CDPRG Internal Report

Health system surge capacity in developing countries: evaluating the feasibility of estimating baseline health resource demand from ICD diagnosis codes for surge response modeling

Sara U. Schwanke Khilji, MD, MPH

23 January 2012; updated 8 November 2012

Introduction

Avian flu, the H1N1 pandemic of 2009, increasingly frequent natural disasters (UNEP/GRID-Arendal 2005): in recent years, emergencies of these sorts have repeatedly revealed how health system capacities worldwide can be stressed, sometimes severely, by catastrophe. In recognition of this challenge, a relatively new body of literature is developing around the concept of “surge” and its related concepts of “surge capacity” and “surge response capability.” Indeed, some authors have claimed that surge capacity “is arguably one of the most important areas of research endeavor for catastrophic events” (Rothman, Hsu et al. 2006).

As noted elsewhere, the literature surrounding surge and its related concepts is complicated by a variety of competing definitions and metrics (Handler, Gillam et al. 2006). For simplicity, we will adhere to the definition of surge as put forth by Rothman *et al.,* defining it as a significant increase in resource demand compared to baseline; in the health care setting, this specifically entails increased demand for public health and medical resources. Surge implies a number of component variables that determine the magnitude of increase in demand, including event-specific characteristics (type, duration, scale), resource demand provoked by the event, and the volume and rate of demand (Kelen and McCarthy 2006). In distinction to “daily surge” (also referred to as “routine surge” or “regular surge”, e.g. in the case of predictable increased resource demand in emergency room settings), we will focus on “extraordinary surge” (also known as “disaster surge,” “catastrophic surge,” or “critical event surge”). The related concept of “surge capacity” is similarly subject to definitional disagreement. The United States Centers for Disease Control and Prevention (CDC), for instance, defines surge capacity as the “ability to obtain additional resources when needed during an emergency” (CDC 2005), while the American College of Emergency Physicians defines it as “a measurable representation of a health care system’s ability to manage a sudden or rapidly progressive influx of patients within the currently available resources at a given point in time” (Handler, Gillam et al. 2006). In recognition of this confusion about the meaning of surge capacity, Kelen *et al.* have introduced a third concept, “surge response capability,” to describe the goodness of fit between existing resources and those demanded by surge. Defining surge capacity as “the maximum potential delivery of required resources, either through augmentation or modification of resource management and allocation,” surge response capability, then, is “the ability of surge capacity to accommodate the surge” (Kelen and McCarthy 2006). It is this definition of surge capacity that will be assumed throughout this paper.

The capacity demanded by surge events can be conceptualized as comprising four essential components: staff (i.e. trained personnel), "stuff" (medical equipment, supplies, and pharmaceuticals), structure (facilities), and systems (policies and procedures integrated both vertically and horizontally) (Barbisch and Koenig 2006). Infrastructure, comprehensive supplies, and expertise, however, remain critically lacking in many developing countries (Oshitani, Kamigaki et al. 2008), often to the point that resource gaps occur even in the absence of surge events. Thus, surge capacity planning can be integrated as a key component for core health system strengthening and disaster preparedness planning; careful analysis of existing resources, from stuff to systems, is essential to these goals.

Despite the fact that developing countries are disproportionately impacted by surge situations, both in terms of direct effects of surge events and the indirect consequences of surge on already limited healthcare capacities (Krumkamp, Kretzschmar et al. 2011; Rudge and al. Unpublished), most of the empirical research to date on surge capacity has focused instead on high-income countries. Data reflecting extant resources, resource needs, and related resource gaps are of paramount importance for supporting developing countries in disaster preparedness planning (Krumkamp, Kretzschmar et al. 2011). Indeed, others have noted that it is not completely self-evident that developing countries should attempt to replicate the highly specialized emergency preparedness programs of the developed world, as chronic under-investment and public health priorities compete for scarce resources (Kruk 2008). There is, then, a pressing need for the development of analytical and conceptual frameworks to assess existing health system capacity in developing countries, in order to identify and address resource gaps under various surge scenarios while at the same time seeking to build core capacity (Kruk 2008; Rudge and al. Unpublished).

While simulation exercises have been created to evaluate adequacy of resources in the setting of pandemic control, many lack any clear consideration of resource availability (i.e. whether they are already in use for regular operations, whether they can be mobilized from other areas/facilities) (Krumkamp, Kretzschmar et al. 2011) or the public health impact of re-allocating resources under surge conditions. Thus, the capacities of health systems in developing countries to meet either normal demand or surge demand remains unclear; the public health implications of surge on health system functioning remains even more poorly defined.

CDPRG's current project, "Surge in demand for health services: evaluating health system impact and capacity to respond in countries with limited resources" was conceived within this context of the challenges surrounding surge capacity assessment and improvement in developing countries. As a region, Southeast Asia is a prime location for surge capacity research due to the diversity of both surge events and health care systems. Indonesia, in many respects, represents a microcosm of the particular difficulties Southeast Asia faces in preparing for surge events: it has been repeatedly subjected to severe infectious and geologic catastrophes in the past decade, while its size, heterogeneous population and decentralized health care system are emblematic of the diversity in the region at large. As such, Indonesia was selected as the focus of the current surge capacity research, which seeks to first develop a new conceptual and analytical framework for evaluating health system surge capacity and then to pilot this framework in the pilot sites of West Sumatra and Bali. The ultimate aim of the project is to provide information about health system surge capacity in resource poor settings that accounts for patterns of baseline health care demand, supply availability, and existing unmet needs, thus enabling health system planners to better understand existing capacity and predict resource mobilization needs.

In developing this framework for surge capacity analysis, both quantitative and qualitative data will be collected in order to address the necessary resources of staff, supplies, structure, and systems. As supplies are the most easily quantified, this will be the starting point for data collection, including data on health service utilization (outpatient visits, inpatient occupancy rates, surgical operation activities) at health facilities as well as health service utilization information based on medical records by International Statistical Classification of Diseases and Related Health Problems, 10th Revision (ICD-10) classifications. The aim of this data collection activity is to obtain utilization statistics by common diagnosis type (ICD- 10) for health care facilities at all levels; these diagnosis codes will subsequently be classified into utilization groups based on a set of clinical patient triage protocols. Ideally, current medical supply usage represented by ICD-10 codes could then be linked to each of the most common diagnosis types, resulting in a database of resources currently in use for baseline treatment functions. This, when compared with information about overall available medical supplies at the institutional and provincial levels, will yield valuable information about baseline capacity, both in terms of spare capacity as well as resources currently in use for regular health care activities which may potentially be used to improve surge capacity through resource allocation from non-essential or elective care.

Although the ICD is considered to be the international standard of diagnostic classification for epidemiological use (WHO 2012), very little health resource utilization research has been carried out using ICD codes in developing countries. Multiple studies have advocated for improved health resource allocation (typically at the national level) through the integration of morbidity data reflected by ICD coding (West 1978; Goldacre and Harris 1980; Andersson, Varde et al. 2000; Asthana, Gibson et al. 2004; Dixon, Smith et al. 2011). Research attempting to link ICD codes to actual resource utilization, however, has been limited almost exclusively to developed countries, primarily the United States. This paper focuses on determining whether ICD codes can be reliably used to gather health utilization data and, if so, how this might be applied in our research and in the developing world, more broadly.

Methods and Results

Review of the literature was performed using the PubMed biomedical database, comprising literature from MEDLINE, life science journals, and online books. Articles were reviewed in full whenever available.

Resource allocation and ICD
The initial search terms employed were “resource allocation” [All Fields] AND ICD [All Fields], yielding 32 candidate articles. Six of these were deemed relevant in that they explicitly linked diagnoses (determined by either ICD-9 or ICD-10, depending on the study) to resource utilization in an effort to improve resource allocation (West 1978; Goldacre and Harris 1980; Mugisha, Kouyate et al. 2002; Simon, Hirsh et al. 2009; Dixon, Smith et al. 2011; Lopez-Bastida, Moreno et al. 2012). Of these, four focused simply on the question of refining current allocation practices to better reflect

morbidity and/or mortality, either at the individual (Dixon, Smith et al. 2011) or population (West 1978; Goldacre and Harris 1980; Simon, Hirsh et al. 2009) level. While the analyses by West, Goldacre and Harris, and Dixon *et al.* all aimed to improve appropriateness of NHS funding through the incorporation of ICD codes into health funding formulas, Simon *et al.* focused on the subpopulation of pediatric emergency department (ED) visitors in the United States to identify individuals with high frequency of unnecessary ED utilization. The report by Lopez-Bastida *et al.* used ICD-9 discharge codes for stroke as the basis for a prevalence-based cost-of-illness analysis encompassing direct healthcare costs, direct non-healthcare costs, costs due to loss of productivity, and patient outcomes. Most pertinent to our purposes, Mugisha *et al.* used ICD-9 codes to estimate costs of health interventions at health facilities in Nouna, Burkina Faso for 33 common diseases by utilizing algorithms for diagnosis and treatment from the Burkina Faso Ministry of Health. According to the published abstract, the study extracted data from four "first line" health facilities in Burkina Faso using a demographic surveillance System (DSS). A cost estimate unit was developed to compare variations in costs across the diagnosis cohort, revealing a range from most costly (case management with hospitalization) to least costly (family planning) interventions. Unfortunately, no electronic version of the full article was available for review via PubMed.

To determine whether the search results were substantively broadened through inclusion of MeSH terms, the search was repeated with the terms "research allocation" [All Fields] AND ICD codes [All Fields], thus triggering Automatic Term Mapping and yielding full search criteria of "resource allocation"[All Fields] AND ("international classification of diseases"[MeSH Terms] OR ("international"[All Fields] AND "classification"[All Fields] AND "diseases"[All Fields]) OR "international classification of diseases"[All Fields] OR ("icd"[All Fields] AND "codes"[All Fields]) OR "icd codes"[All Fields]). Interestingly, the total number of results was comparable (29 articles), but this search returned a number of additional articles not captured by the initial search (15 total) while at the same time omitting 17 articles captured by the original search; of note, the addition of these MeSH terms resulted in the omission of the article by Mugisha *et al*. None of the new articles identified by the modified search were notably different from those reported above, and, more pertinently, none focused on health resource allocation in a developing country.

A final search modifiying these search terms was completed, this time with the addition of MeSH terms for resource allocation, as well (Automatic Term Mapping: "resource allocation"[All Fields] OR "resource allocation"[MeSH Terms]) AND ("international classification of diseases"[MeSH Terms] OR ("international"[All Fields] AND "classification"[All Fields] AND "diseases"[All Fields]) OR "international classification of diseases"[All Fields] OR ("icd"[All Fields] AND "codes"[All Fields]) OR "icd codes"[All Fields]). This search yielded a total of 36 articles: 29 articles were identical to those obtained by the search '"resource allocation" AND ICD codes,' with seven additional articles identified. Of these seven, only one (focused on the policy implications of preventing mother to child HIV transmission in Uganda) initially appeared relevant based on review of titles; however, when the full report was obtained, it became clear that the research did not utilize ICD coding.

"Resource utilization" and ICD

A second search was performed via PubMed with an alteration in the search terms to "resource utilization" [All Fields] AND ICD [All Fields]. This search resulted in 118 candidate articles; articles which linked ICD to “implantable cardiac defibrillator” were excluded, as were articles that focused exclusively on resource utilization as a composite cost quantified by US dollars (i.e. rather than breaking down utilization into relevant components, e.g. pharmaceuticals, transfusions, bed occupancy days, etc.) or were considered excessively clinically complex (e.g. determining differences in healthcare costs between patients on different chemotherapeutic regimens); additionally, articles written in languages other than English were not reviewed. After applying these exclusion criteria, 32 articles were considered relevant to the question at hand. The articles were reviewed in the order presented (i.e. in reverse chronological order), highlighting the fact that the vast majority of such studies have been published quite recently, even after taking into account the fact that PubMed holdings stretch back at most several decades. The oldest relevant reference was from 1989 (Surgenor, Wallace et al. 1989), focusing on the frequency and volume of red blood cell transfusions in patients with ICD-9 codes related to digestive diseases; in contrast, 11 of the candidate articles were published in the 1990s (one of which, by Rutledge & Osler, was considered relevant (Rutledge and Osler 1998), although it explores a model for predicting disease-related costs based on ICD-9 coding rather than analyzing utilization associated with a specific disease and, as such, will be discussed separately; see below) and a total of 80 (27 relevant) were published within the last five years.

The articles share certain key characteristics (see Figure 1, below). First, all are retrospective analyses of some sort (typically cohort or case control). Second, the vast majority of these articles were produced in the United States and, therefore, rely still on ICD-9 rather than ICD-10 codes. Third, many of the studies explore diagnoses with high-resource demands but do not necessarily reflect public health priorities (e.g. fibromyalgia, Dupuytren’s contracture). Fourth, many of the studies are explicitly interested in excess resource use attributable to particular combinations of diseases (i.e. common or expensive comorbidities and complications).

When MeSH terms for ICD codes were added to this search (Automatic Term Mapping: "resource utilization"[All Fields] AND ("international classification of diseases"[MeSH Terms] OR ("international"[All Fields] AND "classification"[All Fields] AND "diseases"[All Fields]) OR "international classification of diseases"[All Fields] OR ("icd"[All Fields] AND "codes"[All Fields]) OR "icd codes"[All Fields]), the results again revealed a mix of additional (56) and omitted articles (44). Review of the additional articles resulting from this additional search similarly showed no difference in their key characteristics, as outlined in the preceding paragraph and the accompanying Figure 1. Specifically, the addition of the MeSH terms did not yield any instances of ICD coding used for research on health resource utilization in developing countries.

A final note: although American spellings are the preferred format in PubMed and for MeSH terms, the possibility that relevant articles might have been missed due to alternative spellings was considered. A subsequent PubMed search using the British spelling of utilization yielded only five additional hits, none of which differed notably from the studies presented in Figure 1 or substantively changed the search results.

Figure 1. Key characteristics of articles identified by search terms "resource utilization" AND "ICD." All studies are US-based and utilize ICD-9 coding, unless otherwise noted.

Author, date	Study design	Data source	Diagnoses	Outcome variables
(Owusu, Liu et al. 2012)	Retrospective database analysis	2009 Kids Inpatient Database	Cleft palate + primary repair	LOS, total hospital charge (by geographic region and teaching status)
(Lopez-Bastida, Moreno et al. 2012) (Spain; ICD-9)	Cross-sectional, retrospective study	Hospital admissions databases and discharge summaries; patient questionnaire	Stroke	Total hospital charges (by DRG), composite outpatient care, number of emergency visits, medications, medical devices, informal care, etc.
(Hung, Lai et al. 2012) (Taiwan; ICD-9)	Retrospective database analysis	Administrative claims data from the Bureau of National Health Insurance	Patients requiring prolonged mechanical ventilation >21 days	LOS, total hospital costs
(Vashishta, Mahalingam-Dhingra et al. 2012)	Retrospective database analysis	Nationwide Inpatient Sample 2009	Thyroidectomy	LOS, total hospital costs
(Crockett, Sperry et al. 2012)	Retrospective database analysis	University of North Caroline hospital databases	Foreign body in esophagus	LOS, medications, procedures, consult fees, total hospital costs
(Collard, Ward et al. 2012)	Retrospective case-control cohort analysis	US insurance claims databases	Idiopathic pulmonary fibrosis	Total direct healthcare costs, hospital costs, outpatient

				costs, medications
(Macaulay, Ivanova et al. 2012)	Retrospective case-control cohort analysis	US private insurance claims databases	Dupuytren's contracture	Comorbidities, ED visits, outpatient visits, inpatient admissions, physical therapy visits, medications, days of work lost, annual direct and indirect costs
(Kollef, Hamilton et al. 2012)	Retrospective case-control cohort analysis	Premier database of hospitals in the US	Ventilator-associated pneumonia	Duration of mechanical ventilation, ICU LOS, inpatient LOS, mean hospitalization costs, mortality
(Udall, Harnett et al. 2012)	Retrospective case-control cohort analysis	MarketScan Commercial Claims and Encounters; Medicare Supplemental Databases	Painful diabetic peripheral neuropathy	Prescription medication costs, total healthcare costs
(Poordad, Theodore et al. 2011)	Retrospective database analysis*	Longitudinal administrative claims database from large US commercial health plan	Severe chronic liver disease (CLD); thrombocyto-penia	CLD-related ambulatory visits, ED visits or admissions; annual CLD-related costs (USD)
(Gore, Tai et al. 2011)	Retrospective case-control cohort analysis	Medical and pharmacy claims data from the LifeLink Database	Osteoarthritis (OA)	Comorbidities; pain-related prescriptions; direct medical costs due to OA (medications, outpatient visits, admissions)

(Patel, Gao et al. 2011)	Retrospective case-control cohort study*	US national data from the Premier Perspective hospital database	Invasive aspergillosis in the setting of chronic obstructive pulmonary disease (COPD)	Total cost of admissions; LOS; intensive care unit utilization; mortality
(Wu, Forsythe et al. 2011)	Retrospective database analysis	Thomson MarketScan databases	Gout	Comorbidity burden; ED visits; admissions; outpatient visits; annual health care costs (USD); "other"
(Sanchez, Uribe et al. 2011)	Retrospective cohort analysis	Humana claims database	Fibromyalgia (FM)	Comorbidity prevalence; utilization and costs of FM-related medications; "health care services"
(Dugar, Lander et al. 2010)	Retrospective database analysis	The Kids' Inpatient Database (KID), a national all-payer database of inpatient stays for children $\leq$20 (2-3 million discharges)	Acute sinusitis	Total hospital charges; LOS; number of diagnoses and procedures; discharge status; comorbidities; geographic distribution
(Bogner, Miller et al. 2010)	Retrospective case-control study*	Academic health care system administrative database; Medicare claims database	Heart failure, primary or secondary; diabetes mellitus	Inpatient costs (including medications); outpatient costs (excluding medications); number of ED visits; number of admissions;

				number of outpatient visits; number of diagnostic/lab procedures
(Dunn and Pill 2009)	Retrospective claims-based analysis	Medical and pharmacy claims from PharMetrics Integrated Patient-Centric Database (>60 million patients)	Osteoarthritis	Annual OA-related health care charges (inpatient and outpatient); medications (narcotics, NSAIDs, steroids, PPIs, H2-blockers)
(Saleh, Fisher et al. 2009)	Retrospective longitudinal cohort study	Managed care administrative claims database	Primary thrombocyto-penia (TCP)	Most frequent TCP-associated treatments (IgGs, steroids, transfusions); medications (antibiotics, anti-hypertensives, analgesics, antidepressants); inpatient costs; ED-associated costs
(Iyer, Saunders et al. 2009)	Retrospective cohort study	Premier Perspective database (inpatient database developed for quality and utilization research; 5 million discharges)	Colectomy-associated post-operative ileus	LOS; inpatient costs. Covariates included patient demographics, mortality risk, disease severity, admission source, payment type, hospital type
(Levin, Chaudhry et al. 2009)	Retrospective cohort	Canadian comp-	Chronic kidney disease (CKD),	Number of physician

(Canada; ICD-10)	analysis*	rehensive administrative billing database at the provincial level (British Columbia)	cardiovascular disease (CVD), diabetes mellitus (DM)	visits; duplicate testing; LOS; medications (ACE-I, ARB, statin)
(Sicras-Mainar, Rejas et al. 2009) (Spain; ICD-10)	Retrospective, multi-center cross-sectional study	HMO database ("public with a private services supply")	Fibromyalgia (FM)	Medications; lab testing; medical visits (all types); referrals; admissions; sick leave; early retirement due to FM-related disability
(Singh and Strand 2009)	Retrospective database analysis	Veterans Integrated Service Network (VISN)13; postal survey; clinical databases	Spondyloarthro-pathies	Comorbidities; outpatient visits; specialty visits; surgical visits; limitations in ADLs
(James, Patel et al. 2008)	Retrospective cohort study*	Discharge abstracts from HCUP-NIS, from AHRQ	Pregnancy-related discharge; obstetrical bleeding-related discharge	LOS; blood transfusions; average total inpatient costs
(Morrison, Patel et al. 2008)	Retrospective cohort study*	Discharge abstracts from AHRQ's HCUP NIS	Heavy uterine bleeding; anemia	Blood transfusions; ED admissions; inpatient costs
(Tencer, Friedman et al. 2007)	Retrospective database analysis	Claims data from PharMetric Patient-Centric database	Hemophilia; hepatitis C (HCV); HIV	Costs for clotting factor; prescription medications; inpatient costs; outpatient costs
(Barron, Quimbo et al. 2008)	Retrospective case control	Claims databases	Breast cancer	All medical and pharmacy

	study	from five US health plans		costs during study year (2004)
(Mirkin, Murphy-Barron et al. 2007)	Retrospective claims analysis	Medstat Marketscan database (HMOs and EPOs excluded to avoid underestimation due to incomplete encounter data)	Endometriosis	Annual admission rate; annual surgical rate; comorbidity prevalence
(Margolis, Barron et al. 2005)	Retrospective claims analysis	Large managed care database of two US health plans	Peripheral arterial disease (PAD), primary or secondary	Clinical complications (myocardial infarction, amputation, etc.); medications; outpatient visits; labs/procedures; ED visits; admissions
(McCollam and Etemad 2005)	Retrospective claims analysis	Unclear from abstract; full article not available	New-onset acute coronary syndrome (ACS)	Revascularization rates; total first year cost of care; inpatient costs; medications (statin, BB, Plavix)
(Desai, Duncan et al. 2003)	Retrospective claims analysis	Medical and pharmacy claims from a mixed-model US health plan	Osteoporosis (OP); OP-related fracture	Total annual cost of OP care; inpatient costs; average medical costs for women with fracture; medication costs
(Surgenor, Wallace et al. 1989)	Retrospective discharge	Unclear from abstract; full article not	Digestive diseases	Blood transfusions

	analysis	available

Key: *excess resource use study. ACE-I = angiotensin converting enzyme inhibitor; ADLs = activities of daily living; AHRQ = Agency for Healthcare Research and Quality, a branch of the US Department of Health & Human Services; ARB = angiotensin receptor blocker; BB = beta-blocker; DRG = diagnosis-related group; ED = emergency department; EPOs = exclusive provider organizations; HCUP-NIS = Healthcare Cost & Utilization Project Nationwide Inpatient Sample; HMOs = health maintenance organizations; LOS = length of stay; NSAIDs = non-steroidal anti-inflammatory drugs; PPIs = proton pump inhibitors; USD = US dollars.

The final relevant article identified using the search terms “resource utilization” and “ICD,” previously singled out as providing a potential alternative model to simple retrospective database analysis for linking ICD codes to specific health resource use was Rutledge & Osler, “The ICD-9-Based Illness Severity Score: A New Model That Outperforms Both DRG and APR-DRG [All Patient Refined DRG] as Predictors of Survival and Resource Utilization” (1998) (Rutledge and Osler 1998). The study is premised on the need to create prospective payment scales for common diagnoses, reflecting the shift at the time in the US health care system from cost-based reimbursement formulae to prospective payment scales that anticipate costs according to patient characteristics and specific conditions. In this setting, there was concern that inaccurate forecasts for treatment costs were hindering care, particularly for the most ill and complicated patients. The authors use this observation as their rationale for developing a simple, easily applied, inexpensive, highly accurate and widely available predictor for resource utilization and outcomes: namely, an ICD-9 based modeling tool.

Prior to this study, Osler, Rutledge, *et al.* had published work in the trauma literature on a modeling tool called the ICD-9-Based Illness Severity Score (ICISS) (Osler, Rutledge et al. 1996), initially created to predict the risk of death in injured patients based on ICD-9 inpatient codes (both diagnosis and procedure codes). In this study, the same tool is extended to predict resource utilization outcomes (LOS and hospital charges) in addition to the original outcome of patient survival; additionally, this follow-up study added a neural network (i.e. a non-linear modeling technique) to test whether this improved accuracy of predicted resource needs.

According to Osler *et al.*, the ICISS model essentially represents a library of selected ICD diagnosis and procedure codes and their associated SRRs (Survival Risk Ratio in the original study, although this variable can be replaced with others of interest, e.g. hospital charges) derived from an existing data set of ICD diagnosis/procedure codes and outcome data according to the following formula: “the number [of] times a given ICD-9 occurs in a surviving patient is divided by the total number of occurrences of that ICD-9 in the [study] registry. A given ICD-9 SRR thus represents the likelihood that any individual patient will survive that particular ICD-9 injury” (Osler, Rutledge et al. 1996). In the subsequent Rutledge & Osler paper, this approach is summarized as “simple[:] if one wishes to predict the outcome of a patient with any particular diagnosis, the best method is to derive the overall outcome of all of the other patients available with that same diagnosis.” They then apply this principle to a set of resource-use outcomes, namely, survival, length of stay (LOS) and total hospital costs. Specifically, the ICISS outcome values were derived by first creating a “training” data set, used to calculate expected outcome values (survival, mean LOS, mean hospital charges) for each ICD-9 diagnosis and procedure code, using the AHRQ's HCUP-NIS (Agency for Healthcare Research and Quality's Healthcare Cost & Utilization Project Nationwide Inpatient Sample) database, which approximates a 20% sample of US community hospitals and including clinical and resource use information in discharge abstracts; this data set was used to construct a "data dictionary", i.e. database table of expected resource use values associated with specific ICD-9 codes. These expected outcome values were then applied to the ICD-9 diagnosis codes from a test data set, using retrospective hospital data from 9,483 trauma patients at US hospitals, with expected versus actual outcomes evaluated by regression analysis. Of note, all patients with a given ICD-9 diagnosis were used when calculating mean values for LOS and hospital charges (i.e. no one was excluded on the basis of comorbidities or associated procedures). A similar procedure was repeated with the training and test sets for DRGs and APR-DRGs, and the accuracy of each approach was compared. Finally, the ICISS-derived outcomes predictions were provided to a commercial neural network development program (NeuroShell 2, Ward Systems Group) to derive the initial ICISS values; the methods used for this procedure were, however, not described in great detail in the article. Comparative analysis of the different predictive approaches revealed that the ICISS model was significantly better than DRG or APR-DRG for predicting all three outcomes, and the neural network outperformed all other methods (Rutledge and Osler 1998).

Resource utilization studies in developing countries

The results from the literature searches described above suggest that minimal research has been performed on resource utilization in the developing country context, whether in regards to disaster preparedness planning more generally. To determine whether this is in fact the case, and not solely an artifact of the search process, a few additional quality checks were performed by targeted searches involving resource utilization, ICD codes, and preparedness planning in developing countries. The PubMed search 'preparedness AND resources AND ICD' yielded no results. A search using the terms 'preparedness AND resources' (Automatic Term Mapping: preparedness[All Fields] AND ("health resources"[MeSH Terms] OR ("health"[All Fields] AND "resources"[All Fields]) OR "health resources"[All Fields] OR "resources"[All Fields])) was too broad, with 680 results, most of which were clearly unrelated to developing settings. The addition of the term "developing countries" to this search (preparedness AND resources AND "developing countries"; Automatic Term Mapping: preparedness[All Fields] AND ("health resources"[MeSH Terms] OR ("health"[All Fields] AND "resources"[All Fields]) OR "health resources"[All Fields] OR "resources"[All Fields]) AND "developing countries"[All Fields])) resulted in 28 articles, four of which contained information about resource assessment techniques (e.g. onsite surveys assessing functional capacity of the primary healthcare system in Orissa to respond to annual flooding (Phalkey, Dash et al. 2012); descriptive surveys to assess preparedness in health care organizations with direct experience with the Bam earthquake of 2003 vs. those without (Seyedin, Ryan et al. 2011); use of the WHO Tool for Situational Analysis to Assess Emergency and Essential Surgical Care for influenza-related respiratory complications at first-referral health facilities in Africa (Belle, Cohen et al. 2010); and CDPRG's 2010 study using the SYSRA toolkit to evaluate pandemic influenza preparedness in Asia (Hanvoravongchai, Adisasmito et al. 2010)); none of these, however, was ICD-related. The final search, resources AND "developing countries" AND ICD (Automatic Term Mapping: ("health resources"[MeSH Terms] OR ("health"[All Fields] AND "resources"[All Fields]) OR "health resources"[All Fields] OR "resources"[All Fields]) AND "developing countries"[All Fields] AND ICD[All Fields]), identified a total of three articles. Of these, two articles examined causes of mortality by ICD codes with no discussion of linked resource usage (Akerman, Campanario et al. 1996; Ahern, Lozano et al. 2011), while the third used ICD codes to identify cases for an epidemiologic study of childhood poisoning by paraffin ingestion (Tagwireyi, Ball et al. 2006). When this search was expanded through the addition of the MeSH term 'ICD codes' (Automatic Term Mapping: ("health resources"[MeSH Terms] OR ("health"[All Fields] AND "resources"[All Fields]) OR "health resources"[All Fields] OR "resources"[All Fields]) AND "developing countries"[All Fields] AND ("international classification of diseases"[MeSH Terms] OR ("international"[All Fields] AND "classification"[All Fields] AND "diseases"[All Fields]) OR "international classification of diseases"[All Fields] OR ("icd"[All Fields] AND "codes"[All Fields]) OR "icd codes"[All Fields])), an additional seven articles were identified; only one of these, a study of psychiatric hospital utilization by ICD code in Papua New Guinea, was relevant (Johnson 1997) (full text unavailable).

Discussion

There is a clear need for information about effective and economically feasible targets to improve surge capacity planning. In developing countries, in particular, there are significant limitations not only on surge response capability but on baseline health care resources (including resources to improve disaster preparedness). CDPRG's current project aims to build on prior resource gap modeling (Krumkamp, Kretzschmar et al. 2011) to create a simple and easily applied framework that takes into account options for modifying or augmenting existing health-service resources in Indonesia in the setting of a surge event. Such a model will ideally also provide useful data for mobilization of resources in surge situations, by incorporating data about baseline resource usage and the potential for re-allocation between various healthcare settings (both geographically and in terms of intensity of care), allowing for improved surge response capability with less increase in the absolute quantity of health resources. Once complete, the model can best support policy-making for surge capacity building by highlighting areas for health resource investment, with the dual goals of maximizing surge response capability and strengthening existing public health structures and resources.

Within this framework, the challenge of identifying a simple, inexpensive, and readily available predictor variable with a proven record of accuracy in estimating baseline resource utilization is imperative, and ICD diagnosis codes do seem to best fulfill this description. It is true that none of the studies described in Figure 1 provides a particularly useful model for applying ICD diagnosis codes to determine resource utilization in settings like Indonesia, due to the significant complexity of the study questions and, crucially, the fact that the existence of these studies is predicated on the availability of extensive and clinically rich databases that are not approximated anywhere in the developing world. The applicability of these studies to the research question at hand, however, is that they do establish solid precedent for the use of ICD diagnosis codes in determining disease-specific resource utilization, with evidence from Rutledge *et al*. that can do so with a high and replicable degree of accuracy (Rutledge and Osler 1998).

It should be noted here that all US-based studies (i.e. the majority of those examined here) rely on ICD-9 codes, as the 10th revision has not yet been widely adopted there due to the enormous complexities involved in converting the country's complicated insurance claims systems over to the ICD-10 system. The rapid increase in number of resource utilization studies during the first decade of the 21st century noted in the results section no doubt belies the transition to using ICD-9 codes as the basis for insurance claims and reimbursement in the US, with a concomitant and exponential increase in data available for resource utilization research at a highly specific level (e.g. types and doses of particular medications by diagnosis). This, again, highlights the imperative need for accurate and comprehensive data collection, both in terms of diagnoses (primary and secondary) and specific associated resource use. While insurance claims databases reflecting millions of discrete health care events exist only in the industrialized countries, accurate and reliable data on a much more limited level is both conceivable and requisite to create the model we propose. Information about the most common and most resource-intense diagnoses, disaggregated by level of healthcare provision, is essential for the type of preparedness modeling we envision. Such information should ideally be collected in a manner that is clinically rigorous and consistent across institutions, with clear guidelines and algorithms for collection and reporting (e.g. clearly specified collection periods; inclusion of primary and secondary diagnoses; routine documentation of all relevant diagnoses at the time of discharge; accounting of frequent comorbidities/complications; appropriate grouping of clinically-related diagnoses, as indicated). Treatment and discharge information should also be subjected to auditing (ideally on a regular basis by independent local entities; if this is not routine practice, then it will need to be performed in the context of research activities). Rather than suggesting that developing countries imitate high-income countries' preparedness planning by prioritizing high-level syndromic surveillance or other types of sophisticated disaster preparedness databases, this research underscores the need and advocates for investment in efficient, accurate, and integrated (at least vertically) health information systems, in keeping with the WHO's global partnership for improved health information systems, the Health Metrics Network[1] (Kruk 2008).

Assessing the literature specifically for potential models linking ICD diagnosis codes to resource utilization amenable to the Indonesian context, it is noteworthy that most of the articles identified seek to answer health resource utilization questions at a level of sophistication far beyond our purposes. This is not meant to disparage the health information resources in Indonesia or similar countries, but rather to suggest that a different paradigm of baseline resource utilization analysis, with much more limited aims, is appropriate here. The ICISS model described by Rutledge and Osler (Osler, Rutledge et al. 1996; Rutledge and Osler 1998)is intriguing, largely due to its reported accuracy in linking ICD diagnosis and procedure codes to survival, LOS and hospital charges. While these outcomes are rather too broad for our purposes, the demonstration that the technique of creating a "data dictionary" of desired ICD codes and resource outcomes from the development of specific training and test data sets is compelling and suggests that ICD codes are sufficiently robust to be used in this type of direct modeling. There are two important caveats to note, however. First and foremost, although the Rutledge article on LOS and hospital charges relied on a relatively small

---

[1] The Health Metrics Network supports developing countries in assessing health information systems and improving vital registration, and provides training on ICD classification and other international health information standards. The Demographic Surveillance Systems (DSSs) in several African countries and in Bangladesh provide good examples of relatively simple but high-performing health information systems in developing countries. In Tanzania, for example, it costs only $0.53 per capita, per year to collect data on 38 demographic and health indicators via the DSS and census. Kruk, M. E. (2008). "Emergency preparedness and public health systems lessons for developing countries." Am J Prev Med **34**(6): 529-534.

data sample, the database used for the test model included hospital data for 9,483 individual patients, and the training set model drew from HCUP-NIS, a massive AHRQ-led collection of (at the time of publication) 54 million inpatient records from 750-900 hospitals annually. Indonesia is in a unique position to emulate this type of data analysis, albeit on a much more limited scale, for predicting resource demand in the specific surge scenarios of earthquake and pandemic influenza, as case series data conceivably exist from recent experiences with each type of surge event. That said, the selection of training and test sets would have to be performed carefully to ensure appropriate sample representativeness and generalizability. Second, the ICISS was validated using inputs of both ICD-9 diagnosis and procedure codes and, it might be argued, would lose some of the accuracy of resource prediction if only diagnosis codes were used. However, one of the advantages of the ICD-10 over ICD-9 appears to be the integration of severity indices into the diagnosis codes themselves (almost entirely lacking in ICD-9) and would ostensibly off-set some of the detail lost by excluding ICD procedure codes from the ICISS model. Additionally, medical procedures are likely both much more common and more costly in the US setting than in Indonesia, and therefore would figure less frequently in the model, if at all.

There are additional likely limitations implicit to our proposed model. First, a high degree of simplification will inevitably be necessary, both in order to achieve a relatively accessible and generalizable tool for surge capacity assessment and also as a function of anticipated limitations on available data. Given this likelihood, the model can only be expected to yield information about relative quantities of resources in use at baseline and potential surge demand in order to highlight resource gaps, rather than predicting absolute quantities. Second, we may find our results from initial ICD diagnosis analysis of existing resources somewhat skewed by mismatch between standard of care and current practice; that is to say, in very resource limited settings, actual resource utilization at baseline may more accurately reflect existing resource availability rather than resources that should ideally be utilized for routine care of baseline health needs. Some objective measure of reasonable standard of care will thus need to be developed and applied in order to evaluate for this possible bias. Indeed, if such trends do exist in routine care, it is critical that there is an established method for identifying them in order to better direct capacity building for baseline public health care provision before moving on to address potential surge scenario resource gaps. Finally, regarding limitations of this paper in particular, it should be plainly stated that this review of the literature pertaining to ICD codes and resource utilization is not exhaustive; instead, as suggested in the presentation of the results, additional literature searches were found only to yield articles similar in type and methodology to those already examined and were not highly applicable to the context of developing countries; therefore additional review was thought likely redundant.

Conclusions

Analyzing and building surge capacity and surge response capability is a crucial component of disaster preparedness planning, particularly in developing countries. As our current project in Indonesia illustrates, an essential foundational step in addressing potential resource gaps is establishing baseline health resource utilization needs across a variety of health care settings. Such a task requires a simple, inexpensive, easily applied and accurate tool to predict resource needs. ICD codes have been proven accurate for this purpose and have been used to link health resource utilization with specific diagnoses in a number of studies, albeit primarily in the United States and other high-income countries. Given the anticipated constraints in terms of health information data in Indonesia and similar countries, a simpler model of ICD code-based estimation of direct costs, using cost estimate profiles derived from a number of data sources, is most feasible.

Ahern, R. M., R. Lozano, et al. (2011). "Improving the public health utility of global cardiovascular mortality data: the rise of ischemic heart disease." Popul Health Metr **9**: 8.

Akerman, M., P. Campanario, et al. (1996). "[Environment and health: an analysis of intra-urban differentials in the city of São Paulo, Brazil]." Rev Saude Publica **30**(4): 372-382.

Andersson, P. A., E. Varde, et al. (2000). "Modelling of resource allocation to health care authorities in Stockholm county." Health Care Manag Sci **3**(2): 141-149.

Asthana, S., A. Gibson, et al. (2004). "The pursuit of equity in NHS resource allocation: should morbidity replace utilisation as the basis for setting health care capitations?" Soc Sci Med **58**(3): 539-551.

Barbisch, D. F. and K. L. Koenig (2006). "Understanding surge capacity: essential elements." Acad Emerg Med **13**(11): 1098-1102.

Barron, J. J., R. Quimbo, et al. (2008). "Assessing the economic burden of breast cancer in a US managed care population." Breast Cancer Res Treat **109**(2): 367-377.

Belle, J., H. Cohen, et al. (2010). "Influenza preparedness in low-resource settings: a look at oxygen delivery in 12 African countries." J Infect Dev Ctries **4**(7): 419-424.

Bogner, H. R., S. D. Miller, et al. (2010). "Assessment of cost and health resource utilization for elderly patients with heart failure and diabetes mellitus." J Card Fail **16**(6): 454-460.

CDC. (2005). "Public Health Guidance for Community-Level Preparedness and Response to Severe Acute Respiratory Syndrome (SARS)." Retrieved 18 January, 2012, from http://www.cdc.gov/sars/guidance/core/app2.html. Accessed 18 January 2011.

Collard, H., A. Ward, et al. (2012). "Burden of illness in idiopathic pulmonary fibrosis." J Med Econ **15**(5): 829-835.

Crockett, S., S. Sperry, et al. (2012). "Emergency care of esophageal foreign body impactions: timing, treatment modalities, and resource utilization." Dis Esophagus **[Epub ahead of print]**.

Desai, S. S., B. S. Duncan, et al. (2003). "The cost of treating osteoporosis in a managed health care organization." J Manag Care Pharm **9**(2): 142-149.

Dixon, J., P. Smith, et al. (2011). "A person based formula for allocating commissioning funds to general practices in England: development of a statistical model." BMJ **343**: d6608.

Dugar, D. R., L. Lander, et al. (2010). "Pediatric acute sinusitis: predictors of increased resource utilization." Laryngoscope **120**(11): 2313-2321.

Dunn, J. D. and M. W. Pill (2009). "A claims-based view of health care charges and utilization for commercially insured patients with osteoarthritis." Manag Care **18**(12): 44-50.

Goldacre, M. J. and R. I. Harris (1980). "Mortality, morbidity, resource allocation, and planning: a consideration of disease classification." Br Med J **281**(6254): 1515-1519.

Gore, M., K. S. Tai, et al. (2011). "Clinical comorbidities, treatment patterns, and direct medical costs of patients with osteoarthritis in usual care: a retrospective claims database analysis." J Med Econ **14**(4): 497-507.

Handler, J. A., M. Gillam, et al. (2006). "Metrics in the science of surge." Acad Emerg Med **13**(11): 1173-1178.

Hanvoravongchai, P., W. Adisasmito, et al. (2010). "Pandemic influenza preparedness and health systems challenges in Asia: results from rapid analyses in 6 Asian countries." BMC Public Health **10**: 322.

Hung, T., Y. Lai, et al. (2012). "Trend Analysis of Hospital Resource Utilization for Prolonged Mechanical Ventilation Patients in Taiwan: A Population Based Study." Respir Care **[Epub ahead of print]**.

Iyer, S., W. B. Saunders, et al. (2009). "Economic burden of postoperative ileus associated with colectomy in the United States." J Manag Care Pharm **15**(6): 485-494.
James, A. H., S. T. Patel, et al. (2008). "An assessment of medical resource utilization and hospitalization cost associated with a diagnosis of anemia in women with obstetrical bleeding in the United States." J Womens Health (Larchmt) **17**(8): 1279-1284.
Johnson, F. (1997). "Ward Six Psychiatric Unit at the Port Moresby General Hospital: a historical review and admission statistics from 1980 to 1989 " PNG Med J **40**(2): 79-88.
Kelen, G. D. and M. L. McCarthy (2006). "The science of surge." Acad Emerg Med **13**(11): 1089-1094.
Kollef, M., C. Hamilton, et al. (2012). "Economic impact of ventilator-associated pneumonia in a large matched cohort." Infect Control Hosp Epidemiol **33**(3): 250-256.
Kruk, M. E. (2008). "Emergency preparedness and public health systems lessons for developing countries." Am J Prev Med **34**(6): 529-534.
Krumkamp, R., M. Kretzschmar, et al. (2011). "Health service resource needs for pandemic influenza in developing countries: a linked transmission dynamics, interventions and resource demand model." Epidemiol Infect **139**(1): 59-67.
Levin, A., M. R. Chaudhry, et al. (2009). "Diabetes, kidney disease and cardiovascular disease patients. Assessing care of complex patients using outpatient testing and visits: additional metrics by which to evaluate health care system functioning." Nephrol Dial Transplant **24**(9): 2714-2720.
Lopez-Bastida, J., J. Moreno, et al. (2012). "Social and economic costs and health-related quality of life in stroke survivors in the Canary Islands, Spain." BMC Health Serv Res **12**: 315-323.
Macaulay, D., J. Ivanova, et al. (2012). "Direct and indirect costs associated with Dupuytren's contracture." J Med Econ **15**(4): 664-671.
Margolis, J., J. J. Barron, et al. (2005). "Health care resources and costs for treating peripheral artery disease in a managed care population: results from analysis of administrative claims data." J Manag Care Pharm **11**(9): 727-734.
McCollam, P. and L. Etemad (2005). "Cost of care for new-onset acute coronary syndrome patients who undergo coronary revascularization." J Invasive Cardiol **17**(6): 307-311.
Mirkin, D., C. Murphy-Barron, et al. (2007). "Actuarial analysis of private payer administrative claims data for women with endometriosis." J Manag Care Pharm **13**(3): 262-272.
Morrison, J., S. T. Patel, et al. (2008). "Assessment of the prevalence and impact of anemia on women hospitalized for gynecologic conditions associated with heavy uterine bleeding." J Reprod Med **53**(5): 323-330.
Mugisha, F., B. Kouyate, et al. (2002). "Costing health care interventions at primary health facilities in Nouna, Burkina Faso." Afr J Health Sci **9**(1-2): 69-79.
Oshitani, H., T. Kamigaki, et al. (2008). "Major issues and challenges of influenza pandemic preparedness in developing countries." Emerg Infect Dis **14**(6): 875-880.
Osler, T., R. Rutledge, et al. (1996). "ICISS: an international classification of disease-9 based injury severity score." J Trauma **41**(3): 380-386; discussion 386-388.
Owusu, J., M. Liu, et al. (2012). "Resource utilization in primary repair of cleft palate." Laryngoscope **[Epub ahead of print]**.
Patel, D. A., X. Gao, et al. (2011). "US hospital database analysis of invasive aspergillosis in the chronic obstructive pulmonary disease non-traditional host." J Med Econ **14**(2): 227-237.
Phalkey, R., S. Dash, et al. (2012). "Prepared to react? Assessing the functional capacity of the primary health care system in rural Orissa, India to respond to the devastating flood of September 2008." Glob Health Action **5**.
Poordad, F., D. Theodore, et al. (2011). "Evaluating medical resource utilization and costs associated with thrombocytopenia in chronic liver disease patients." J Med Econ.
Rothman, R. E., E. B. Hsu, et al. (2006). "Research priorities for surge capacity." Acad Emerg Med **13**(11): 1160-1168.

Rudge, J. W. and e. al. (Unpublished). Health system resource gaps and associated mortality from pandemic influenza across six Southeast Asian territories.
Rutledge, R. and T. Osler (1998). "The ICD-9-based illness severity score: a new model that outperforms both DRG and APR-DRG as predictors of survival and resource utilization." J Trauma **45**(4): 791-799.
Saleh, M. N., M. Fisher, et al. (2009). "Analysis of the impact and burden of illness of adult chronic ITP in the US." Curr Med Res Opin **25**(12): 2961-2969.
Sanchez, R. J., C. Uribe, et al. (2011). "Longitudinal evaluation of health care utilization and costs during the first three years after a new diagnosis of fibromyalgia." Curr Med Res Opin **27**(3): 663-671.
Seyedin, H., J. Ryan, et al. (2011). "Lessons learnt from the past and preparedness for the future: how a developing country copes with major incidents." Emerg Med J **28**(10): 887-891.
Sicras-Mainar, A., J. Rejas, et al. (2009). "Treating patients with fibromyalgia in primary care settings under routine medical practice: a claim database cost and burden of illness study." Arthritis Res Ther **11**(2): R54.
Simon, H. K., D. A. Hirsh, et al. (2009). "Pediatric emergency department overcrowding: electronic medical record for identification of frequent, lower acuity visitors. Can we effectively identify patients for enhanced resource utilization?" J Emerg Med **36**(3): 311-316.
Singh, J. A. and V. Strand (2009). "Health care utilization in patients with spondyloarthropathies." Rheumatology (Oxford) **48**(3): 272-276.
Surgenor, D. M., E. L. Wallace, et al. (1989). "Utility of DRG and ICD-9-CM classification codes for the study of transfusion issues. Transfusions in patients with digestive diseases." Transfusion **29**(9): 761-767.
Tagwireyi, D., D. E. Ball, et al. (2006). "Toxicoepidemiology in Zimbabwe: admissions resulting from exposure to paraffin (kerosene)." Clin Toxicol (Phila) **44**(2): 103-107.
Tencer, T., H. S. Friedman, et al. (2007). "Medical costs and resource utilization for hemophilia patients with and without HIV or HCV infection." J Manag Care Pharm **13**(9): 790-798.
Udall, M., J. Harnett, et al. (2012). "Costs of pregabalin or gabapentin for painful diabetic peripheral neuropathy." J Med Econ **15**(2): 361-370.
UNEP/GRID-Arendal. (2005). "Trends in natural disasters." UNEP/GRID-Arendal Maps and Graphics Library Retrieved 18 January, 2012, from http://maps.grida.no/go/graphic/trends-in-natural-disasters.
Vashishta, R., A. Mahalingam-Dhingra, et al. (2012). "Thyroidectomy Outcomes: A National Perspective." Otolaryngol Head Neck Surg **[Epub ahead of print]**.
West, R. (1978). "Bed usage and disease specific mortality within ICD chapters." J Epidemiol Community Health **32**(1): 38-40.
WHO. (2012). "International Classification of Diseases (ICD)." Programmes and Projects: Classifications Retrieved 19 January, 2012, from http://www.who.int/classifications/icd/en/.
Wu, E. Q., A. Forsythe, et al. (2011). "Comorbidity Burden, Healthcare Resource Utilization, and Costs in Chronic Gout Patients Refractory to Conventional Urate-Lowering Therapy." Am J Ther.