Commonwealth of Massachusetts
Center for Health Information & Analysis (CHIA)
Non-Governmental Application for Case Mix Data
[Exhibit A: Data Application]

This form is required by all Applicants, except Government Agencies as defined in 957 CMR 5.02. All Applicants must also complete the Data Management Plan, attached to this Application. The Application and the Data Management Plan must be signed by an authorized signatory of the organization. This Application and the Data Management Plan will be used by CHIA to determine if your organization may receive CHIA data. Please be sure the documents are completed fully and accurately. You may wish to consult the Evaluation Guide that CHIA will use to review your documents. Prior to receiving CHIA Data, the organization must execute the Data Use Agreement. You may wish to review that document as you complete these forms. This application should be completed by the Primary Investigator, and must be signed by a party with authority to bind the organization seeking CHIA Data for the purposes described herein.

NOTE: In order for your application to be processed, you must submit the required application fee. Please consult the fee schedule for the appropriate fee amount. A remittance form with instructions for submitting the application fee is available on the CHIA website.

All attachments must be uploaded to IRBNet with your Application. All applications documents can be found on the CHIA website in Word and/or PDF format.

I. GENERAL INFORMATION

APPLICANT INFORMATION	
Applicant Name: (Primary Investigator)	Sharon-Lise Normand, PhD
Title:	Director, Massachusetts Data Analysis Center (Mass-DAC) Professor of Health Care Policy (Biostatistics), Department of Health Care Policy
Organization Requesting Data: (Recipient)	Massachusetts Data Analysis Center (Mass-DAC) at Harvard Medical School Dept. of Health Care Policy
Project Title:	Comparing the cost-effectiveness of PCI, standard medical therapy, and CABG using a population of patients with Chronic Kidney Disease.
IRBNet ID:	610321
Address, City/Town, Zip Code	180A Longwood Ave, Boston, MA 02115
Telephone Number:	617-432-3260
Email Address:	sharon@hcp.med.harvard.edu
Names of Co-Investigators:	David Charytan, MD
Email Addresses of Co-Investigators:	dcharytan@partners.org
Original Data Request Submission Date:	
Dates Data Request Revised:	FY2013 and FY2014 case mix data
Project Objectives (240 character limit):	Patients with chronic kidney disease (CKD) have a high incidence of coronary artery disease and a high risk of cardiovascular (CV) death yet are less likely than those with preserved renal function to undergo PCI or receive standard medical therapy (MTX). This project

	will analyze the comparative clinical effectiveness, costs, and cost-effectiveness of PCI, MTX, and CABG using a population of patients with CKD.
Project Research Questions (if applicable) Business Use Case(s):	1. To determine the association between initial choice of therapy and subsequent all-cause mortality, and ESRD in patients with CKD. 2. To determine the association between initial choice of therapy and comparative life-time healthcare-associated costs. 3. To determine the associations between initial choice of therapy and differences in 1) quality-adjusted survival and 2) incremental cost-effectiveness using decision and cost-analytic models informed by estimates from Aims 1 and 2.

II. PUBLIC INTERST & PROJECT SUMMARY

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

CKD affects >10% of the population and is associated with a high risk of developing and dying from cardiovascular disease (CVD). The risk of death or a recurrent CV event following myocardial infarction (MI) increases by 10% for each 10/mL/min/1.73m2 decrement in glomerular filtration rate (GFR)5, and even minor reductions in GFR are strongly associated with the risk of post-MI death and of developing de novo CAD. Although multiple contributing factors have been implicated in this high risk of CV death, the exceedingly high prevalence (>50%) of obstructive coronary atherosclerosis (which is often diffuse) in patients with advanced CKD is undoubtedly a major and treatable contributing factor.

It is increasingly apparent that the optimal management of CKD requires a focus on CVD as well as on preventing the progression of CKD. Elderly CKD patients, for example, are 5-10 times more likely to die from CVD than to reach ESRD. Conversely, among those surviving to initiate dialysis, the annual incidence non-fatal MI and CVD death exceeds 10%. Aggressive application of standard CV therapies in this high-risk population might decrease CV risk, but the use of standard therapies is paradoxically lower in CKD than in patients with preserved renal function. Following MI, for example, those with pre-dialysis CKD or ESRD are 30-40% less likely to undergo coronary angiography or revascularization than patients with normal renal function. Angiography is frequently withheld in moderate-severe CKD even when standard indications for coronary angiography or revascularization are present4, and low utilization of MTX such as β-blockers, renin-angiotensin axis inhibitors, statins and aspirin has also been repeatedly observed.

This selective underutilization of potentially life-saving CV therapies in a population at high risk of CV death has been referred to as “renalism”. Although, tempting to advocate increased use of standard therapies as a means of decreasing CV morbidity, whether this is indicated is uncertain. Several factors including the higher overall and all-cause mortality rates in patient with CKD, the altered and reversed associations of typical risk factors with CV outcomes in CKD, and the routine exclusion of patients with CKD from the majority of clinical trials testing CVD therapies engender significant reservations about the relevance of current guidelines for use of MTX in the treatment of patients with CKD. In fact, several randomized clinical trials have confirmed the unique nature of CAD in patients with CKD by demonstrating an altered response to standard therapies in patients with advanced CKD. Statins, for example, failed to improve mortality in the SHARP, 4D and AURORA trials, large, placebo-controlled randomized clinical trials in moderate-advanced CKD or ESRD, despite the nearly universally benefits in other populations.
The efficacy of coronary revascularization—PCI and CABG—is even less certain in CKD. Trials conducted in the general population suggest that CABG prevents death and MI more effectively than either medical therapy or PCI among

patients with high-risk coronary anatomy or multiple clinical risk factors. CABG is nevertheless, often used as a second-line therapy. In contrast, PCI, despite less certain mortality benefits, is frequently the initial choice in patients when left main, 3-vessel, or proximal left anterior descending disease are absent. Although less invasive than CABG, there are potential drawbacks to PCI. Repeat revascularization is required more frequently, and mortality appears to be higher following PCI than CABG in the setting of multi-vessel disease, diffuse CAD or left main disease. In clinical practice, CABG frequently is reserved for cases with high-risk coronary anatomy, while patient and physician preferences guide the choice between CABG and PCI in other settings. As with MTX, differences in risk factors and pathophysiology of CAD as well as the routine exclusion of patients with CKD from relevant trials raise important concerns about the extrapolation of guidelines for revascularization to individuals with CKD. More serious concerns are raised by the high overall and CV mortality rates in CKD. Roughly 40% of individuals with moderate or severe CKD die with 3-years after MI, and annual all-cause mortality in Medicare patients with CKD is 15%—rates several-fold higher than in the general population. Post revascularization mortality is similarly increased—the 2-year risk of death after CABG is 53% higher with a serum creatinine ≥2.0 vs. <1.0 mg/dL. Similarly, post-PCI mortality is 1.68-fold higher in stage 4-5 CKD compared with normal function or mild CKD. In absolute terms, the 3-year post-CABG mortality rate of 34.8% in elderly CKD patients is nearly equal to the 44.9% incidence observed at 10-years in the control arms of the landmark trials comparing CABG with medical therapy and vastly exceeds the 10% incidence at 5-years in major trials comparing CABG and PCI. These striking contrasts in long-term, post-procedure outcomes, suggest that generalization from clinical trials performed in patients without CKD to the CKD population should be done with extreme caution.

Consideration of peri-procedural mortality leads to a similar conclusion. In the aforementioned trials, operative mortality was only 3.2%. Contemporary rates have fallen under 2%, but GFR is an important risk factor for peri-CABG morbidity and mortality. Indeed, operative mortality in one recent series was only 1.7% in stage 1-2 CKD but rose to 7.5 % in stage 4 CKD and 9.8% in stage 5 CKD. In other studies, the risk of operative death was 3 to 7-fold higher in patients with advanced CKD (ESRD or stage 4 CKD) than in patients without CKD. Given the shortened overall survival in patients with CKD, this level of operative mortality may effectively preclude the potential to derive long-term survival benefits from CABG.

Fatal and non-fatal procedural complications occur less frequently after PCI than CABG making an attractive alternative means of revascularization in CKD. However, target vessel restenosis is more frequent after PCI and revascularization of non-target lesions is less complete. PCI may thus provide less overall protection from CV events. As a result, repeat revascularization is needed more often after PCI than CABG, and the lower peri-procedural risks with an index PCI may be partly counterbalanced by incurred downstream risks related to subsequent procedures. In light of clinical trials that have yet to demonstrate mortality benefits of PCI compared to medical therapy or CABG and in which medium-term survival in patients without critical left main disease was not different after CABG vs. PCI, these considerations raise important concerns about a hypothetical preference for PCI over CABG or MTX in CKD.

The high risk of contrast nephropathy in advanced CKD is another important issue. Acute kidney injury (AKI) is associated with increased morbidity, costs, death, and progression to ESRD. A need for additional procedures after PCI is a non-trivial factor that could directly lead to worse long-term outcomes following PCI compared with CABG. Conversely, MTX is unlikely to cause AKI and this may be a distinct advantage of up-front medical therapy compared with PCI or CABG in CKD. Although unique to CKD, the risk of peri-procedural AKI and accelerated CKD progression is of primary importance in evaluating the merits of PCI and CABG in the setting of CKD. Avoiding dialysis is of critical importance for many patients, and in our experience individuals with CKD frequently express strong preferences for preserving kidney function over relief of angina or lowering long-term CV risks. Revascularization is frequently deferred by patients or their clinicians with a goal of avoiding ESRD and recognition that even a successful revascularization with complete relief of angina may decrease overall quality of life when dialysis-dependence is accelerated by PCI or CABG.

At the present time, only limited data are available comparing the costs, mortality benefits, risks of repeat revascularization, and risks of AKI or permanent dialysis dependence following, MTX, PCI and CABG in individuals with CKD. Extrapolating from the standard of care adopted for the general population to those with CKD is possible, but as noted above, unique, CKD-related factors make such extrapolation of questionable value. As a result, clinicians remain unable to provide evidence-based guidance for critical, patient-centered concerns regarding therapy of CAD in the

setting of CKD. CKD-specific data on both the potential benefits and the potential risks of, MTX, PCI, and CABG are critically needed or in order for clinicians to optimize treatment approaches to CAD and improve CV outcomes in the setting of CKD.

A few retrospective analyses in CKD have been performed and these have generally favored CABG over MTX or PCI. However, findings have been inconsistent. For example, an analysis of all patients in New York undergoing PCI or CABG between 1993 and 1998, demonstrated that CABG did not increase overall survival among patients with moderate to severe CKD (creatinine >2.5 mg/dL—RR 0.86, P=0.50). Similarly, the most recent analysis suggested that benefits of CABG relative to PCI are attenuated as GFR declines. Additionally, published studies suffer from deficiencies such as a small sample size, and definition of MTX as the absence of PCI or CABG rather than as optimal MTX or active MTX. They are likely to also be confounded by the selective referral of the fittest CKD patients to CABG and shunting of sicker patients to PCI or MTX—the substantial underutilization of coronary interventions in patients with CKD makes it probable that the CKD patients referred for CABG in these studies were not broadly representative of the overall CKD population. Adequate control of these factors requires methods not used in the majority of studies such as propensity score matching or the separate estimation of risk ratios within differential subgroups of baseline CV risk. Differential use of medications in patients receiving CABG, the most aggressive therapy, vs. PCI is another factor that could partly underlie observed benefits of CABG. However, despite widespread underutilization of CV medications in patients with CKD, studies comparing CABG and PCI have not adjusted for background medication use. Finally, these studies primarily predate the widespread use of drug eluting stents (DES), off-pump CABG, thienopyridines, and contemporary blood pressure or cholesterol targets—therapies with important influences on CV outcomes—and they therefore provide limited insight on the relative merits of MTX, PCI, and CABG in the contemporary era.

Furthermore, available analyses have not addressed risks of non-fatal outcomes such as MI, stroke, progression to ESRD, quality of life, costs, or cost-effectiveness. These outcomes are critical to patients, and assessing the optimal CV therapy from a patient-centered perspective requires incorporating better estimates of these risks into clinical decision making. Similarly, the disproportionate share of health care resources (particularly the Medicare budget) allocated to care of CKD, the high costs of PCI and CABG (which increase further as GFR declines), and the suboptimal allocation of U.S. health-care resources suggest that better estimates of the comparative costs and cost-effectiveness of MTX, PCI and CABG in the setting of CKD will have an important impact on health-care utilization and delivery in a growing, and costly population of patients.

In summary, despite the high incidence of CVD in CKD and well-designed clinical trials conducted in the general population, definitive, high-quality evidence comparing, MTX, CABG and PCI in patients with CKD is lacking. Existing studies have important limitations and are not informative about important non-fatal outcomes such as CKD progression, costs of care and quality of life. Given the high prevalence of CKD, the expected growth in the prevalence of advanced CKD, the high risk of CV death in ESRD, and the widespread use of these costly and potentially harmful therapies, better evidence to guide treatment decisions is urgently needed. Funding of this application will facilitate the construction of a unique, multi-disciplinary project team, a well-powered novel database that will facilitate well-adjusted, comparative analyses of the risks of death, ESRD, costs cost-effectiveness and overall quality of life of CV therapies in patients with CKD thereby providing answers to key clinical questions which are currently inadequately addressed. Although large-scale randomized trials specifically enrolling patients with advanced CKD are desirable, it is unlikely that they will be designed to answer all of these questions, funded, or completed in the near future. Furthermore, the proposed analyses will optimize choices of appropriate endpoints, estimates of event rates, and the precision of power calculations thereby improving the design of the necessary randomized trials in this population. Failure to fund this proposal, on the other hand, will perpetuate a situation in which nephrologists, cardiac surgeons and cardiologists remain unable to provide evidence-based treatment of CAD in patients with CKD.

2. Has an Institutional Review Board (IRB) reviewed your project?

☒ Yes, a copy of the approval letter and protocol must be included with the application package on IRBNet

☐ No, this project is not human subject research and does not require IRB review.

3. **Research Methodology**: Applicants must provide a written description of the project methodology (typically 1-2 pages), which should state the project objectives and/or identify relevant research questions. This document must be included with the application package on IRBNet, and must provide sufficient detail to allow CHIA to understand how the data will be used to meet objectives or address research questions. Applications that do not include this methodology statement cannot be reviewed or approved.

Please see document MDAC-DCharytan-Research-Meth-SecII3.docx

III. DATA FILES REQUESTED [*Applicants seeking 2015 data only should skip to Question 2*]

1. *FY 2004 – 2014 Data*: Please indicate the Case Mix files from which you seek data, the Level(s), the year(s) of data requested, and your justification for requesting *each* file. Please refer to the Case Mix Data Specifications for details of the file contents.

CASE MIX FILES	Levels 1 – 6 All Levels contain Core Elements plus the following in each Level	Years Available 2004 - 2014
Hospital Inpatient Discharge Database	☐Level 1: 3 Digit Zip Code, YYYYMM of Admission; Discharge; Significant Procedures	**Year(s) of Data Requested:**
	☐Level 2: 5 Digit Zip Code, Unique Physician Number (UPN), YYYYMM of Admission; Discharge; Significant Procedures	
	☐Level 3: 5 Digit Zip Code, Unique Health Information Number (UHIN), YYYYMM of Admission; Discharge; Significant Procedures	FY2013, FY2014
	☐Level 4: 5 Digit Zip Code, UHIN, UPN, YYYYMM of Admission; Discharge; Significant Procedures	
	☒Level 5: 5 Digit Zip Code , UHIN, UPN, YYYYMMDD of Admission; Discharge; Significant Procedures	
	Please describe how your research objectives require the requested Level of Hosptial Inpatient Discharge data: The inpatient data has diagnosis codes for AMI and cancer that are not available in the MASS-DAC PCI and CABG registry. CHIA data will be used to find a cancer diagnosis at the time of PCI or CABG and it will allow us to find subsequent cases where the patient had an AMI or later cancer diagnosis after the index PCI or CABG procedure. Please refer to MDAC-DCharytan-Research-Meth-SecII3.docx referenced in Section II.3.	
Outpatient Observation Database	☐Level 1: 3 Digit Zip Code, YYYYMM of Admission; Discharge; Significant Procedures	**Year(s) of Data Requested:**
	☐Level 2: 5 Digit Zip Code, Unique Physician Number (UPN), YYYYMM of Admission; Discharge; Significant Procedures	

	☐Level 3: 5 Digit Zip Code, Unique Health Information Number (UHIN), YYYYMM of Admission; Discharge; Significant Procedures	
	☐Level 4: 5 Digit Zip Code, UHIN, UPN, YYYYMM of Admission; Discharge; Significant Procedures	FY2013, FY2014
	☒Level 5: 5 Digit Zip Code , UHIN, UPN, YYYYMMDD of Admission; Discharge; Significant Procedures	
	Please describe how your research objectives require the requested Level of Outpatient Observation data: The outpatient data has diagnosis codes for AMI and cancer that are not available in the MASS-DAC PCI and CABG registry. CHIA data will be used to find a cancer diagnosis at the time of PCI or CABG and it will allow us to find subsequent cases where the patient had an AMI or later cancer diagnosis after the index PCI or CABG procedure. Please refer to MDAC-DCharytan-Research-Meth-SecII3.docx referenced in Section II.3.	
Emergency Department Database	☐Level 1: 3 Digit Zip Code, YYYYMM of Admission; Discharge; Significant Procedures	**Year(s) of Data Requested:**
	☐Level 2: 5 Digit Zip Code, Unique Physician Number (UPN), YYYYMM of Admission; Discharge; Significant Procedures	
	☐Level 3: 5 Digit Zip Code, Unique Health Information Number (UHIN), YYYYMM of Admission; Discharge; Significant Procedures	
	☐Level 4: 5 Digit Zip Code, UHIN, UPN, YYYYMM of Admission; Discharge; Significant Procedures	
	☐Level 5: 5 Digit Zip Code , UHIN, UPN, YYYYMMDD of Admission; Discharge; Significant Procedures	
	Please describe how your research objectives require the requested Level of Emergency Department data:	

2. *FY 2015 Data*: Beginning *with fiscal year 2015, Massachusetts Acute Care Hospital and Case Mix and Charge Data (collectively Case Mix Data) are released in* ***Limited Data Set (LDS) files***. Please refer to the Case Mix Data Specifications for details of the file contents.

Please indicate the Case Mix files from which you seek data, the year(s) of data requested, and your justification for requesting *each* file.

CASE MIX LIMITED DATA SET FILES	Year(s) Of Data Requested Current Yrs. Available in LDS ☐ 2015
☐**Hospital Inpatient Discharge Database**	**Please describe how your research objectives require Inpatient Discharge data:**
☐ **Outpatient Observation Database**	**Please describe how your research objectives require Outpatient Observation data:**
☐ **Emergency Department Database**	**Please describe how your research objectives require Emergency Department data:**

Sections IV-IX must be completed by all Applicants requesting 2015 data. Applications that only include requests for prior years of data can skip to Section X.

IV. GEOGRAPHIC DETAIL

Limited Data Set files include zip codes in the following formats for CT, MA, ME, NH, RI, VT, and NY only. Please choose *one* of the following geographic options.

☐ 3 Digit Zip Code (Standard)	☐ 3 Digit Zip Code & City/Town ***	☐ 5 Digit Zip Code ***	☐ 5 Digit Zip Code & City/Town ***
***Please provide justification for the chosen level of geographic detail if requesting something other than 3-Digit Zip Code only. Refer to specifics in your methodology:**			

V. DEMOGRAPHIC DETAIL

Please choose *one* of the following demographic options:

☐ Not Requested (Standard)	☐ Race & Ethnicity***
***** If requested please, provide justification for requesting Race and Ethnicity. Refer to specifics in your methodology:**	

VI. DATE DETAIL

Please choose *one* option from the following options for dates of admissions, discharges, and significant procedures:

☐ Year (YYYY)(Standard)	☐ Month (YYYYMM) ***	☐ Day (YYYYMMDD)***
*****Please provide justification for the chosen level of date detail if requesting Month or Day. Refer to specifics in your methodology:**		

VII. PHYSICIAN IDENTIFICATION NUMBERS (UPN)

Please choose *one* of the following options for Provider Identifier(s):

☐ Not Requested (Standard)	☐ Hashed ID ***	☐ Board of Registration in Medicine # (BORIM) ***
*****If requested please, provide justification for requesting Hashed ID or BORIM #. Refer to specifics in your methodology:**		

VIII. HASHED UNIQUE HEALTH IDENTIFICATION NUMBER (UHIN)

Please choose *one* of the following:

☐ Not Requested (Standard)	☐ UHIN Requested ***
***** If requested please, provide justification for requesting UHIN. Refer to specifics in your methodology:**	

IX. HASHED MOTHER'S SOCIAL SECURITY NUMBER

Please choose *one* of the following:

☐ Not Requested (Standard)	☐ Hashed Mother's SSN Requested ***
***** If requested please, provide justification for requesting Hashed Mother's SSN. Refer to specifics in your methodology:**	

X. DATA LINKAGE AND FURTHER DATA ABSTRACTION

Note: Data linkage involves combining CHIA data with other databases to create a more extensive database for analysis. Data linkage is typically used to link multiple events or characteristics within one database that refer to a single person within CHIA data.

1. Do you intend to link or merge CHIA Data to other datasets?

☒ Yes

☐ No linkage or merger with any other database will occur

2. If yes, please indicate below the types of database to which CHIA Data be linked. [Check all that apply]

☒ Individual Patient Level Data (e.g. disease registries, death data)

☐ Individual Provider Level Data (e.g., American Medical Association Physician Masterfile)

☐ Individual Facility Level Data level (e.g., American Hospital Association data)

☐ Aggregate Data (e.g., Census data)

☐ Other (please describe):

3. If yes, describe the data base(s) to which the CHIA Data will be linked, which CHIA data elements will be linked; and the purpose for the linkage(s):

The purposes of the linkage In this research proposal is to determine long-term outcomes and subsequent conditions for a patient from an index procedures documented in the Mass-DAC data sets. Hospitals submit fully identified patient level clinical data for PCI and cardiac surgery procedures to Mass-DAC on a quarterly basis. CHIA Case Mix data contains billing data for the same patients and includes additional diagnosis and procedures with information not available in the Mass-DAC data collection process.

When a link is found, we will use the valid CHIA case mix unique patient identifier number (UHIN) to find subsequent diagnosises related to the index PCI and CABG procedures, including subsequent acute myocardial infarctions, cancer and chronic kidney disease. For patients in both the Mass-DAC and case mix data sets, we can then identify an index procedure and subsequent procedures to determine outcomes related to the safety and effectiveness of medical devices used in patients with cardiac surgery or PCI procedures.

The final data set for the analysis will only contain flag variables indicating if a condition was present, 1=present, 0=not present. None of the CHIA case mix data fields will be stored with the Mass-DAC data.

Data Flow Process:

The steps below describe the major methods used in the linkage process. Each sub-process uses brackets [] to identify the related container in **Error! Reference source not found.** found in response to question 4.

1. **Mass-DAC Subset Data [FP.1.1]:** Create a subset [S.1.a] from the master Mass-DAC harmonized PCI and cardiac surgery data registries [D.1] that contains PCI and CABG records for current fiscal year.

The following lists the minimum set of fields needed to complete the merge.

a.	Site:	Short name identifies hospital
b.	AdmitDate	Admission date to hospital
c.	ProcDate/SurgDate	Date PCI or CABG procedure was done
d.	DschDate	Discharge date from hospital
e.	DschStatus	Patient status at discharge, alive or dead
f.	PatID_MDAC	Unique patient identifier created for Mass-DAC records
g.	RecordID_MDAC	Unique record id within the PCI and cardiac surgery data
h.	DOB_Submit	Patient date of birth submitted by hospital
i.	DOB_MDAC	Patient date of birth assigned by Mass-DAC
j.	MedRecN	Hospital patient medical record number
k.	Gender	Patient gender

2. **CHIA Subset Data [FP.1.2]:** Create two subsets, one with PCI and CABG patients only [S.2.a] and one with all valid UHINs and associated admission, procedure, and discharge dates, diagnosis/DRG codes and discharge disposition [S.2.b]. All three CHIA databases are used to maximize the chance of finding a PCI or CABG record after the Mass-DAC initial procedure date. UHINs for the PCI and CABG cases are identified in the inpatient and outpatient observation room databases [D.2] where at least one of 15 procedure codes or principal procedure contains an ICD9-CM code for a PCI procedure (3601-3607) or CABG surgery (3610-3619) or a CPT-4 code (92980-92994, 92995-92996, and 92920-92944 *[for 2013+]*) for a PCI procedure. The following lists the minimum set of fields needed to complete the merge.

a.	MDPHHospNum:	MDPH hospital number determined from all hospital IDs in the case mix data
b.	AdmitDt	Admission date to hospital
c.	ProcDates	Possible dates for PCI or CABG procedure *(15 in inpatient, 3 in OOR data)*
d.	ProcedureCodes	ICD-9-CM and CPT (OOR only) codes for PCI and CABG records *(only 3 ICD-9-CM in OOR data)*
e.	PrincipalProcDate	Date for principal procedure (OOR only)
f.	PrincipalProcedure	ICD9-CM code for PCI and CABG records
g.	Diagnosis Codes	Diagnosis and DRG codes for conditions
h.	DischDate	Discharge date from hospital
i.	DischDispo	Patient status at discharge, alive or dead

j. RecordType20ID CHIA Record Id Control Number
k. UHIN Unique patient identifier from CHIA
l. DOB Patient date of birth submitted by hospital
m. MedicalRecordNum Hospital patient medical record number
n. Gender Patient gender

The subset [S.2.b] contains the minimum set of fields needed to determine a subsequent episode of care. The fields saved in the temporary data set are UHIN, date of birth, gender, admission date, latest procedure date, discharge date, and discharge disposition. Records with an invalid UHIN are excluded, (i.e., UHIN is '-', ' ', or '000000001').

3. **Link Subsets to get PCI and CABG UHINs [FP.1.3]:** Perform multiple links with Mass-DAC subset [S.1.a] and CHIA Case Mix subset [S.2.a] keeping all records that merge successfully in a temporary data set [D.3]. A successful record merge is determined by one of the field sets listed below.

   a. Site, MedRecN, AdmitDate
   b. Site, MedRecN, DischDate
   c. Site, MedRecN, DOB, ProcDates
   d. Site, DOBs, AdmitDate, DischDate, ProcDates
   e. Site, DOBs, AdmitDate, DischDate
   f. Site, DOBs, AdmitDate, ProcDates
   g. Site, DOBs, DischDate, ProcDates

   The temporary data set [D.3] contains the minimum set of fields needed to find subsequent patient episodes of care in the CHIA databases using UHINs. The fields saved are UHIN, date of birth, gender, Mass-DAC procedure date, Mass-DAC unique patient identifier, and Mass-DAC unique record identifier.

4. **Link Data to CHIA UHINs to get long-term outcome [FP.1.4]:** The merged PCI and CABG data set, [D.3] UHINs are merged with UHIN subset [S.2.b] on UHIN number. For all records that match on UHIN, only records where the CHIA procedure date is after the Mass-DAC procedure date are retained to determine long-term outcome. . If a subsequent episode of care is found, then the new Mass-DAC field is created that will flag the long-term outcome (e.g., Cancer), 0=Absent, 1=Present. A subsequent episode of care is found in less than 20% of cases, so most remain unknown in the

resulting data set.

The following fields are saved in the results data set [D.4] used in the Mass-DAC analysis database.

a. PatID_MDAC Unique patient identifier created for Mass-DAC records
b. RecordID_MDAC Unique record id within the PCI and cardiac surgery data
c. Flag_outcome Flag whether the condition is present or absent after the index procedure.

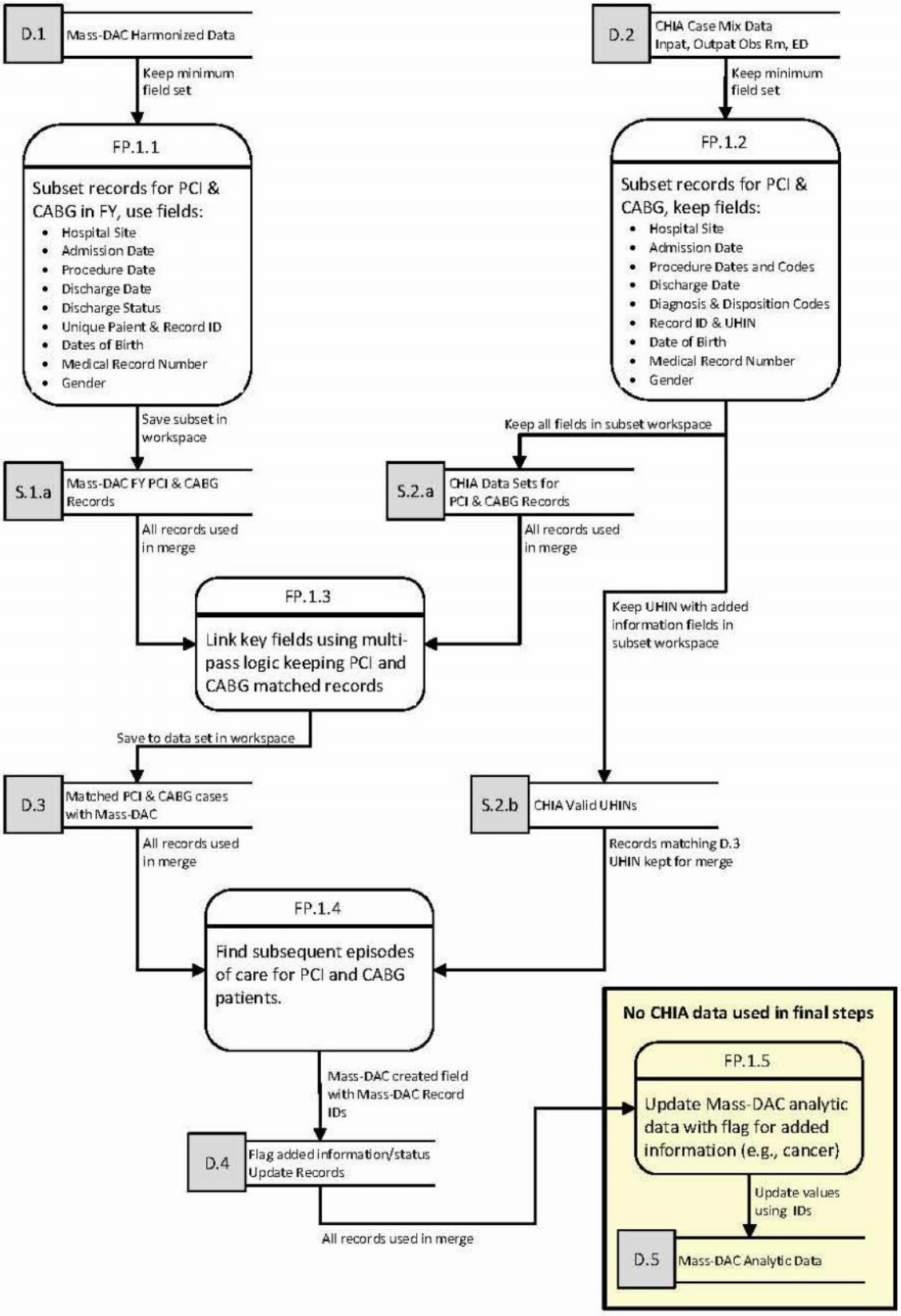
5. **Update Mass-DAC Registry Databases [FP.1.5]:** Use the results data set [D.4.] to update the Mass-DAC analysis database. **No CHIA data fields or records are merged with the Mass-DAC registry data, which contain patient identifiers.**

4. If yes, for each proposed linkage above, please describe your method or selected algorithm (e.g., deterministic or probabilistic) for linking each dataset. If you intend to develop a unique algorithm, please describe how it will link each dataset.

It is deterministic. A flowchart of how this is done is included on he following page.

CHIA Case Mix Linkage with Mass-DAC Data *Revision 2015.06.22*

Flowchart 1: Linkage Process between Case Mix and Mass-DAC Data



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

The flow chart in question 4 identifies the steps taken to prevent identification in the final analytic file. The linkage is done on a stand-alone PC which is not used by the analysts. The linkage process uses only tempoaray data sets that are immediately removed once the program has completed. The final analytic data set does not contain any patient names, address, or social security numbers. It also does not contain any CHIA case mix data fields.

6**.** Once the linkage/merge is made, what non-MA Case Mix data elements will appear in the new linked file?

The final analytic data set will contain many Mass-DAC data along with the analytic variables created during the link process that identify a cancer or AMI condition is present. No case mix variables are saved in the final analytic data set.

XI. PUBLICATION / DISSEMINATION / RE-RELEASE

1. Describe your plans to publish or otherwise disclose CHIA Data, or any data derived or extracted from such CHIA Data, in any paper, report, website, statistical tabulation, seminar, conference, or other setting. All publication of CHIA Data must comply with CHIA's cell size suppression policy, as set forth in the Data Use Agreement. Please explain how you will ensure that any publications will not display a cell less than 11, and no percentages or other mathematical formulas will be used if they result in the display of a cell less than 11.

We plan to use derived aggregate data only in presentations and manuscripts for both public and internal use. Cell suppression (<11 cases) will be used for any CHIA data disclosed, or derived, or extracted.

2. Do you anticipate that the results of your analysis will be published and/or publically available to any interested party? Please describe how an interested party will obtain your analysis and, if applicable, the amount of the fee, that the third party must pay.

Aggregate data only would be used in a manuscript or presentation that would be available publicly.

3. Will you use CHIA Data for consulting purposes?

☐ Yes
☒ No

4. Will you be selling standard report products using CHIA Data?

☐ Yes
☒ No

5. Will you be selling a software product using CHIA Data?

☐ Yes
☒ No

6. Will you be reselling CHIA Data in any format?

☐ Yes

☒ No

If yes, in what format will you be reselling CHIA Data (e.g., as a standalone product, incorporated with a software product, with a subscription, etc.)?

N/A

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

N/A

8. If you have answered "yes" to questions 4, 5, or 6, what is the fee you will charge for such products, services or studies?

N/A

XII. APPLICANT QUALIFICATIONS

1. Describe your qualifications (and the qualifications of your co-investigators) to perform the research described.

Sharon-Lise Normand, Ph.D., is Professor of Health Care Policy (Biostatistics), Department of Health Care Policy, Harvard Medical School; and Professor of Biostatistics, Department of Biostatistics, Harvard School of Public Health. She has been the Director of Mass-DAC since 2002. Dr. Normand has been utilizing claims and clinical registry data for more than two decades; was a member of the Harvard Medical School IRB for several years; and serves as an investigator to the Center for Devices and Radiological Health at the U.S. Food and Drug Administration.

2. **Attach** résumés or curricula vitae of the Applicant/principal investigator, and co-investigators. (These attachments will not be posted on the internet.)

XIII. USE OF AGENTS AND/OR CONTRACTORS

Please note: by signing this Application, the Organization assumes all responsibility for the use, security and maintenance of the CHIA Data by its agents, including but not limited to contractors.

Provide the following information for all agents and contractors who will work with the CHIA Data. *Add agents or contractors as needed.*

Company Name:	No Agents or Contractors are used
Contact Person:	
Title:	
Address, City/Town, Zip Code	
Telephone Number:	
E-mail Address:	
Organization Website:	

1.Will the agent or contractor have access to or store the CHIA Data at a location other than the Applicant's location, off-site server and/or database?

☐ Yes, a separate Data Management Plan **must** be completed by each agent or contractor
☒ No

2. Describe the tasks and products assigned to this agent for this project; their qualifications for completing the tasks; and the Organization's oversight of the agent, including how the Organization will ensure the security of the CHIA Data to which the agent has access.

Company Name:	
Contact Person:	
Title:	
Address, City/Town, Zip Code	
Telephone Number:	
E-mail Address:	
Organization Website:	

1.Will the agent or contractor have access to or store the CHIA Data at a location other than the Applicant's location, off-site server and/or database?

☐ Yes, a separate Data Management Plan **must** be completed by each agent or contractor
☐ No

2. Describe the tasks and products assigned to this agent for this project; their qualifications for completing the tasks; and the Organization's oversight of the agent, including how the Organization will ensure the security of the CHIA Data to which the agent has access.

XIV. FEE INFORMATION

Please consult the fee schedules for Case Mix Data and select from the following options:

☒ Single Use
☐ Limited Multiple Use
☐ Multiple Use

Are you requesting a fee waiver?

☒ Yes
☐ No

If yes, please refer to the Application Fee Remittance Form and submit a letter stating the basis for your request (if required). Please refer to the fee schedule for qualifications for receiving a fee waiver. If you are requesting a waiver based on the financial hardship provision, please provide documentation of your financial situation. Please note that non-profit status alone isn't sufficient to qualify for a fee waiver.

XV. ATTESTATION

By submitting this Application, the Data Applicant attests that it is aware of its data use, privacy and security obligations imposed by state and federal law *and* is compliant with such use, privacy and security standards. The Data Applicant further agrees and understands that it is solely responsible for any breaches or unauthorized access, disclosure or use of any CHIA Data provided in connection with an approved Application, including, but not limited to, any breach or unauthorized access, disclosure or use by its agents.

Applicants requesting data from CHIA will be provided with data following the execution of a Data Use Agreement that requires the Data Applicant to adhere to processes and procedures aimed at preventing unauthorized access, disclosure or use of data.

By my signature below, I attest to: (1) the accuracy of the information provided herein; (2) that the requested data is the minimum necessary to accomplish the purposes described herein; (3) the Data Applicant will meet the data privacy and security requirements describe in this Application and supporting documents, and will ensure that any third party with access to the data meets the data use, privacy and security requirements; and (4) my authority to bind the organization seeking CHIA Data for the purposes described herein.

Signature: (Authorized Agent)	
Printed Name :	
Title:	
Signature (Applicant/Primary Investigator)	
Name:	Sharon-Lise Normand
Title:	
Original Data Request Submission Date:	
Dates Data Request Revised:	

Attachments. Please indicate below which documents have been attached to the Application and uploaded to IRBNet:

☐ 1. IRB approval letter and protocol (if applicable)

☐ 2. 1-2 page Research Methodology

☐ 3. Resumes of Applicant and co-investigators

☐ 4. Data Management Plan (including one for each agent of contractor that will have access to or store the CHIA Data at a location other than the Applicant's location, off-site server and/or database)

☐ 5. Fee Remittance Form (including any required documentation if a fee waiver is being requested)