<div style="background:blue">

Commonwealth of Massachusetts
Center for Health Information & Analysis (CHIA)
Non-Government Agency Application for Data

</div>

NOTE: *This application is to be used by all applicants, except Government Agencies, as defined in 957 CMR 5.02.*

I. GENERAL INFORMATION

APPLICANT INFORMATION	
Applicant Name:	Amresh Hanchate
Title:	Assistant Professor
Organization:	Boston University School of Medicine
Project Title:	801 Massachusetts Avenue, 2nd floor, Boston, MA 02118
Date of Application:	8/16/2013
Project Objectives (240 character limit)	National Estimates for Inpatient Care, Outcomes and Hospital Effect Among Hispanics.
Project Research Questions	1. Estimate population-level rates among Hispanics, non-Hispanic Whites and non-Hispanic Blacks for AHRQ Prevention Quality Indicators (PQIs), AHRQ Referral sensitive inpatient surgeries and AHRQ Inpatient Quality Indicator (IQIs) surgical and admission conditions 2. Estimate risk-adjusted rates of inpatient mortality, 30-day mortality and 30-day readmissions following AHRQ IQI surgical procedure and medical admissions by ethnicity and race. 3. Estimate the extent to which differences in risk-adjusted outcomes for Hispanics are associated with the hospitals where they are treated.

Please indicate if you are a Researcher, Payer, Provider or Provider Organization and you are seeking data pursuant to 957 CMR 5.04 (De-Identified Data) or 957 CMR 5.05 (Direct Patient Identifiers for Treatment or Coordination of Care).

☒ Researcher	☐ 957 CMR 5.04 (De-identified Data)
☐ Payer	☐ 957 CMR 5.05 (Direct Patient Identifiers)
☐ Provider / Provider Organization	

All other requests are subject to 957 CMR 5.06.

II. PROJECT SUMMARY
Briefly describe the purpose of your project and how you will use the CHIA data?

Hispanics now form the largest ethnic or racial minority in the US. Evidence based on national surveys or death records indicates significant differences in health behaviors and risk compared to non-Hispanic Whites or non-Hispanic Blacks. Hispanics are also at greater risk for healthcare access barriers due to higher rates of uninsurance, poverty and inadequate health literacy. However, little is known about inpatient care access, quality or discharge outcomes among Hispanics nationally. This is largely due to poor identification of Hispanic subjects in national inpatient care data sources. This gap in knowledge limits the ability "to set and monitor national objectives, to formulate policy, and to design legislation to address the disparities".

We propose to use a novel approach of combining state inpatient discharge (SID) data along with census population data to obtain representative estimates of inpatient care utilization for Hispanic adults and contrast them with those for non-Hispanic Whites and non-Hispanic Blacks. We will examine SID data (2010-11) from 15 states that together account for over 87 percent of the national Hispanic adult population.

We will examine a range of **inpatient care** indicators of distinct domains of care. As indicators of **access to care**, we will use the ACSC conditions identified by Agency for Healthcare Research Quality (AHRQ) Prevention Quality Indicators (PQIs) (e.g., diabetes and hypertension) and AHRQ *referral sensitive surgeries* (e.g., percutaneous transluminal coronary angioplasty [PTCA] and knee replacement). To evaluate **quality** and patient **outcomes** we will use AHRQ Inpatient Quality Indicators (IQI), comprising of eight surgical (e.g., esophageal resection) and six medical admission conditions (e.g., acute myocardial infarction [AMI]).

Our aims are to estimate the following indicators for Hispanics, non-Hispanic Blacks and non-Hispanic Whites aged ≥ 21: (a) population-level rates of admission for PQIs, referral sensitive surgeries and IQIs and (b) risk-adjusted rates of inpatient mortality, 30-day mortality and 30-day readmissions following admission for IQI medical or surgical admission. We will also estimate the extent to which differences in risk-adjusted rates Hispanics are associated with the hospitals where they are treated.

The proposal is innovative and significant: representative estimates of inpatient care utilization and their outcomes for Hispanics have never been presented; they will establish measurable target indicators, help set national priorities, and will provide a mechanism with which to judge the relative impact of new policy initiatives for minorities by ethnicity and race.

III. FILES REQUESTED

Please indicate the databases from which you seek data, the Level(s) and Year(s) of data sought.

DATABASE	Level 1[1] or 2[2]	Single or Multiple Use	Year(s) Of Data Requested Current Yrs. Available 2009 - 2011
☐ Medical Claims	☐ Level 1 ☐ Level 2	Select... ▼	☐ 2009 ☐ 2010 ☐ 2011
☐ Pharmacy Claims	☐ Level 1 ☐ Level 2	Select... ▼	☐ 2009 ☐ 2010 ☐ 2011
☐ Dental Claims	☐ Level 2	Select... ▼	☐ 2009 ☐ 2010 ☐ 2011
☐ Member Eligibility	☐ Level 2	Select... ▼	☐ 2009 ☐ 2010 ☐ 2011
☐ Provider	☐ Level 2	Select... ▼	☐ 2009 ☐ 2010 ☐ 2011
☐ Product	☐ Level 2	Select... ▼	☐ 2009 ☐ 2010 ☐ 2011

CASEMIX	Level 1 - 6	Fiscal Years Requested
Inpatient Discharge	☐ Level 1 – No Identifiable Data Elements ☐ Level 2 – Unique Physician Number (UPN) ☐ Level 3 – Unique Health Information Number (UHIN) ☐ Level 4 – UHIN and UPN ☑ Level 5 – Date(s) of Admission; Discharge; Significant Procedures ☐ Level 6 – Date of Birth; Medical Record Number; Billing Number	1998-2012 Available (limited data available 1989-1997) 2010 (already aquired) 2011
Outpatient Observation	☐ Level 1 – No Identifiable Data Elements ☐ Level 2 – Unique Physician Number (UPN) ☐ Level 3 – Unique Health Information Number (UHIN) ☐ Level 4 – UHIN and UPN ☑ Level 5 – Date(s) of Admission; Discharge; Significant Procedures ☐ Level 6 – Date of Birth; Medical Record Number; Billing Number	2002-2011 Available 2010 (already aquired) 2011

---

[1] Level 1 Data: De-identified data containing information that does not identify an individual patient and with respect to which there is no reasonable basis to believe the data can be used to identify an individual patient. This data is de-identified using standards and methods required by HIPAA.

[2] Level 2 (and above) Data: Includes those data elements that pose a risk of re-identification of an individual patient.

Emergency Department	☐ Level 1 – No Identifiable Data Elements ☐ Level 2 – Unique Physician Number (UPN) ☐ Level 3 – Unique Health Information Number (UHIN) ☐ Level 4 – UHIN and UPN; Stated Reason for Visit ☑ Level 5 – Date(s) of Admission; Discharge; Significant Procedures ☐ Level 6 – Date of Birth; Medical Record Number; Billing Number	<u>2000-2011 Available</u> 2010 (already acquired) 2011

IV. REQUESTED DATA ELEMENTS [APCD]

State and federal privacy laws limit the use of individually identifiable data to the minimum amount of data needed to accomplish a specific project objective. Please use the APCD Data Specification Workbook to identify which data elements you would like to request and attach this document to your application.

V. REQUESTED DATA ELEMENTS [CASE MIX]

Please use the CASE MIX DATA SPECIFICATION WORKBOOK to identify which deniable data elements (from Level 2 or above) you would like to request and attach this to your application.

Workbook not yet available, as per Adam Tapply email 8/16/13

VI. MEDICAID DATA

Federal law (42 USC 1396a(a)7) restricts the use of individually identifiable data of Medicaid recipients to uses that benefit the administration of the Medicaid program. If you are requesting Medicaid data from Level 2 or above, please describe in detail why your use of the data benefits the administration of the Medicaid program.

VII. MEDICARE DATA

Medicare data may be disseminated to state agencies and/or entities conducting research projects that are directed and partially funded by the state if such research projects would allow for a Privacy Board or an IRB to make the findings listed at 45 CFR 164.512(i)(2)(ii) if the anticipated data recipient were to apply for the data from CMS directly. If you are requesting Medicare data, please explain how your research project is directed and partially funded by the state and describe in detail why your proposed project meets the criteria set forth in 45 CFR 164.512(i)(2)(ii). Applicants must describe how they will use the data and inform CHIA where the data will be housed. CHIA must be informed if the data has been physically moved, transmitted, or disclosed.

VIII. DIRECT PATIENT IDENTIFIERS[3]

State and federal privacy laws may require the consent of Data Subjects prior to the release of any Direct Patient Identifiers. If you are requesting data that includes Direct Patient Identifiers, please provide documentation of patient consent or your basis for asserting that patient consent is not required.

---

[3] <u>Direct Patient Identifiers</u>. Personal information, such as name, social security number, and date of birth, that uniquely identifies an individual or that can be combined with other readily available information to uniquely identify an individual.

IX. REQUESTS PURSUANT TO 957 CMR 5.04

Payers, providers, provider organizations and researchers seeking access to Level 1 (de-identified) data are required to describe how they will use such data for the purposes of lowering total medical expenses, coordinating care, benchmarking, quality analysis or other administrative research purposes. Please provide this information below.

N/A

X. FILTERS

If you are requesting APCD elements from Level 2 or above, describe any filters you are requesting to use in order to limit your request to the minimum set of records necessary to complete your project. (For example, you may only need individuals whose age is less than 21, claims for hospital services only, or only claims from small group projects.

APCD FILE	DATA ELEMENT(S) FOR WHICH FILTERS ARE REQUESTED	RANGE OF VALUES REQUESTED
Medical Claims		
Pharmacy Claims		
Dental Claims		
Membership Eligibility		
Provider		
Product		

XI. PURPOSE AND INTENDED USE

1. Please explain why completing your project is in the public interest.

In its landmark report, *Unequal Treatment*, the Institute of Medicine (IOM) reported that "a central concern of the committee … has been the relative paucity of data on non-African-American racial and ethnic minority groups." We propose a unique study, titled "**National Estimates for Inpatient Care, Outcomes & Hospital Role among Hispanics**", that assembles a near-national data set for Hispanic adults, estimates rates of access to care (e.g., ACSC admissions), assesses relative quality of discharge outcomes (e.g., 30-day mortality and readmissions) and evaluates the role of hospitals where most Hispanics receive care. All estimates for Hispanics will be contrasted with those for non-Hispanic Whites and non-Hispanic Blacks. We will examine medical admissions and surgical procedures that cover a range of acute clinical areas (including endocrine, cardiovascular, cerebrovascular, cancer, digestive, musculoskeletal and respiratory care) and capture three critical domains of health disparities in inpatient care for Hispanics: i) **access**, ii) **quality** and iii) **outcomes**.

2. **Attach** a brief (1-2 pages) description of your research methodology. (This description will not be posted on the internet.)

    See attachment entitled Methodology

3. Has your project received approval from your organization's Institutional Review Board (IRB)?

☑ Yes, and a copy of the approval letter is attached to this application. See attachment entitled IRBApprove2013_Hanchate.

☐ No, the IRB will review the project on _____.

☐ No, this project is not subject to IRB review.

☐ No, my organization does not have an IRB.

XII. APPLICANT QUALIFICATIONS

1. Describe your qualifications to perform the research described or accomplish the intended use of CHIA data.

> Amresh Hanchate, Ph.D., Principal Investigator, is an Assistant Professor of Medicine (BUSM) and Health Economist (VA-Boston), and has extensive experience in the use of large databases for epidemiologic disparities research. He is PI & co-PI of two ongoing NIH-funded studies that utilize data from the same sources as those in the proposed study. One study (R21 NINDS) uses state inpatient discharge data from eight states and census population data to estimate population-level incidence rates of ischemic stroke, and evaluate discharge outcomes, by race and ethnicity. Another study (co-PI of NHLBI center grant project) uses state inpatient discharge and census data from five states, to evaluate the impact of Massachusetts health reform on racial/ethnic disparities in access to inpatient healthcare utilization and their outcomes. He is also PI and co-investigator of other ongoing studies examining (a) gains in model prediction from adding clinical fields (laboratory tests and vital signs) to administrative-data based discharge outcome models (PI), and (b) statistical properties of commonly used models of hospital ratings based on administrative data (PI) and (c) impact of health care insurance on time to treatment resolution following breast and cervical cancer abnormal tests (co-investigator). Dr. Hanchate's research has been published in high-impact peer-reviewed journals, including *Archives of Internal Medicine, Health Services Research,* and *Medical Care.* He has extensive experience in the design and estimation of hierarchical statistical models involving merged patient-level data with broader level data at the zip code, county and state levels.

2. Attach résumés or curriculum vitae of the applicant/principal investigator, key contributors, and of all individuals who will have access to the data. (These attachments will not be posted on the internet.)

> **See attached forms for :**
> Biosketch_Paasche-Orlow
> Nancy Kressin_biosketchHispanicDB
> Biosketch_Hanchate_NIH_201205 NIMHD R01
> Rosen_biosketch
> Chen Feng-CV

XIII. DATA LINKAGE AND FURTHER DATA ABSTRACTION

1. Does your project require linking the CHIA Data to another dataset?   YES ☑   NO ☐

6

2. If yes, will the CHIA Data be linked to other patient level data or with aggregate data (e.g. Census data)?

Patient Level Data ⌐   Aggregate Data   ☑

3. If yes, please identify all linkages proposed and explain the reasons(s) that the linkage is necessary to accomplish the purpose of the project.

	Source	Source Organization	Time Period	Data Description
				Table 5. Data Sources
1.	State Inpatient Discharge with Vital Status	Individual state healthcare agency or AHRQ	2010-11	Universe of all discharges from non-Federal short-term acute care hospitals; patient ethnicity, race, demographics, admission diagnosis and procedure codes, hospital identifier, zip code and vital status (death date)
2.	Population Data	Census Bureau	2010-11	County population by sex, age, race/ethnicity
3	Medicare Fee for Service	CMS	2010-11	Inpatient and outpatient claims records
4.	American Hospital Association (AHA) Annual Survey Database	AHA	2010-11	Hospital data on facilities, services offered, revenue from main payers and organizational structure
5.	Area Resource File (ARF)	HRSA	2010-11	County-level data on health resources (hospital beds, physicians by specialty)
6.	State Health Facts	Kaiser Foundation	2010-11	State-level public expenditure and assistance data
7.	Small Area Income & Poverty Estimates (SAIPE)	Census Bureau	~2010-11	County-level population estimates of selected income and poverty statistics
8.	Small Area Health Insurance Estimates (SAHIE)	Census Bureau	~2010-11	County-level population estimates of uninsured
9.	Local Area Unemployment Statistics (LAUS)	Bureau of Labor	~2010-11	County-level unemployment estimates

While SID data will be the primary data source, we will obtain other complementary data from a variety of sources (Table 5). An important data source will be the Census Bureau (CB); in addition to estimation of census population stratified by state, ethnicity, race, sex and age, we will use zip-code median income to stratify cohort patients based on socioeconomic status. Annual Census estimates adjust for immigration and emigration, and are available at state and county levels. As noted earlier we will obtain Medicare Fee for Service data covering inpatient stays and outpatient visits for the selected AHRQ surgery and medical admission cohorts. While 100 percent of the inpatient records will be included, we will limit outpatient data to a significant random sample (10 to 20 percent), stratified to ensure sample adequacy by ethnicity and race, to reduce acquisition and processing costs. In addition we will obtain hospital and county-level data from several sources to provide more information for better characterizing the site of care and the socioeconomic environment of study cohorts. The American Hospital Association (AHA) Annual Survey Database will provide information about hospital characteristics relevant to patient outcomes. We will use Area Resource Files (ARF) to obtain county-level data on health care resources (e.g., # hospitals beds per capita, # Medicaid enrollees) that could potentially impact procedure volume. Kaiser Family Foundation's State Health Facts will provide data on differences across states in public assistance, e.g., Medicaid eligibility rules, enrollments and expenditures, disproportionate share and other public health assistance spending over the study period. Annual changes in the uninsured population, employment, income and poverty are also measured using Census Bureau sources. Two programs, Small Area Health Insurance Estimates (SAHIE) and Small Area Income & Poverty Estimates (SAIPE), have developed county- and state-level

estimates of the percentage of persons who are uninsured, or who are poor (based on federal poverty level) and median incomes. While the ethnic and racial distribution of the patient cohorts is of primary interest in this study, the other measures will also be crucial in adjusting for a) baseline differences across counties and states in their economic conditions, and b) changes in these conditions following the recent national economic downturn. We will obtain corresponding unemployment estimates from the Bureau of Labor Statistic's Local Area Unemployment Statistics (LAUS) program.

4. If yes, please identify the specific steps you will take to prevent the identification of individual patients in the linked dataset.

We will not attempt to identify individual patient records. All data kept on hospital hard drives will be de-identified with unique patient identifiers. Any key that exists will be kept in a locked file drawer in a locked office until the study has been completed. At that time all data and key will be destroyed; CDs will be shredded, paper will be shredded with cross hatch shredding and on-line data will be written over.

XIV. PUBLICATION / DISSEMINATION / RE-RELEASE

1. Describe your plans to publish or otherwise disclose CHIA Data, or any data derived or extracted from such data, in any paper, report, website, statistical tabulation, or similar document.

Given the high impact potential of our research, we have considered a broad dissemination plan for our findings. This plan includes:
- Presentation of research findings at national research meetings such as the American Heart Association, Society for General Internal Medicine, Academy Health and AHRQ.
- Submission for publication in high-impact peer-reviewed medical and health policy journals
- Coordination with media offices at the authors' home institutions to write and disseminate press releases about our findings.
- Posting findings on the website of the Boston University School of Medicine and Harvard Medical School (home of collaborators for the proposed project).
- Presentation of health policy briefs to MA and US elected and appointed officials with key roles in health policy planning.

Our collaboration with health economists and health policy experts will support this dissemination plan.

2. Will the results of your analysis be publicly available to any interested party? Please describe how an interested party will obtain your analysis and, if applicable, the amount of the fee.

Results will be published in peer reviewed publications that will be available on-line free of cost.

3. Will you use the data for consulting purposes?　　　YES ☐　　　NO ☑

4. Will you be selling standard report products using the data?　　YES ☐　　NO ☑

5. Will you be selling a software product using the data? YES ☐ NO ☑

6. If you have answered "yes" to questions 3, 4 or 5, please describe the types of products, services or studies.

N/A

XV. USE OF AGENTS AND/OR CONTRACTORS

<u>Third-Party Vendors</u>. Provide the following information for all agents and contractors who will work with the CHIA Data.

Company Name:	N/A
Contact Person:	
Title:	
Address:	
Telephone Number:	
E-mail Address:	
Organization Website:	

1. Will the agent/contractor have access to the data at a location other than your location or in an off-site server and/or database? YES ☐ NO ☑
   N/A

2. Describe the tasks and products assigned to this agent or contractor for this project.

N/A

3. Describe the qualifications of this agent or contractor to perform such tasks or deliver such products.

N/A

4. Describe your oversight and monitoring of the activity and actions of this agent or subcontractor.

N/A