Title:
An Assessment of the Feasibility of Call Center Data for Syndromic Surveillance

Type of Research Project:
T4, “Preliminary Study Grant”

Applicants:
Douglas W. Roblin, PhD (PI)
   Professor, Health Management and Policy, School of Public Health, Georgia State University
   Affiliate Investigator, Kaiser Permanente Georgia
   Adjunct Professor, Health Policy and Management, Rollins School of Public Health, Emory University
   droblin@gsu.edu

Lisa M. Casanova, PhD
   Assistant Professor, Environmental Health, School of Public Health, Georgia State University
   lcasanova@gsu.edu

Christina Hemphill Fuller, ScD
   Assistant Professor, Environmental Health, School of Public Health, Georgia State University
   cfuller@gsu.edu

Frederic J. Grant IV, PhD
   Chief Administrative Officer, School of Public Health, Georgia State University
   Associate Professor, Goizueta School of Business, Emory University
   fgrant@gsu.edu

Hongyuan Zha, PhD
   Professor, College of Computing, Georgia Institute of Technology
   zha@cc.gatech.edu

Alexander Isakov, MD, MPH
   Executive Director, CEPAR
   Associate Professor, Emergency Medicine, Emory University School of Medicine
   aisakov@emory.edu

Melissa Butler, PhD, PharmD
   Assistant Investigator, The Center for Research / Southeast, Kaiser Permanente Georgia
   Melissa.G.Butler@KP.Org

Requested Award Amount (Direct Costs):
$23,478
Abstract:
Conceptual and statistical models for assessing the presence and scale of disease outbreaks – that is, models of “syndromic surveillance” – have relied on evaluation of rates of emergency room and hospital admission rates. These models suffer from limitations that might affect conclusions about the timing, magnitude, and geographic scope of a disease outbreak. Nurse call centers provide a complementary data source to improve estimates of the timing, magnitude, and scope of a disease outbreak. When first experiencing symptoms – such as respiratory or gastrointestinal distress – that might represent a disease outbreak, most adults consult with staff of a nurse advice center before presenting to an emergency room or hospital for evaluation and treatment. Syndromic surveillance using the British National Health Service nurse advice program (NHS Direct) has been investigated for almost 15 years. Little equivalent research has been conducted in the U.S.

In this ACTSI “Preliminary Study Grant”, we propose to assemble a rich, innovative dataset for the period 2005 through 2013 that can be used to characterize patterns of nurse advice center calls, emergency room visits, and hospital admissions among Kaiser Permanente patients in the metropolitan Atlanta area in response to generally naturally occurring environmental stressors that provoke syndromic responses – such as respiratory responses to pollen season and both respiratory and gastrointestinal responses to influenza season. We will conduct preliminary spatio-temporal analyses of responses to these environmental stressors, both in the population overall and in selected subsets (e.g. adult vs. pediatric patients, patients with comorbidities such as asthma and CHF that might exacerbate a response). We will use findings from this study as preliminary data to support a larger NIH or NSF research application.
An Assessment of the Feasibility of Call Center Data for Syndromic Surveillance

A. Introduction
Following the World Trade Center attack, the United States government invested heavily in funding projects for development of models for surveillance of public health disasters that might be associated with bioterrorism. Much of that effort has been focused on identification of sentinel event rates using physician visit, emergency department (ED) visits, or hospital admissions for syndromic-related medical conditions.1-4

There are some obvious reasons for relying on these types of data for syndromic surveillance. First, ED visit and hospital admission rates which rapidly increase relative to adjusted historic rates represent a possible acute outbreak of a disease. Second, virtually all areas of the United States are covered by some institutions which provide emergent or acute care. Third, these acute events are captured in hospital electronic medical records when the patient presents for emergent treatment and, if needed, is subsequently admitted for hospital care. Fourth, emergency room personnel are sensitive to the general patterns of medical conditions with which patients have historically presented for emergent care, and, therefore, can quickly recognize and alert public health officials to atypical conditions or volumes. Thus, reports from ED personnel that suggest an unanticipated increase in specific types of presenting conditions can provide timely indication of a significant outbreak of an untoward medical crisis.

There are also some significant disadvantages to relying on rates of patients presenting for acute treatment as sentinel events of a disease outbreak. This approach assumes that 1) the syndrome is manifest through acute symptoms and 2) the rate of acute events represents the scale and scope of the outbreak. Both assumptions may be inherently wrong and impede an appropriate public health response in time or in space. First, many disease outbreaks may be manifest through sub-acute signs and symptoms such as respiratory complaints (e.g. cough or wheezing) and gastrointestinal complaints (e.g, nausea, vomiting, or diarrhea) – particularly if the disease’s prodromal phase is prolonged and mild before onset of an acute phase. Many patients with these complaints may not report for emergent care but may obtain care elsewhere (e.g. primary care visits, urgent care center visits). Second, many patients who present with sub-acute complaints initially may never present for acute treatment – for reasons that might be due to access (travel time), insurance coverage (copayments), susceptibility, or other factors such as presence of comorbidities.

B. Objective and Specific Aims
Thus, we believe that reliance of syndromic surveillance solely on “spikes” in rates of acute events may put public health response at risk. Acute event reports may be untimely as to the initiation of an outbreak, and may lag identification of the outbreak by several days. Acute event reports may under-represent the scale or scope of an outbreak. If comorbidities or access to acute treatment affect the likelihood that a patient presents for acute treatment, then the initial rates of acute treatment may significantly under-estimate the magnitude or geographic spread of an epidemic. Deployment of appropriate and adequate resources in response to a potentially evolving epidemic, whether unintended (nature) or intended (terrorism), requires time. But adequate response must be timely. The sooner an epidemic, its character, location, and geographic scope can be identified, the sooner an appropriate deployment and medical response can be initiated. Nurse call centers appear to have great potential to accelerate 1) the timeliness with which the outbreak of an epidemic can be identified, 2) the character of the outbreak, and 3) the geographic locations and spread of an outbreak.

Nurse hot lines and related call center services are already in place throughout the United States (US) and are integral to routine health care delivery. Virtually all adults in the US have access to a telephone and routinely begin their interaction with a health care system when a person – or those under the care of the caller – begin to experience untoward signs and symptoms, such as respiratory, dermatologic, or gastrointestinal complaints. Unlike ED visits and hospital admissions, access is not limited to provision of services at a finite number of locations that require time and transportation to obtain—even for information and guidance; and, unlike many other health care services, telephone advice and triage is provided without patient cost-sharing, which should further promote access.

Beginning in the 1990s, health services researchers in the United Kingdom (UK) recognized the potential surveillance benefits of monitoring patterns of signs and symptoms provided by patients when calling “NHS Direct”, a telephone health advice services system that is part of the UK’s National Health Service (NHS). Studies have explored patterns in call center data for monitoring of seasonal trends in respiratory and
gastrointestinal symptoms indicative of timing and spread of influenza\textsuperscript{5-15}, gastrointestinal symptoms in response to contamination of food or drinking water supplies\textsuperscript{16-18}, and the public health impact of natural events with potential to exacerbate syndromic response.\textsuperscript{19} To date, only one similar study has been conducted in the US\textsuperscript{20}, and that study examined “trends” over only one year and without linking patients back to their residential areas to monitor spread of influenza.

The overall objective of our proposed “Preliminary Study Grant” is to assess the feasibility of call center records as a complement to other established databases (e.g. ED visits and hospital admissions) for estimation of the timing, magnitude, and geographic scope of disease outbreaks. An innovative longitudinal database will be constructed from data collected during routine provision of medical care at Kaiser Permanente Georgia (KPGA), a group-model HMO which currently provides health care to approximately 235,000 residents of the metropolitan Atlanta area. The specific aims of our proposal are to:

1) Build a retrospective, observational database on KPGA members from 2005-2013 that links basic demographic and residential area information to call center records, visit records, ED records, and hospital admission records related to respiratory and gastrointestinal symptoms.
   a. Develop an “exposures” matrix for the Atlanta metropolitan area, defining timing and location of seasonal (pollen, influenza) and atypical (smog days, “boil water advisories”) events for the period 2005-2013.

2) Compare rates of the proportion of call center records, visit records, and ED visit/hospital admission records for respiratory and gastrointestinal symptoms during the period 2005-2013 by time and by space for estimation of the magnitude and geographic spread in response to seasonal cycles of pollen and influenza.
   a. Evaluate the moderating effects of comorbidities (e.g. asthma/COPD and cardiovascular disease) and access (e.g. travel distance to ED/hospital) on daily rates of call center records, ambulatory visit records, and ED visit/hospital admission records.

3) Assess the potential for developing a large multi-site NIH or NSF application and, if feasible, develop a plan for development of this research application.

C. Approach
C.1. Study Setting
KPGA is one of eight regions of the Kaiser Permanente Medical Care Program, a federally qualified, prepaid group health maintenance organization and the largest non-governmental health care provider in the world. Incorporated in 1985, Kaiser Georgia currently provides comprehensive medical care services to a diverse cross-section of residents of the metropolitan Atlanta area. The membership reflects the racial and socioeconomic diversity of Atlanta. Approximately 55% of the adult membership is Caucasian and 40% is African American.

The Kaiser Georgia Call Center is available to HMO members 24 hours a day, 7 days a week. The call center assists members by scheduling or canceling appointments, providing nurse advice, and accessing after hours care services. When a member calls the Call Center, the service representative verifies the member’s eligibility and ascertains the reason for the call (e.g. chief complaint or symptoms). During regular office hours (Monday through Friday, generally 8 am to 6 pm), the call is then routed directly to the member’s primary care team. The team’s advice nurse then ascertains the member’s status and works with the patient to schedule a visit with the team (approximately 80% of appointments are scheduled for a time within 24 hours of the call) or to obtain other appropriate services. In recent years, the Call Center has logged approximately 2,000,000 calls per year. Each year, there are approximately 500,000 completed primary care visits among adult patients and 250,000 among pediatric and adolescent patients.

C.2. Study Data
KPGA’s Research Department maintains a datamart for purposes of supporting research projects requiring computerized records. All data are maintained in SAS datasets. These datasets are populated from KPGA’s electronic medical record system and its computerized business system, which collectively contain all files used for purposes of tracking enrollment eligibility, insurance benefits, and medical services (primary care, specialty care, emergency and hospital care, outpatient pharmacy dispensings, laboratory tests and results,
radiology services). Records of medical care services include ICD-9 coded diagnoses, ICD-9 and CPT-4 coded services, and the dates, locations, and providers of the services. The datamart also contains a set of disease “registries” (e.g. diabetes, CHF, asthma/COPD) that estimate an approximate “onset date” from assessment of ICD-9-CM diagnosis codes, pharmacy dispensing, and laboratory test results.

In 2005, KPGA implemented the EpicCare medical record system throughout the group-model primary care practices. This electronic medical record serves as the main portal by which primary care services are now entered into KPGA's main computerized datawarehouse. Authorized KPGA staff have access to legacy medical record and computerized business system databases.

The advice nurse call records in the EMR contain the enrollee identifier, date of the call, and nurse assessment of reason for the call. The assessments are coded in structured text regarding reason for the call, for example "URI/HARSH COUGH", “VOMITING/DIARRHEA”.

Primary care visit records in the EMR contain the enrollee identifier, date of visit, and reason for the visit. Reason for the visit is represented by ICD-9-CM diagnosis codes. ED visits and hospital admissions are primarily provided at 3 primary contract hospitals in the metropolitan area (Northside, Piedmont, Gwinnett). Records from those hospitals – or any other area facilities which submit bills for services – are captured in a claims database. Claims for service contain the enrollee identifier, dates of service, and ICD-9-CM codes associated with the services provided (which are, de facto, reasons for the service).

The KGPA Research Department retains residential addresses as of the end of each year. Residential address is geocoded using MapInfo Professional v. 8.5 (MapInfo, Troy, NY) to the exact latitude and longitude. Residential records contain, in addition, the enrollee identifier.

All databases can be linked by the unique enrollee identifier; and, events can be sequenced by event dates. After linking study datasets by enrollee identifiers, we will de-identify the final analysis dataset – first, by substituting a unique study identifier for the enrollee identifier, and second, by modifying latitude and longitude for each enrollee by a random perturbation to latitude and longitude of approximately ± 1000 feet.

C.3. Measurement and Methods
C.3.1. Aim 1: Build a retrospective, observational database on KPGA members from 2005-2013 that links basic demographic and residential area information to call center records, visit records, ED records, and hospital admission records related to respiratory and gastrointestinal symptoms

Under the direction of the PI, the KPGA programmer will extract all call center records for the period 2005-2013. Frequency distributions will be developed for the reasons for the call. The investigator team will review those frequencies and identify the reasons most relevant to 1) respiratory responses to high pollen days or smog alert days, and 2) gastrointestinal responses to influenza season and to contamination of drinking water supplies.

Under the direction of the PI, the KPGA programmer will extract all primary care visit records for the period 2005-2013. A primary care visit record will be defined as any face-to-face visit in the departments of pediatrics, adult medicine, OB/Gyn, provided in a KPGA clinic. The KPGA programmer will also extract all urgent care visits (i.e. “after hours” visits) provided at selected KPGA facilities, ED visits, hospital observation stays (HOBS), and hospital admissions from the KPGA research datamart (which aggregates claims into utilization events). Frequency distributions will be developed on the ICD-9-CM diagnosis codes that represent the reasons for these event types. The investigator team will review those frequencies and identify the reasons most relevant to 1) respiratory responses to high pollen days or smog alert days, and 2) gastrointestinal responses to influenza season or to contamination of drinking water supplies.


Under the direction of the investigator team, the GSU graduate research assistant (GRA) will develop a week by week matrix in order to identify the general times of onset of different types of events that might affect temporal trends and intensity of population responses to respiratory or gastrointestinal triggers. Those triggers include timing of pollen seasons in the Atlanta area, smog alert days, “boil water” advisories, and CDC and KPGA announcements on “flu season”. The matrix will be linkable to the datasets being generated at KPGA.
Sources of information for developing this matrix will include pollen counts from the Atlanta Allergy Center, National Weather Service records, CDC bulletins, etc.

Under the direction of the PI, the KPGA programmer will create a de-identified dataset at the patient-month level of observation. Upon approval of a data transfer agreement, the de-identified dataset will be transferred from KPGA to the PI for analysis.

C.3.2. Aim 2: Compare rates of the proportion of call center records, visit records, and ED visit/hospital admission records for respiratory and gastrointestinal symptoms during the period 2005-2013 by time and by space for estimation of the magnitude of disease outbreak and geographic.

Under the direction of the PI, the GSU GRA will initially generate a series of tables and graphs that describe:

1) Daily volume of calls by weekday for comparing key “trigger months” (e.g. high pollen months, influenza season months) with other months. Calls will be subset to relevant events (e.g. respiratory symptoms for high pollen season).

2) Daily volume of primary care visits by weekday for comparing key “trigger months” (e.g. high pollen months, influenza season months) with other months. Primary care visits will be subset to relevant events (e.g. respiratory symptoms for high pollen season).

3) Daily volume of ED visits/hospital admissions by weekday for comparing key “trigger months” (e.g. high pollen months, influenza season months) with other months. Calls will be subset to relevant events (e.g. respiratory symptoms for high pollen season).

In addition, the GSU GRA will generate maps, using SAS Graph functions, to display the general locations and spread of these 3 types of events in the metropolitan Atlanta area.

The primary dependent variables will be: 1) the proportions of calls, primary care visits, ED visits, and hospital admissions per day attributable to respiratory disease, and 2) the proportions of calls, primary care visits, ED visits, and hospital admissions per day attributable to gastrointestinal disease. Following the methods developed in analysis of the NHS Direct data 5, 7, 9, 10, we will estimate time-series models that adjust for weekday, holidays, season, and membership population size – all of which might affect spatiotemporal patterns in the distributions of these dependent variables. Preliminary analyses by Dr. Roblin suggest these factors – as also found in the NHS Direct data – require adjustment (Figures 1 and 2).

Analyses will include estimation of 95% confidence intervals. To the extent possible, we will also compute an estimate of “exceedance” – a measurement concept developed in analysis of the NHS Direct data which indicates whether a season or event might generate a syndromic response that statistically exceeds historic rates adjusted for weekdays, holidays, and season.
C.3.2.a. Evaluate the moderating effects of comorbidities (e.g. asthma/COPD and cardiovascular disease) and access (e.g. travel distance to ED/hospital) on event rates.

Under the direction of the PI and advice of the investigator team, the GSU GRA will replicate the analyses described in Aim 3 to relevant subsets of patients. For example, we suspect that adults with asthma/COPD and CHF will have a more dramatic response in terms of respiratory symptoms (as represented in calls, visits, or acute events) than other adults. We will also contrast differences in “straight line distance” between residence and nearest ED/hospital facility and rates of calls, visits, and acute events – under the presumption that distance will have no effect on call rates but will dramatically affect acute event rates. If both of these effects are observed to influence dependent variable rates (comorbidities and distance affect acute event rates), then this will provide prima facie evidence that use of acute event rates are biased toward limiting the sense of timing and geographic scope of an outbreak toward higher rates in communities with higher rates of comorbidities and better access to acute care facilities.

D. Anticipated Results
We anticipate two products from this “Preliminary Study Grant”. The first product is a set of data tables and graphs which can be used as preliminary data in support of an NIH application, either an R21 or an R01, or an NSF application. For an R21, we would request further funding to better model the dataset collected through this grant. For an R01, we would propose a multi-site study that would better incorporate development of a “trigger matrix” (i.e. a site specific space-time dataset to describe the onset and intensity of a trigger set, such as pollen count) and its relation to the outcomes that we are studying in this proposed project. That type of grant would help to better establish a “dose-response” comparison in the use of complementary datasets (i.e. call center records, primary care visit records, and acute event records) for understanding the magnitude and spread of a disease outbreak.

In addition, depending on the results of our proposed scope of work, we anticipate the possibility of one or more manuscripts that could be submitted to peer reviewed journals. Given the absence of any work to date that compares the potential use of call center data to more traditional data (ED records, hospital admission records) to support syndromic surveillance, we believe that our work will be sufficiently significant and innovate to warrant publication.

E. Project Management and Timeframe
E.1. Project Management
The project will be directed by Douglas Roblin, PhD. The investigator team will consist of Dr. Roblin and additional senior investigators with expertise in statistical computing (Dr. Zha), emergency medicine (Dr. Isakov), environmental health (Drs. Casanova and Fuller), medical informatics (Dr. Grant), and the Kaiser health care system (Dr. Butler). Dr. Roblin will have primary responsibility for oversight of the KPGA subcontract to acquire the needed data and supervision of the GSU GRA to analyze the data. The investigator team will have 4 face-to-face meetings and monthly conference calls to advise on study design and to review and interpret study results.

E.2. Project Timeframe

<table>
<thead>
<tr>
<th>Tasks</th>
<th>Project Month</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Investigator Face-to-Face Meetings</td>
<td>X</td>
</tr>
<tr>
<td>Investigator Conference Calls</td>
<td>X</td>
</tr>
<tr>
<td>IRB Submissions</td>
<td></td>
</tr>
<tr>
<td>1. Assemble KPGA Call Center Data</td>
<td>X</td>
</tr>
<tr>
<td>2.a. Assemble KPGA Visit Data</td>
<td></td>
</tr>
<tr>
<td>2.b. Assemble KPGA ED/Admission Data</td>
<td></td>
</tr>
<tr>
<td>2.c. Develop Exposure Matrix</td>
<td></td>
</tr>
<tr>
<td>2.d. Transfer Data to GSU</td>
<td></td>
</tr>
<tr>
<td>3. Tabulate and Graph Data Trends</td>
<td></td>
</tr>
<tr>
<td>4. Tabulate and Graph by Patient Groups</td>
<td></td>
</tr>
<tr>
<td>Write Final report / Draft R21/R01</td>
<td></td>
</tr>
</tbody>
</table>

7
References


