1000 samples of 1006 individuals, and compute the bias and root mean squared error (rmse) by comparing the estimated caries prevalence to the true prevalence. For census estimates the bias was -0.028 with a rmse of 0.115, both of which

are low.

We use a similar procedure for the health conditions needed in the model except that we base OBGyn demand on the number of births in each county as given in the US Census. Full logistic regression results are included in Scherrer (2005) and prevalence results are available upon request from the authors.

Given that we have estimated the disease need in a particular county, we need to determine how many people are likely to use a CHC. NHANES contains a survey question asking what type of place a patient uses for their primary healthcare services. We assume an uninsured individual is likely to use a CHC if they chose clinic. Note that this result may underestimate since a person would only choose “clinic” if one were available to him, and not all areas may have a clinic in close proximity. There may also be some bias by sociodemographic status, since clinics may more likely be based in low-income areas. Also, there are many more clinics than CHCs, so it is not a reflection only of the current CHC network. To attempt to minimize this bias, we only use the value for uninsured persons and then scale the percentages of insured (government and private) persons likely to use a clinic from the percentages of each category actually visiting the clinics in Georgia in 2002. We then use the same statistical technique described above to estimate the number of people in each county who are likely to visit a CHC.

Once the number of people with each health condition are estimated, we convert that value into the number of people who will likely go to a county’s CHC with that condition by computing the product of population, prevalence of condition, and likelihood of CHC visit. To estimate the demand in *number* of patient encounters, we multiply the number of people who will likely go to a county’s CHC for a condition by the average number of encounters for that condition. Taking Georgia as an example, where data sources other than Bureau of Primary Health Care (2002) are specified:

- To estimate the demand for general services, we multiply the number of people likely to use a clinic estimated by demographics by the likelihood of visiting a doctor in the last year, which is 81% according to the National Center for Health Statistics (2001),
- To estimate the demand for dental services, we scale the dental demand by likely clinic use and then multiply by the average number of annual dental encounters per dental user in 2002, which was 1.60.
- For OBGyn services, we scale the number of pregnancies by likely clinic use and then multiply by seven (half of the number of visits recommended by the American College of Obstetricians and Gynecologists, as many women receiving care at CHCs do not visit until their second or

third trimester and may not visit as frequently as recommended), to obtain an estimate for number of encounters.

- To estimate the demand in number of encounters for mental health and substance abuse services, we scale by likely clinic use and multiply by 3.28, which was the average number of annual mental health encounters per mental health user in Georgia in 2002.

Along with people who live in the county where a CHC is located, some additional demand will come from nearby counties. To account for this, we determine distances between counties from census latitude and longitude data, measured from the center of each county. Health Resources and Services Administration (HRSA) has guidelines that access is desired to be within 40 minutes of travel (<http://bhpr.hrsa.gov/shortage/hpsacritmental.htm>) and define this as between 20-30 miles of travel. We made the likelihood of visiting a CHC a decreasing function of distance, with a maximum distance defined by the HRSA guidelines. We assume that willingness to travel does not differ by service, which is reasonable since CHCs do not offer many specialized services not available elsewhere. Travel is described in detail in Section 2.3. Patients from a county can be served by more than one CHC location. However, the total number of a county's patients served by all CHCs is constrained by the total demand from that county.

2.2.2 Costs

We estimate costs using CHC reports. We illustrate the steps in this section using Georgia as an example. For the experiment where we locate facilities nationally, we use the average costs from the national version of the roll-up report used for the state of Georgia estimates.

From the 2002 Georgia CHC report (Bureau of Primary Health Care, 2002) we determine the average number of patients served by the different types of practitioners required for the services we are considering (including support nursing staff). We use these figures along with Georgia's census median salary data to obtain estimates of the variable cost of each service. The exception was mental health and substance abuse encounters, which are a separate line item in the Georgia CHC report. All costs for Georgia are summarized in Table 2.

Also from the CHC report, dividing the dollars spent on equipment and supplies by the total dollars spent on staffing, we estimated the portion of the cost of equipment and supplies. Similarly, we estimate the average cost for laboratory and x-ray. Dividing the total cost for pharmacy, administration, and enabling (translation during visits, transportation, etc.) by the sum of the costs above per service type, we estimate that allocation as well.

We classify CHCs by size to determine appropriate service levels for the four service types using

data from the CHC report (Bureau of Primary Health Care, 2002). For example, for the general category, total encounters ranged from 3,391 to 97,633 with a mean of 29,438 in Georgia, so we set the capacity of a small clinic at 8,000 encounters, medium 30,000, and large 70,000. See Table 2 for the other capacity values.

The final variable cost of service of a patient from a county is dependent upon the estimated percentage with each type of insurance coverage estimates from above weighted by the probabilities of each insurance type to use a CHC. To determine the variable cost of service in a county, we calculate the weighted average of the variable costs for the service for each type of insurance coverage. In 2002, CHC users with public insurance (Medicare, Medicaid, etc.) paid 74.0% of their charges, users with private insurance paid 64.1% of their charges, and uninsured users paid 22.3% of their charges (Bureau of Primary Health Care, 2002). We assume these reimbursement values are on the above costs.

Finally, the fixed cost per service (FS_{jk}) is \$5,000 per small service, \$10,000 per medium service and \$15,000 per large sized service offered, and the fixed cost per location (FL_i) is estimated to be \$100,000 per location. We estimate these from the annual facility costs for Georgia (Bureau of Primary Health Care, 2002) and apply them to other states in the national model.

2.3 Optimization Model

The optimization model for location and service selection of CHCs uses cost, demand, and geographical considerations in the constraints. Costs are split into their fixed and variable components as detailed previously, and the model has an overall budget constraint. The model is constrained by the demand for services as a function of estimated county-level need data. We also determine how many people a CHC serves by placing limits on how far a person would travel as discussed in Section 2.2.1. Based on the HRSA guidelines previously mentioned in the Section, we use four distance levels, indexed by l , with $P_1 = 1.0$, $P_2 = 0.75$, $P_3 = 0.5$, and $P_4 = 0.25$. People in category $l = 1$ are not willing to travel outside of their location, $l=2$ up to 10 miles, $l = 3$ up to 20 miles and $l = 4$ up 30 miles. For distances greater than 30 miles, the probability was set to 0. For example, if the distance between i and z is 15 miles, that corresponds to level 3 and 50% of location i 's population is willing to travel to location z for service.

We examine the impact of increasing the budget on the number of patients that we are able to serve and the cost per encounter for the corresponding location choices. We also examined sensitivity with respect to the maximum distance a patient is willing to travel.

In addition to previously stated assumptions, we assume that a person's willingness to travel

does not vary by service. We also assume that adding new CHC locations will not change the expected behavior of an individual. That is, the prediction of whether a person will use a CHC is based on demographic estimates that do not vary with the location of available CHCs. The overall optimization model, then, makes two assumptions about behavior. First, the number of potential patients that would seek CHC service is estimated from county-level demand described earlier. Second, for those estimated as candidates for visiting a CHC, the probability of a visit is defined by their willingness to travel.

The notation for the model is as follows.

Indices

i = locations index (Note: we also use z when we are comparing two locations)

j = services index

k = index of levels for each service

l = index on distance levels (Note: we also use q when summing on a subset of distance levels)

Decision variables

y_{izjl} = patients from z who are served by a center in i for service type j and within distance category l from their location (relaxed to linear)

s_{ijk} = binary indicator variable of service type j at level k in location i

c_i = integer variable for number of centers in location i

Data

w_j = the weight associated with serving a customer of type j

P_l = the maximum percentage of one county's population that can be served in another county if those two counties are distance level l apart

n_{ij} = need (demand) for service j in i

d_{izjl} = maximum demand for service type j in location i that can be served in location z .

(= $P_l * n_{zj}$ if the distance between i and z corresponds to level l , 0 otherwise)

CAP_{jk} = number of patients of service type j that can be served at level k

B = budget

FL_i = Fixed cost for location i

FS_{jk} = Fixed cost for service j at level k

VS_{ij} = Variable cost for service j at i after patient/insurance reimbursement

Model

$$\max \sum_{izjl} w_j y_{izjl} \quad (1)$$

$$\text{s.t.} \sum_i FL_i c_i + \sum_{ijk} FS_{jk} s_{ijk} + \sum_{izjl} VS_{ij} y_{izjl} \leq B \quad (2)$$

$$\sum_{zl} y_{izjl} \leq \sum_k CAP_{jk} s_{ijk} \quad \forall i, j \quad (3)$$

$$\sum_k s_{ijk} \leq c_i \quad \forall i, j \quad (4)$$

$$y_{izjl} \leq d_{izjl} \quad \forall i, z, j, l \quad (5)$$

$$\sum_{l \geq q, i} y_{izjl} \leq P_q n_{z,j} \quad \forall z, j, q \quad (6)$$

$$y_{izjl} \geq 0 \quad (7)$$

$$s_{ijk} \in \{0, 1\} \quad (8)$$

$$c_i \text{ integer} \quad (9)$$

This is a mixed integer optimization problem with linear constraints. The objective is to maximize the total weighted number of patients served. Constraint (2) is the budget constraint and constraint (3) that patients can only be served if there is capacity available for them at that service level. Constraint (4) states that there can only be as many locations offering service type j as there are open locations, and, combined with constraint (3), implicitly requires that patients of type j can be served at facility i only if that center is open and offering service j . Constraint (5) only allows the proportion of patients that are eligible based on the distance calculation to be served. Constraint (6) enforces the maximum total percentage of location i 's population served by locations more than each distance level away.

CHCs are specified to be located in medically underserved areas, so in some cases we add a constraint for this. This also addresses locating CHCs in areas where physicians are needed. However, it is also useful to solve the optimization model without this constraint, to see how different the solution is and evaluate this policy rule.

There are several limitations to our model. First, the solutions resulting from our model are dependent upon the input data for costs and demands, much of which we had to estimate. Second, solving the model for counties assumes that all of the population of a county is located at the center of the county. While we do not believe this to be a significant problem for Georgia's relatively small counties, this could result in bad solutions where counties are large and sparsely populated or when counties are large and the population is densely located in a small off-center area. Third, we

assume that there are no economies of scope or scale for service offerings, and that all services could be offered at any location. Finally, we assume that there are sufficient medical resources, namely physicians, dentists and nurses, in all locations chosen. Rosenblatt et al. (2006) found that for federally funded CHCs in urban areas, that were vacancies of 13% for family practice physicians, 21% for OBGyn, and 23% for psychiatrists. If data were available on medical personnel shortages, this could easily be incorporated into the model. In addition, this illustrates even more the importance of using good models for location.

3 Results

3.1 Georgia Solution

Current CHC locations for the state of Georgia are shown in Figure 1 along with the uninsured quartiles of each county; the demographics suggest that the network does not cover the maximum number of uninsured persons. This network had a government subsidy budget of \$44M in 2002. We compare to this case by using the same budget of \$44M and a maximum driving distance of 30 miles. To keep solution times reasonable, all of the models below were run for solutions guaranteed to be within 1% of optimal.

Summary results from the optimization model for the number of locations, number of encounters, cost per encounter, and the number of CHCs offering each service are included in Table 3 for a \$44M subsidy. The MUA constrained solution (Column (B)) considers only counties that are medically underserved (MUA) but where locations are not constrained at all by the current network. It should be noted that the current CHC network in Georgia (Column (A)) does have some CHCs located in counties that are not currently MUA counties, although it is possible the counties were MUA when the centers were initially located. As can be seen from the solution, 27 CHCs are located in solution (B), the cost per encounter drops by almost 20% compared to the current network, and the number of encounters increases by over 20%. In addition, more of the demand is satisfied, and all of the services are located in more locations than currently. If locations are not-constrained to MUAs (Column (C)), the solution in all factors improves. Figure 2 shows the non-MUA constrained optimal solution along with the population quartile of each county; the map also shows that population density may be one factor but is not the only driver for location choice. Comparing to Figure 1 and Column (A), we see that only 5 of the counties have CHCs in both solutions.

Finally, if we restrict current CHC locations to be open, but allow additional facilities and

reallocation of services provided (Column (D)), we see that the solution is only slightly worse in the cost per encounter and number of encounters than Column (C), which has no restrictions. Since it is highly unlikely that a CHC would be closed down, Column (D) would be the easiest to implement and hence the most likely solution. Notice that even when constrained to keeping all existing facilities, significant improvements in the number of encounters and services offered can be made. This implies that it is important to not only choose locations for the overall facility, but also the appropriate services and capacity levels.

As a comparison to the optimal results, we also considered several heuristics to choose the locations of CHCs based on measures that might be used by policymakers to locate facilities. Specifically, we choose counties in ranked order of the number of uninsured persons, number of total persons, or MUA index of each county. In each case we allowed the optimization model to choose the exact services and capacities offered. Interestingly, locating based on density while choosing services and capacities optimally performed quite well in the state of Georgia, although the services chosen were fairly different than the other solutions; this also suggests that it is important to choose services and capacities well. Locating based on the number of uninsured persons or the county MUA index did not do as well on any of the measures, with MUA performing the worst. This suggests that the MUA index alone is not sufficient to capture characteristics of the network. It should be noted that locating based primarily on population density is not guaranteed to do well for states with other geographical characteristics, and it does not meet the goals of the CHC program to provide care in rural (less-dense) areas.

In order to compare how well the model represented the current CHC network, we constrained the model to choose the exact same locations as are currently open, with the same annual budget, and assumed that patients were willing to travel at most 30 miles. In this case, the model chooses to offer each service at every location with approximately 898,000 total patient encounters (compared to actual encounters of 803,106 in Column (A)). The difference between the solutions was less than 12%, which could in part be explained by allowing different selections of services. Note that the estimates of the number of improved encounters in the optimized solutions is 18.4% suggesting that some improvement is due to the optimization itself, not to data differences. However, if we assume that patients will only visit a CHC in their own county (solution not shown), the number served decreases to approximately 605,000 which is quite different. It appears then that it is important to explicitly model distance, and that the estimates used match fairly closely to actual values.

3.1.1 Sensitivity to the Budget Constraint

It is of interest to study the impact of increasing the budget constraint, since this will allow for analysis of expansion to the program as well as other policy questions such as access versus coverage. We start with the current CHC network, and increase the budget incrementally from the current subsidy of \$44M up to \$132M (a value that allows service of more than 99% of the demand likely to visit a clinic). One issue is that access becomes more expensive as CHCs are added since some of the state's population is in low density communities where economies of scale are difficult to achieve. We therefore determine the cost per encounter and total number served for each budget level. These are shown as a function of budget in Figures 3(a) and 3(b), respectively.

As seen from the figure, the cost per encounter remains fairly flat up to approximately \$110M. Beyond that point, there are diminishing returns to adding CHCs. However, increasing the budget to this level, would satisfy over 90% of the demand as compared to less than 40% satisfied by the current budget. It therefore appears that an increased use of CHCs in Georgia would have a significant impact on general health, and in addition be relatively cost effective.

3.1.2 Impact on Needy Populations

As one of the purposes of the CHC program is to reach uninsured persons, it is of interest to see how well the different solutions given in Table 3 performed with respect to this metric. The current CHC network (Column (A)) has locations in 4 counties among the 25% with the highest uninsured populations, while the solutions for the not restricted (Column (B)) and MUA restricted (Column (C)) have locations in 5 of these counties. We also calculate the number of uninsured persons served by each solution; all of the optimized solutions perform better than the current network even though they are optimizing a different objective. An alternative measure is the average of the percentage of uninsured persons per county location. The benefit of this measure is that it is not biased towards solutions with more locations. The values are 14.1% for the current solution, 14.3% for both the non-restricted and MUA restricted solutions and 14.5% for the MUA-restricted solution. In both measures, therefore, our model solutions better serve the uninsured populations, which is one purpose of a healthcare safety net.

There is also a preference for CHC locations to be in rural communities since there are fewer opportunities for other types of access. The current CHC network has one location in a county with the lowest 25% population density, while the non-restricted solution has locations in three of these counties and the MUA-restricted solution has locations in 5 of these counties. If we compute the average population density in people per square mile for each solution, we get 691.3 for the

current network, 370.9 for the non-restricted solution, and 331.0 for the MUA-restricted solution. The non-restricted optimization solution (Column (B)) gives the solution with the best results in both of these measures. However, it is important to note that the range is quite large in each of the three cases, and the non-restricted average is significantly impacted by having approximately half as many locations in the five most urban counties as the other two solutions. Therefore our model solutions perform better in both measures, with the MUA-restricted performing the best.

Another needy population is people who have no source of primary care. As one of the goals of the Healthy People initiative is to improve access to primary care for those who do not already have it, we consider this measure as well. See Griffin et al. (2006b) for a full discussion of this objective, where we specifically optimize CHCs towards this and compare increased access (through adding CHCs) to increased coverage (by increasing Medicaid coverage); we find that CHCs are a cost-effective mechanism for increasing access. Although in the current paper we optimize weighted disease, we also calculate the improvement to the number of persons who gain a source of primary care for the current network and optimized solutions (see Table 3). The optimized solutions show a 15 - 20% improvement over the current network in increasing those with a source of primary care.

3.2 National Solution

We apply the same optimization model on data from the entire US, where decisions are for each county. The data for the model was collected in the same way as for Georgia and is available from the authors upon request. The government subsidy budget used was the 2004 total CHC grant revenue of approximately \$2.48 Billion. Figure 4 shows the CHC locations along with population quartiles.

Several observations can be made. First, in comparing any of the solutions from Section 3.1 for the state of Georgia to those in the national solution, it is clear that Georgia is severely medically underserved; and resources from other states should be shifted to Georgia. 89.9% of Georgia's counties are MUA, while nationally the rate is lower. Second, this demonstrates that the network effect of CHC location and service location is not only seen at the scale of a state, but also at the national level. This of course is driven in large part by the "demand" for services. Figure 4 shows that much of this unmet demand is in southern states such as Alabama, Georgia, and Mississippi. Third, we found the MUA constraint did reduce the total number of CHCs, but only had a big impact on a few states; namely Colorado, New Mexico, North Dakota, Texas, and West Virginia. Of these, the western states all gained CHCs in the MUA-constrained solution compared to the not-constrained solution. Note that they all have large areas of land with relatively small population

densities, and it was some of the counties in those areas that gained under the constrained solution. There was therefore some shift of CHCs from urban to rural areas. Clearly the value of MUA is to ensure that these populations get preference. One of the benefits of the national CHC location model is that it allows for quantifying the services delivered, enabling policy makers to weigh the cost and benefit of their decisions such as requiring CHC locations be in MUA counties or that a certain percentage of locations be in rural communities.

4 Conclusions and Future Research

In this paper we have presented a systematic approach to CHC location and service selection. We have also shown how publicly available data can be used to estimate demand for services at a local level and cost of services provided including measures of disease improvement. Finally, we applied the approach to the state of Georgia and then to the entire United States.

When making comparisons between our model and the current CHC locations, it is clear that the current location and service decisions are not the most cost-efficient. Locations and services chosen by our model offered improvements in all measures of 15 - 20%, such as for encounters per year, and our solutions better target uninsured and rural populations, even though this was not an explicit goal of the optimization model. Our results also show that it is important to choose not only locations, but also the services and capacities of the locations. Overall our approach demonstrates a benefit to being able to consider the location and allocation decision as part of a larger network approach to decision making.

In addition to the benefits of using the model to choose location or services, our approach offers a way to evaluate a number of policy questions. For example, results allow one to quantify the impact of requiring CHCs to serve rural or MUA areas, and our results can be used to quantify the cost and benefits gained from targeting needy populations such as the uninsured or people without a source of primary care. This allows policymakers to explicitly compare different modes of delivery for healthcare access such as increasing access or increasing coverage. For instance, it appears clear from the Georgia example that increasing the CHC government subsidy (to at least double) would be an effective way to improve the general health of the state as the cost per encounter would remain relatively constant over this range.

There are many who would argue that Community Health Centers should be more financially self-sufficient, and certainly many struggle with that given that they are supposed to be providing a safety net for certain segments of the population. This paper illustrates some of the negative

ramifications associated with that prospect. Making decisions only on budget considerations leaves out some of those most in need of services - namely those who are uninsured and those who live in very rural areas. Rural areas do not have enough demand to achieve the economies of scale that overcome the fixed costs. It is less cost effective to serve the uninsured population, since they tend to pay significantly less for services than those who are insured. This paradox must be handled appropriately in policy decisions. The results from this work demonstrate many of the tradeoffs that are inherent in this problem. The work could also be used to locate facilities in a way that would improve the self-sufficiency of CHCs. We should point out that insured patients take much less of the government subsidy than uninsured patients for the same capacity. Part of the benefit of our model is that communities with high levels of both types of individuals are good candidates for CHC location, something that is typically not currently considered in choosing potential sites.

There are at least three important areas that we are interested in pursuing in future work. One is to further evaluate mechanisms of improving healthcare delivery such as improving access (adding CHCs or clinics in Walmart) versus increasing coverage (expanding Medicaid coverage or reimbursement rates) (Griffin et al., 2006b). Second, the model we develop could also be used to help make decisions that take future migration patterns into consideration. For example, the Hispanic population is currently increasing in Georgia. Being Hispanic is a risk factor for certain health conditions, and Hispanics are more likely to be uninsured than white individuals (National Center for Health Statistics, 2001). By using forecasts of future demographics, we can forecast future need for CHCs' services that can be taken into account in policy decision making. Finally, we have aggregated several services provided into four main categories. It would be of interest to examine in a more detailed way individual services (such as breast cancer screening) in the context of the CHC network. This could then be compared to other approaches to screening, and help to make decisions about what services CHCs should simply refer to other providers.

References

American Academy of Family Physicians. Medicaid shortfalls prompt concerns, 2002.

<http://www.aafp.org/fpr/20020700/2.html>, Accessed July 2006.

American College of Obstetricians and Gynecologists. <http://www.acog.org>. Accessed June 2005.

M. O. Ball and F. L. Lin. A reliability model applied to emergency service vehicle location. *Operations Research*, 41(1):18–36, 1993.

- W. D. Bennett. A location-allocation approach to health-care facility location - a study of the undoctored population in Lansing, Michigan. *Social Science and Medicine Part D-Medical Geography*, 15(2D):305–312, 1981.
- S. Berman, J. Dolins, S. F. Tang, and B. Yudkowsky. Factors that influence the willingness of private primary care pediatricians to accept more Medicaid patients. *Pediatrics*, 110(2):239–248, 2002.
- D. Blumenthal, E. Mort, and J. Edwards. The efficacy of primary-care for vulnerable population groups. *Health Services Research*, 30(1):253–273, 1995.
- C. Brecher and M. Forman. Financial viability of community-health centers. *Journal Of Health Politics Policy And Law*, 5(4):742–768, 1981.
- Bureau of Health Professions. Rural health professions facts: Supply and distribution of health professions in rural America, 1992.
- Bureau of Primary Health Care. Georgia roll-up report. Technical report, United States Department of Health and Human Resources, 2002. Available by request at <http://bphc.hrsa.gov/uds/>, Accessed July 2005.
- Bureau of Primary Health Care, Division of Shortage Designation. Guidelines for medically underserved area and population designation. <http://bhpr.hrsa.gov/shortage/muaguide.htm>, June 1995.
- R. L. Church and C. S. ReVelle. The maximal covering location problem. In *Papers of the Regional Science Association*, volume 32, pages 101–118, 1974.
- P. Cunningham and J. Hadley. Expanding care versus expanding coverage: How to improve access to care. *Health Affairs*, 23(4):234–244, 2004.
- P. Cunningham and P. Kemper. Ability to obtain medical care for the uninsured: How much does it vary across communities? *JAMA: The Journal of the American Medical Association*, 280:921–927, 1998.
- P. J. Cunningham and L. M. Nichols. The effects of Medicaid reimbursement on the access to care of Medicaid enrollees: A community perspective. *Medical Care Research And Review*, 62(6):676–696, 2005.

- J. R. Current and J. E. Storbeck. Capacitated covering models. *Environment And Planning B-Planning and Design*, 15(2):153–163, 1988.
- M. S. Daskin. *Network and Discrete Location: Models, Algorithms, and Applications*. Wiley-Interscience, 1995.
- M. S. Daskin and L. K. Dean. Location of health care facilities. In M. L. Brandeau, F. Sainfort, and W.P. Pierskalla, editors, *Operations Research and Health Care: A Handbook of Methods and Applications*, pages 43–76. Kluwer Academic Publishers, USA, 2004.
- A. Dievler and T. Giovannini. Community health centers: Promise and performance. *Medical Care Research And Review*, 55(4):405–431, 1998.
- P. Dolan, R. Shaw, A. Tsuchiya, and A. Williams. QALY maximisation and people’s preferences: a methodological review of the literature. *Health Economics*, 14(2):197–208, 2005.
- M. Falik, J. Needleman, J. Korb, and N. McCall. ACSC experience by usual source of health care: Comparing Medicaid beneficiaries who rely on CHCs with other primary care providers: Final report, 1998.
- P. Griffin, S. Griffin, J. Swann, and N. Zlobin. A new method for local estimates of dental disease, 2006a. Working paper, Georgia Institute of Technology.
- P.M. Griffin, C. R. Scherrer, and J.L. Swann. Access through community health centers or coverage through Medicaid: A geographical and mathematical analysis of the state of Georgia, 2006b. Working paper, Georgia Institute of Technology.
- S.O. Griffin, P.M. Griffin, K. Jones, J.L. Swann, and J. Etchison. The impact of health conditions on self-reported general health, 2006c. Working Paper, Centers for Disease Control and Prevention.
- D. Hawkins and M. Proser. Special topics issue brief, a nation’s health at risk: A national and state report on America’s 36 million people without a regular healthcare provider, 2004.
- E. L. Idler and Y. Benyamini. Self-rated health and mortality: A review of twenty-seven community studies. *Journal Of Health And Social Behavior*, 38(1):21–37, 1997.
- D. A. Jacobs, M. N. Silan, and B. A. Clemson. An analysis of alternative locations and service areas of American Red Cross blood facilities. *Interfaces*, 26(3):40–50, 1996.

- R. Kington, L. Lillard, and J. Rogowski. Reproductive history, socioeconomic status, and self-reported health status of women aged 50 years or older. *American Journal Of Public Health*, 87(1):33–37, 1997.
- E. L. Knight, N. Ofsthun, M. Teng, J. M. Lazarus, and G. C. Curhan. The association between mental health, physical function, and hemodialysis mortality. *Kidney International*, 63(5):1843–1851, 2003.
- J. S. McAlearney. The financial performance of community health centers, 1996-1999. *Health Affairs*, 21(2):219–225, 2002.
- National Center for Health Statistics. Fast stats: Ambulatory care use/physician visits. <http://www.cdc.gov/nchs/fastats/docvisit.htm>, 2001. Accessed July 2005.
- National Center for Primary Care. Georgia’s health safety net: Access to primary care for Georgia’s uninsured and underserved. <http://www.msm.edu/ncpc/publications/GAHealthSafety.pdf>, Accessed March 2007.
- Office of the Press Secretary. Fact sheet: President Bush’s plan to make health care more affordable, September 2004. <http://www.whitehouse.gov/news/releases/2004/09/20040902.html>, Accessed August 2006.
- H. Pirkul and D. A. Schilling. The siting of emergency service facilities with workload capacities and backup service. *Management Science*, 34(7):896–908, 1988.
- H. Pirkul and D. A. Schilling. The maximal covering location problem with capacities on total workload. *Management Science*, 37(2):233–248, 1991.
- R. A. Rosenblatt, C. H. A. Andrilla, T. Curtin, and L. G. Hart. Shortages of medical personnel at community health centers - implications for planned expansion. *Jama-Journal Of The American Medical Association*, 295(9):1042–1049, 2006.
- C. R. Scherrer. *The Allocation of Scarce Resources in Public Health*. PhD thesis, Georgia Institute of Technology, 2005.
- Z. Sinuany-Stern, A. Mehrez, A.-G. Tal, and B. Shemuel. The location of a hospital in a rural region: The case of the Negev. *Location Science*, 3:255–266, 1995.
- J. Taylor. The fundamentals of community health centers, 2004. National Health Policy Forum Background Paper. The George Washington University, Washington, D.C.

U. S. Census Bureau. Small area health insurance estimates: Model-based estimates for counties and states, 2005. <http://www.census.gov/hhes/www/sahie/>, Accessed August 2006.

U.S. Department of Health and Human Services. Healthy people 2010: Understanding and improving health, 2000. <http://www.healthypeople.gov/>, Accessed January 2007.

V. Verter and S. D. LaPierre. Location of preventive health care facilities. *Annals Of Operations Research*, 110(1-4):123–132, 2002.

Acknowledgments

We would like to thank Professor Joel Sokol of the Georgia Institute of Technology on his advice that improved the formulation of the optimization model. We would also like to acknowledge Sebastian Urbina for his help in the initial work on this problem. This work is supported in part by grant NSF DMI-0348532. The findings and conclusions in this manuscript are those of the authors and do not necessarily represent the views of the National Science Foundation or other sponsors.

Biosketches

Paul Griffin is a Professor in the Stewart School of Industrial and Systems Engineering at the Georgia Institute of Technology. His research and teaching interests are in the application of economic decision analysis to supply chain and public health settings.

Christina Scherrer is an Assistant Professor in the Industrial Engineering Technology department at Southern Polytechnic State University. She received her Ph.D. in Industrial Engineering from the Georgia Institute of Technology. Her research interests are in the application of economic decision analysis and operations research to public health and non-profit sectors. She is the faculty co-advisor for the SPSU IIE student chapter.

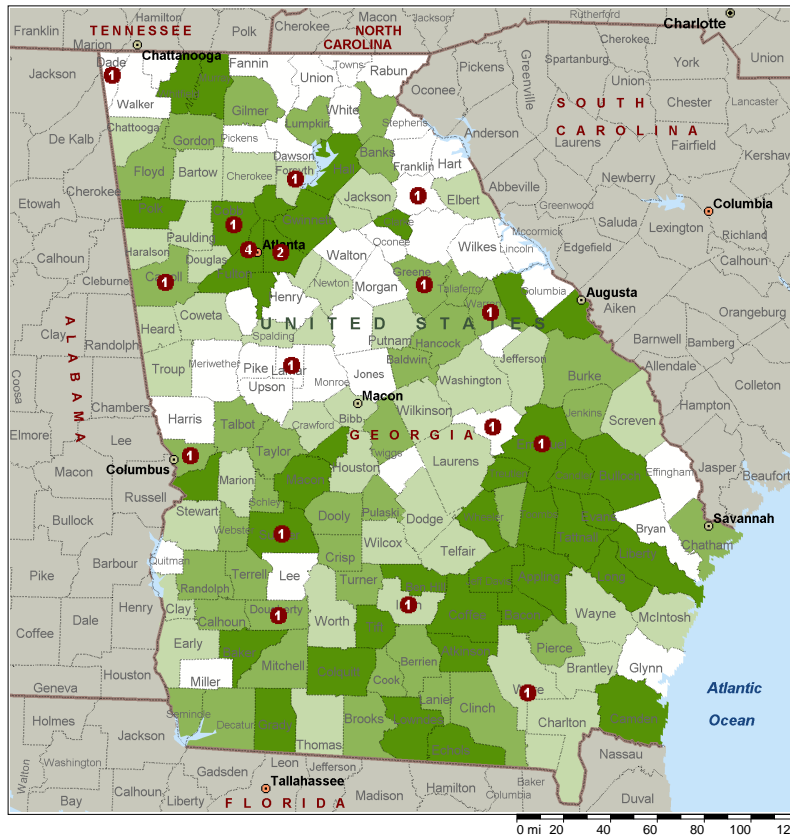
Julie L. Swann is an Assistant Professor in the Stewart School of Industrial and Systems Engineering at the Georgia Institute of Technology who received her Ph.D. in Industrial Engineering and Management Science from Northwestern University. Her research interests are in integrating methods from optimization and economics to solve problems in supply chain management and healthcare policy. Her teaching interests include these areas as well as improving ethics in engineering curricula. She is the faculty advisor for the Georgia Tech student chapter of IIE.

Disease	Odds Ratio	Weights	Prevalence (SE)
Intercept	15.01		
Arthritis	0.65	1.54	21.80 (0.71)
Asthma or bronchitis	0.56	1.79	6.12 (0.31)
Back or neck problem	0.47	2.13	8.76 (0.40)
Blood Pressure/Hypertension	0.57	1.75	23.19 (0.84)
Body Mass Index	0.96	1.04	30.17 (0.93)
Cancer	0.59	1.69	8.09 (0.41)
Cardiovascular	0.42	2.38	6.93 (0.42)
Depression	0.59	1.69	9.35 (0.48)
Diabetes	0.34	2.94	6.68 (0.33)
Emphysema	0.21	4.76	1.64 (0.17)
Flu/Pneumonia/Ear Infection	0.89	1.06	4.81 (0.33)
Hearing	0.48	2.08	1.42 (0.16)
Hepatitis C	0.59	1.69	2.02 (0.19)
Poor Oral Health	0.46	2.17	34.75 (1.12)
Stomach/Intestinal	0.81	1.23	9.24 (0.48)
Stroke	0.49	2.04	2.46 (0.20)
Thyroid condition	0.79	1.27	4.94 (0.29)

Table 1: Association between health conditions and self reported general health (Logistic Regression, all significant at the 5% level), weights, and prevalence

	COSTS					CAPACITY		
	Staffing	Equipment/ Supplies	Lab/ X-ray	Admin/ Pharmacy/ Enabling	TOTAL COST	Small	Med.	Large
General	\$37.38	\$10.92	\$4.78	\$16.98	\$70.06	8000	30000	70000
OBGyn	\$37.38	\$10.92	\$4.78	\$16.98	\$70.06	2100	4200	8400
Dental	\$48.50	\$16.14	\$6.20	\$22.04	\$92.88	1320	3960	6600
M/SA	\$26.64			\$9.37	\$36.01	300	1000	3000

Table 2: Variable costs per encounter and capacity levels for each service type

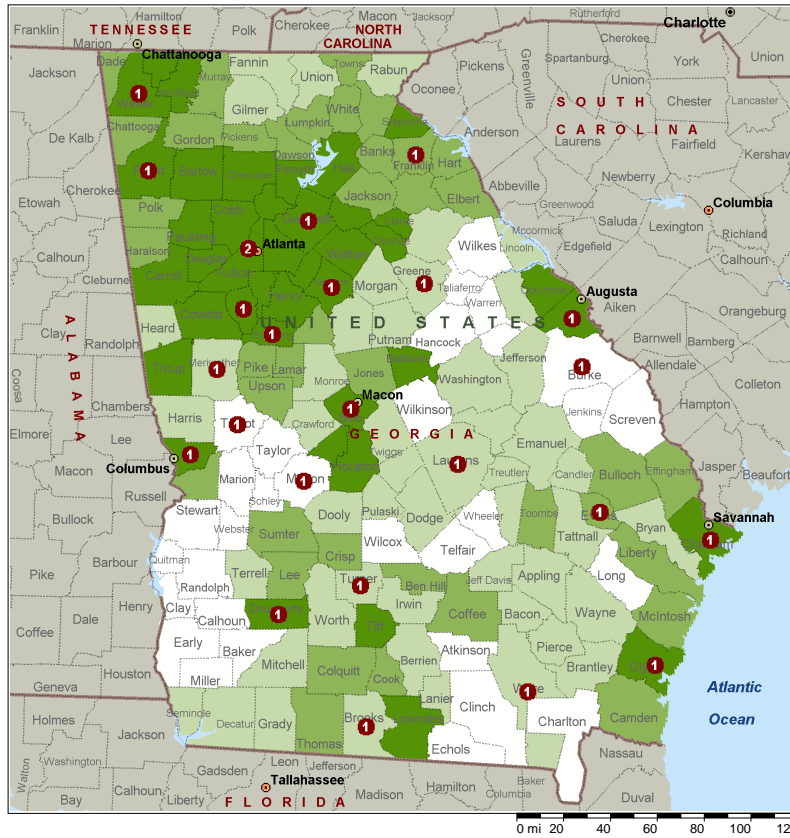


Copyright © 1988-2001 Microsoft Corp. and/or its suppliers. All rights reserved. <http://www.microsoft.com/mappoint>
 © Copyright 2000 by Geographic Data Technology, Inc. All rights reserved. © 2000 Navigation Technologies. All rights reserved. This data includes information taken with permission from Canadian authorities © Her Majesty the Queen in Right of Canada.
 © Copyright 2000 by CompuSearch Micromarketing Data and Systems Ltd.

Figure 1: Current Georgia CHC locations and uninsured persons quartiles by county (darker is more uninsured persons)

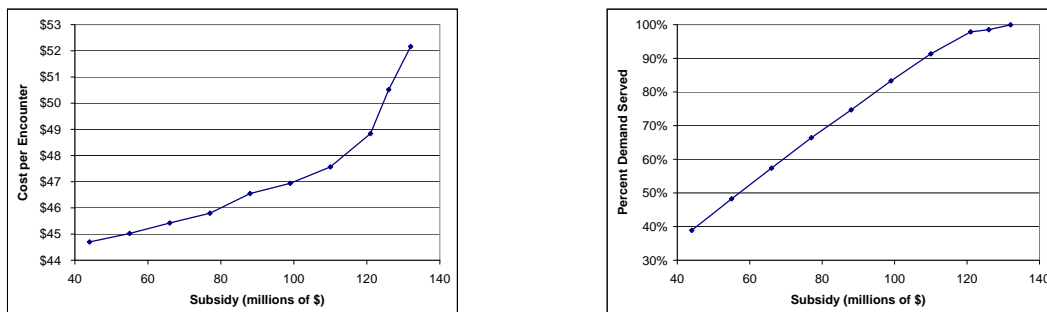
	(A) Current	(B) MUA Constrained	(C) Not MUA Constrained	(D) Current, Constrained
Total encounters	803,106	980,858	988,173	984,352
Cost per encounter	\$ 54.79	\$ 45.73	\$ 44.53	\$ 44.69
% of total demand served	31.7%	37.9%	39.0%	38.9%
Number CHC locations	22	27	25	31
Number offering general	22	22	24	26
Number offering OBGYN	11 (all offer referral)	20	24	26
Number offering Dental	13	16	17	15
Number offering M/SA	12	23	25	31
Number uninsured served (general visits)	35000	44266	44620	42660
Number people served with new place for primary care (general visits)	40791	49420	48958	46890

Table 3: CHC solutions for Georgia



Copyright © 1988-2001 Microsoft Corp. and/or its suppliers. All rights reserved. <http://www.microsoft.com/mappoint>
 © Copyright 2000 by Geographic Data Technology, Inc. All rights reserved. © 2000 Navigation Technologies. All rights reserved. This data includes information taken with permission from Canadian authorities © Her Majesty the Queen in Right of Canada.
 © Copyright 2000 by CompuSearch Micromarketing Data and Systems Ltd.

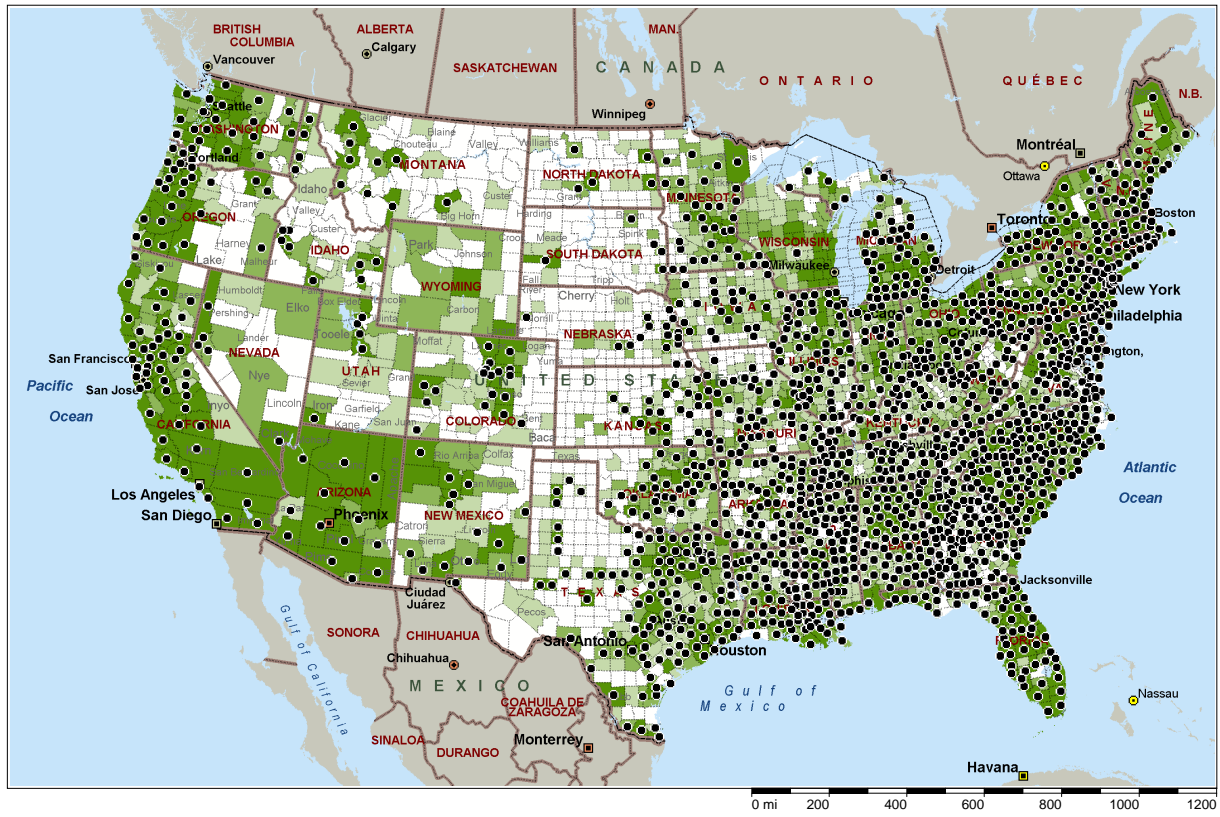
Figure 2: Optimal Georgia CHC locations when not MUA-constrained with population quartiles by county (darker is more populated)



(a) Cost per encounter

(b) Percentage of total estimated demand served

Figure 3: Effect of increase in government subsidy of CHC budget



Copyright © 1988-2001 Microsoft Corp. and/or its suppliers. All rights reserved. <http://www.microsoft.com/mappoint>
 © Copyright 2000 by Geographic Data Technology, Inc. All rights reserved. © 2000 Navigation Technologies. All rights reserved. This data includes information taken with permission from Canadian authorities © Her Majesty the Queen in Right of Canada.
 © Copyright 2000 by CompuSearch Micromarketing Data and Systems Ltd.

Figure 4: National solution with population quartiles by county