

CENTER FOR HEALTH INFORMATION AND ANALYSIS

OVERVIEW OF THE MASSACHUSETTS

**ACUTE HOSPITAL
CASE MIX
DATABASES**

DECEMBER 2019



Overview of the Massachusetts Acute Hospital Case Mix Databases (December 2019)

Table of Contents

Introduction 3

Background of the Massachusetts Acute Hospital Case Mix Databases 3

Data Governance 4

Data Submission 4

Data Enhancement 4

Case Mix Compared to the All-Payer Claims Database 5

Data Release 5

Uses of Case Mix Data 6

Appendix A 8

Appendix B 11

Appendix C 15

Introduction

The mission of the Center for Health Information and Analysis (CHIA) is to be the agency of record for Massachusetts health care information, to responsibly steward sensitive and confidential data, and to objectively report reliable and meaningful information about the quality, affordability, utilization, access, and outcomes of the Massachusetts health care system. In support of this mission, CHIA offers a variety of data and analytic products to support continual improvement in health care quality, affordability, access, and outcomes. The Massachusetts Acute Hospital Case Mix Databases (Case Mix) are critical to this effort, enabling CHIA, its partners and stakeholders to pursue a wide variety of projects, including complex research and analyses that support state agency operations and enhance the ability of payers and providers to deliver care.

This paper provides an overview of the Case Mix databases, as well as how CHIA manages and uses the databases to promote the public interest.

Background of the Massachusetts Acute Hospital Case Mix Databases

Section 8 of Chapter 12C of the Massachusetts General Laws grants CHIA authority to collect data from Massachusetts hospitals. CHIA, and its predecessor agency the Division of Health Care Finance and Policy, have collected patient-level data from Massachusetts acute care hospitals for more than twenty years. The following three main data types constitute the Case Mix database:

1. **Hospital Inpatient Discharge Database (HIDD)**—contains discharge-level inpatient diagnostic data which describe the socio-demographic characteristics of the patient, the reason for the admission, the treatment and services provided to the patient, the charges, and the duration of the patient's stay.
2. **Outpatient Observations Database (OOD)**—consists of discharge-level outpatient observation stay data which, like the HIDD, have socio-demographic information, reason for the observation, the treatment and services provided, the charges, and the duration of stay.
3. **Outpatient Emergency Department Database (ED)**—contains visit-level data for all emergency department visits. Similar to the HIDD and OOD, each record includes socio-demographic information, medical reason for the visit, the procedures and services provided to the patient, the charges, and the duration of the visit.

Since 1988, state acute hospital Case Mix data—including the databases from Massachusetts—has been aggregated by the Agency for Healthcare Research and Quality (AHRQ) and released as the Healthcare Cost and Utilization Project (HCUP). [Appendix C](#) describes the data available from HCUP.

A key distinguishing feature of Case Mix is the number of years it has been collected in a consistent format. While some data elements have been added over the years, Case Mix can be used to observe longitudinal trends in acute hospital utilization which is useful for cost trends, epidemiological, and health planning studies.

Data Governance

Pursuant to Chapter 12C, CHIA issued Regulation 957 CMR 8.00 to govern the submission of Case Mix and charge data by Massachusetts hospitals. Administrative bulletins and submission guides issued by CHIA further detail data elements that hospitals are required to submit.

CHIA collaborates with hospitals on data definitions and submission standards through periodic Technical Advisory Groups.

Data Submission

Regulation 957.8.03(3) outlines data submission requirements for the Case Mix files. Massachusetts acute care hospitals, non-acute hospitals and psychiatric facilities are required to submit data electronically on a quarterly basis to CHIA following data submission specifications. Hospitals are required to report data within 75 days following the close of the reporting period. Each data element in the submission specifications is characterized by its importance to CHIA and Case Mix data users. CHIA assigns an “A” level to data elements that are the most important and a “B” level to less important elements.

Submissions must meet quality standards set by CHIA. CHIA’s intake systems electronically check each file prior to acceptance for valid codes, correct formatting, and the presence of required data elements. Hospitals receive an error report that displays invalid discharge information. Generally speaking, if more than 1% of the records in a file have errors,¹ the files are rejected and hospitals must re-submit until the standard is met.

In addition, hospitals may amend their submissions in accordance with 957 CMR 8.04(4). Profile reports are available to the hospitals on a quarterly basis to enable the hospitals to correct and resubmit their data. Hospitals are provided summary reports of their submissions for their review and verification on an annual basis. Their comments are included in an annual documentation manual which accompanies the data made available to applicants under 957 CMR 5.00: Health Care Claims, Case Mix and Charge Data Release Procedures.

Data Enhancement

CHIA enhances data submitted by the hospitals to increase its analytic value while maintaining patient privacy. Examples of such enhancements include:

- Encryption of patient social security number into a Unique Health Information Number (UHIN). The encrypted UHIN allows for tracking of patients across facilities.
- Application of the following classification Diagnosis Related Grouping (DRG) systems:
 - 3M’s All-Patient DRG;
 - 3M’s All-Patient Refined DRG—which expands the basis DRG structure to reflect the patient’s severity of illness and risk of mortality; and

¹ An error in one “A” field or errors in two “B” fields causes rejection of a discharge record.

- CMS's Medicare DRG.
- Calculation of a field that indicates the chronological order of inpatient admissions for patients (unique UHINs) with multiple stays in a given reporting year (e.g., first admission, second admission). This allows analysts to examine strings of multiple admissions without date of service.
- Calculation of pre-operative days for each procedure.
- Calculation of the number of days between each discharge and each consecutive admission for each UHIN. Again, it allows analysts to examine strings of multiple admissions without dates of service.

Case Mix Compared to the All-Payer Claims Database

In addition to patient-level data submitted by hospitals for Case Mix, CHIA collects claim-level data from payers for the Massachusetts All-Payer Claims Database (MA APCD). While the hospital and payer data are complementary, they differ in a few notable ways:

- Hospitals report charges billed to payers in their Case Mix submissions while payers report paid amounts in addition to charges in the MA APCD.
- Race and ethnicity data are collected by hospitals in a standardized methodology and reported to CHIA. While there are Race and Ethnicity data elements in the MA APCD, payers have not been able to consistently submit these data to CHIA.
- Massachusetts hospitals submit information on discharges, emergency department visits, and observation stays, including services paid by Workers Compensation, auto and other non-medical insurance programs, as well as services paid by TriCare and the Federal Employee Benefits of Veterans Affairs. These types of insurers do not presently report to the MA APCD.

Data Release

In order to realize the highest value of the Hospital Case Mix and Charge Data, CHIA is committed to releasing data extracts to a wide variety of users. Requests to use CHIA data are considered through a multi-layer process that considers risks to patient privacy, research or project objectives, and the public interest.

Release is determined by *who* is requesting the data, their *purpose*, and *the type* of data sought. Pre-FY 2015 data elements in the Case Mix and Charge Data are grouped into six levels to facilitate data extraction and release. Subsequent years' data are only available as Limited Data Sets (LDS), and non-government users may request one of several LDS.² The LDS excludes HIPAA-defined direct identifiers and certain quasi-identifiers. All applicants receive the "Core" LDS with options, depending on the purpose for which the data are sought, for date granularity, geographic granularity, and physician ID numbers to applicants. See [Appendix A](#) for a complete listing of available data elements. Government Users may request all data elements, including those with direct identifiers.

² Data for FY 2014 and earlier are organized into six levels or groupings. See [Appendix B](#) for details.

All applications for the Case Mix data—including those from Massachusetts state agencies—are reviewed by the CHIA Data Privacy Committee. The committee, which is composed of CHIA staff, reviews each application for compliance with state and federal privacy laws, including establishing that the data sought is the minimum amount necessary to achieve a specific public purpose. The Data Privacy Committee is chaired by CHIA's Chief Privacy Officer and staffed by the Chief Information Security Officer and other data specialists.

Certain non-governmental applications, including any applications from researchers and consultants, are further reviewed by the Data Release Committee, a group of experts representing carriers, providers, researchers, and consumers. This Data Release Committee builds on the work of the Data Privacy Committee by considering whether the request is in the public interest as defined by CHIA regulations. These applications are posted to the CHIA [website](#) to promote transparency.

The final decision for release is made by the Executive Director based on the recommendations of these committees. Before receiving the data, applicants are required to sign CHIA's Data Use Agreement, which imposes requirements for maintaining data security and protecting patient privacy.

Uses of Case Mix Data

Many organizations—both government and non-government—rely on Case Mix data for research, policy development, and evaluation. Here are several examples that demonstrate the breadth of its use:

- **The Massachusetts Department of Public Health has used Case Mix data to support a variety of its programmatic and research efforts:** The **Health and Disability Program** identified a cohort of individuals with disabilities to track for hospitalizations over time as part of its mission to promote the health and well-being of people with disabilities in Massachusetts. The **Occupational Health Surveillance Program** monitors Case Mix files to identify potential cases with work-related health conditions and work-related injuries. The **Bureau of Health Statistics** links Case Mix data to birth certificate data and health plan data to examine the completeness and quality of responses on data sets as well as to develop performance measures. The **Bureau of Family and Community Health** identifies maternal morbidities, pregnancy-associated mortalities in their Maternal Mortality and Morbidity Reviews. The **Bureau of Substance Abuse Services** analyzes treatment needs and gaps in services.
- **Massachusetts Health Data Consortium** and several analytic firms have developed a reporting product that hospitals use for planning and benchmarking. Additionally, all of the academic medical centers in Massachusetts, as well as the surrounding states, request Case Mix data directly for service area planning, benchmarking, and population health management.
- **Harvard University Kennedy School of Government** studied trends in health care utilization before and after the Massachusetts health reform efforts, and evaluated whether expanded coverage increased demand for medical care and made access to care more difficult.

A complete list of non-government uses is available through CHIA's website:

<http://www.chiamass.gov/case-mix-applications-received-and-under/>

For further information on CHIA's Case Mix databases, please refer to the following sources:

CHIA website

The Case Mix [webpage](#) hosts a wealth of documentation about Case Mix data annual releases. Release documentation provides details at the data element level as well as information about intake edits and standardization.

Case Mix Data Submission Guides

Some users also avail themselves of the Data Submission Guides. While these guides are intended for the providers that submit Case Mix data, they include information useful to users including record specifications, data elements definitions and data record descriptions.

<http://www.chiamass.gov/information-for-data-submitters-acute-hospital-case-mix-data/>

<http://www.chiamass.gov/behavioral-health-facilities-case-mix-data/>

Appendix A

Data Available to Non-Government Users in Limited Data Sets from FY 2015

Admission Day of Week
Admission Source Code 1, 2
Admission Type
Admission Year
Age LDS
Birth weight in grams
Condition Present ECode
Days Between Stays
Discharge Year
Discharge Day Of Week
Discharge Passed
DRGs
Ecode (External Cause of Injury Code)
ED Flag Code (Emergency Department Flag)
Hispanic Indicator
Homeless Indicator
ICD Indicator
IdOrgFiler (Identifies filer organization)
IdOrgHosp (Identifies the main hospital organization)
IdOrgSite (Identifies the clinical site of care)
IdOrgTransfer (Identifies transfer organization)
Leave Of Absence Days
Length of Stay

New Born Age
Number of ANDs (Administratively Necessary Days)
Number of Diagnosis Codes
Number of Hours in Emergency Department
Number of Procedure Codes
Other Care Giver Code
Outpatient Observation Stay Flag Code
Patient Status
Permanent Patient State of Residence (MA and border states only)
Permanent Patient ZIP Code (3 digit MA only)
Permanent Patient Country (USA [US], Mexico [MX], Canada [CN] – All Others ZZ)
Payer Code 1, 2
Period Ending Year
Period Starting Year
Principal Preoperative Days
Primary Condition Present
Primary Diagnosis Code and Associated Diagnosis Codes
Primary Payer Type
Principal Procedure Code and Other Procedure Codes
Quarter
Record Type 201D
Secondary Payer Type
Patient Sex
Special Condition Indicator

Submission Control ID
Submission Passed Flag
Temporary Patient State
Temporary Patient ZIP Code (3 digit MA only)
Total Charges All
Total Charges Ancillaries
Total Charges Routine
Total Charges Special
Year

Applicants can also request the following: Zip Code -MA only: 5 digits, Dates of Service either Month and Day or Month Only, MA Municipality Name, Physician ID numbers, Race and Ethnicity, Unique Health Information Number (encrypted patient ID), Emergency Department Registration and Discharge Dates.

Appendix B

Definition of Levels for HIDD Data for FY 2014 and Earlier

Hospital Inpatient Data Organized into Levels for Release

Data Element	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6
Admission Date	No	No	No	No	Yes	Yes
Admission Day of Week	Yes	Yes	Yes	Yes	Yes	Yes
Admission Month	Yes	Yes	Yes	Yes	Yes	Yes
Admission Source Code 1, 2	Yes	Yes	Yes	Yes	Yes	Yes
Admission Type	Yes	Yes	Yes	Yes	Yes	Yes
Patient Age	Yes	Yes	Yes	Yes	Yes	Yes
Birth weight in grams	Yes	Yes	Yes	Yes	Yes	Yes
Condition Present 1-15	Yes	Yes	Yes	Yes	Yes	Yes
Condition Present ECode	Yes	Yes	Yes	Yes	Yes	Yes
Days Between Stays	No	No	Yes	Yes	Yes	Yes
Diagnosis Code1-15*	Yes	Yes	Yes	Yes	Yes	Yes
Discharge Date	No	No	No	No	Yes	Yes
Discharge Day Of Week	Yes	Yes	Yes	Yes	Yes	Yes
Discharge ID	No	No	No	No	No	No
Discharge Month	Yes	Yes	Yes	Yes	Yes	Yes
Discharge Passed	Yes	Yes	Yes	Yes	Yes	Yes
DNR Status (Do Not Resuscitate Status)	Yes	Yes	Yes	Yes	Yes	Yes
DOB (Date of Birth)	No	No	No	No	No	Yes
Ecode (External Cause of Injury Code)	Yes	Yes	Yes	Yes	Yes	Yes
ED Flag Code (Emergency Department Flag)	Yes	Yes	Yes	Yes	Yes	Yes
Employer Zip Code	Yes	Yes	Yes	Yes	Yes	Yes
Encrypted Attending Physician Number	No	Yes	No	Yes	Yes	Yes

Encrypted Operating PhysicianP 1-14	No	Yes	No	Yes	Yes	Yes
Ethnicity 1, 2	Yes	Yes	Yes	Yes	Yes	Yes
Hispanic Indicator	Yes	Yes	Yes	Yes	Yes	Yes
Homeless Indicator	Yes	Yes	Yes	Yes	Yes	Yes
Hospital Bill Number	No	No	No	No	No	Yes
IdOrgFiler (Identifies filer organization)	Yes	Yes	Yes	Yes	Yes	Yes
IdOrgHosp (Identifies the main hospital organization)	Yes	Yes	Yes	Yes	Yes	Yes
IdOrgSite (Identifies the clinical site of care)	Yes	Yes	Yes	Yes	Yes	Yes
IdOrgTransfer (Identifies transfer organization)	Yes	Yes	Yes	Yes	Yes	Yes
Leave Of Absence Days	Yes	Yes	Yes	Yes	Yes	Yes
Length of Stay	Yes	Yes	Yes	Yes	Yes	Yes
Medical Record Number	No	No	No	No	No	Yes
Mother's Medical Record Number	No	No	No	No	No	Yes
Mother's SSN (encrypted)	No	No	Yes	Yes	Yes	Yes
New Born Age	Yes	Yes	Yes	Yes	Yes	Yes
Number of ANDs (Administratively Necessary Days)	Yes	Yes	Yes	Yes	Yes	Yes
Number of Diagnosis Codes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Procedure Codes	Yes	Yes	Yes	Yes	Yes	Yes
Other Care Giver Code	Yes	Yes	Yes	Yes	Yes	Yes
Other Ethnicity Code	Yes	Yes	Yes	Yes	Yes	Yes
Other Race	Yes	Yes	Yes	Yes	Yes	Yes
Outpatient Observation Stay Flag Code	Yes	Yes	Yes	Yes	Yes	Yes
Patient Status	Yes	Yes	Yes	Yes	Yes	Yes
Payer Code 1, 2	Yes	Yes	Yes	Yes	Yes	Yes
Period Ending Date	Yes	Yes	Yes	Yes	Yes	Yes

Period Starting Date	Yes	Yes	Yes	Yes	Yes	Yes
Permanent Street Address	No	No	No	No	No	No
Permanent Patient City	Yes	Yes	Yes	Yes	Yes	Yes
Permanent Patient Country	Yes	Yes	Yes	Yes	Yes	Yes
Permanent Patient State	Yes	Yes	Yes	Yes	Yes	Yes
Permanent Patient ZIP Code	Yes	Yes	Yes	Yes	Yes	Yes
Preoperative Days 1-3	Yes	Yes	Yes	Yes	Yes	Yes
Primary Payer Type	Yes	Yes	Yes	Yes	Yes	Yes
Procedure Code 1-15*	Yes	Yes	Yes	Yes	Yes	Yes
Procedure Code Date 1-15*	No	No	No	No	Yes	Yes
Race 1 and 2	Yes	Yes	Yes	Yes	Yes	Yes
Record Type 201D	Yes	Yes	Yes	Yes	Yes	Yes
Secondary Payer Type	Yes	Yes	Yes	Yes	Yes	Yes
Patient Sex	Yes	Yes	Yes	Yes	Yes	Yes
Special Condition Indicator	Yes	Yes	Yes	Yes	Yes	Yes
Submission Control ID	Yes	Yes	Yes	Yes	Yes	Yes
Submission Passed Flag	Yes	Yes	Yes	Yes	Yes	Yes
Temporary Patient Address	No	No	No	No	No	No
Temporary Patient City	Yes	Yes	Yes	Yes	Yes	Yes
Temporary Patient State	Yes	Yes	Yes	Yes	Yes	Yes
Temporary Patient Zip Code	Yes	Yes	Yes	Yes	Yes	Yes
Total Charges All	Yes	Yes	Yes	Yes	Yes	Yes
Total Charges Ancillaries	Yes	Yes	Yes	Yes	Yes	Yes
Total Charges Routine	Yes	Yes	Yes	Yes	Yes	Yes
Total Charges Special	Yes	Yes	Yes	Yes	Yes	Yes
UHIN (Unique Health Information Number)	No	No	Yes	Yes	Yes	Yes
UHIN Sequence No	No	No	Yes	Yes	Yes	Yes

Veterans Status	Yes	Yes	Yes	Yes	Yes	Yes
-----------------	-----	-----	-----	-----	-----	-----

*Note: Diagnosis and Procedure codes are unlimited from FY 2015

Appendix C

Healthcare Cost and Utilization Project (HCUP) Databases

Presently, 48 states collect and submit hospital data to AHRQ. Data from HCUP is widely used by researchers and policy analysts. The following databases constitute the HCUP family of databases:

The Nationwide HCUP Databases

- **National Inpatient Sample (NIS)** is the largest all-payer inpatient care database containing over seven million hospital stays. The NIS comprises data elements in the Core and 3 or more additional files at the following link: <https://www.hcup-us.ahrq.gov/db/nation/nis/nisdde.jsp>
- **Kids' Inpatient Database (KID)** is the only all-payer pediatric inpatient care database in the United States and has two to three million hospital stays each year. KID's data elements in a core dataset and one or more additional files are accessible here: <https://www.hcup-us.ahrq.gov/db/nation/kid/kiddde.jsp>
- **Nationwide Emergency Department Sample (NEDS)** is the largest all-payer emergency department (ED) database in the United States that produces national estimates of hospital-based ED visits. It comprises about 31 million unadjusted ED visits per year. If adjusted, it yields an estimated 143 million ED visits. Data elements in NEDS' core data and three additional files are accessible here: <https://www.hcup-us.ahrq.gov/db/nation/neds/nedsdde.jsp>
- **Nationwide Readmission Database (NRD)** is a database designed to support different types of analyses of national readmission rates for all payers and the uninsured. The NRD description of data elements is accessible here: <https://www.hcup-us.ahrq.gov/db/nation/nrd/nrddde.jsp>

State-Specific HCUP Databases³

- **State Inpatient Databases (SID)** are state-specific files that contain all-payer inpatient records in participating states—encompassing more than 95% of all hospital discharges in the United States. The following link is the description of the SID file composition: https://www.hcup-us.ahrq.gov/db/state/siddist/sid_multi.jsp
- **State Emergency Department Databases (SEDD)** are state-specific databases that consist of information on all emergency department visits that do not result in hospitalization. Information on SEDD documentation can be accessed in the following link: <https://www.hcup-us.ahrq.gov/db/state/sedddbdocumentation.jsp>

³ State Ambulatory Surgery and Services Databases (SASD) are omitted from this document because Massachusetts does not participate in such data collection.



For more information, please contact:

CENTER FOR HEALTH INFORMATION AND ANALYSIS

501 Boylston Street
Boston, MA 02116

www.chiamass.gov
[@Mass_CHIA](https://twitter.com/Mass_CHIA)

(617) 701-8100