HEALTH CARE COST INSTITUTE

This Data Dictionary applies to the HCCI data in the data Enclave for data years 2008-2016.

Two Statistically De-identified Views are available for selection, Data Set #1 and Data Set #2. Fields available under each view are denoted with an 'X'.

DISCLAIMER: The contents of this document are for informative purposes only and could be subject to change. HCCI cannot warrant how fully populated a variable listed below may be.
Note: Two statistically de-identified data views are available for research, Data Set #1 and Data Set #2. To ensure the data remains de-identified, researchers may only request a single data set, Data Set #1 OR Data Set #2, for a research project. The differences between Data Set #1 and #2 are age and geography.

Member Enrollment Fields

Data Set #1	Data Set #2	Field Name	Field ID	Data Type	Description
X	X	Patient Identifier (encrypted)	Z_PATID	Integer	Encrypted, unique identifier for all members in data set. PATID is consistent over time and unique across HCCI data contributors.
X	X	Enrollment Month	MNTH	Character	Calendar Month of active member enrollment.
X	X	Enrollment Year	YR	Character	Calendar Year of active member enrollment.
X	X	Gender	GDR	Character	Member Gender: Male (1), Female (2), Unknown (9).
X		Year of Birth	YBIRTH	Character	Year of Member birth.
	X	Age Band Code	AGE_BAND_CD	Character	A code identifying the age range of the member. Bands: 0-17, 18-24, 25-34, 35-44, 45-54, 55-64, 65-74, 75-84, 85+.
X	X	Relationship Code	REL_CD	Character	Identifies relationship of member to policy holder.
	X	State	STATE	Character	Standard two character indicator of Member's state of residence.
X		State (for Rural or micropolitan CBSAs)	STATE_RURAL	Character	Standard two character indicator of Member's state of residence where CBSA is masked.
	X	Member Zip Code	MBR_ZIP_5_CD	Character	The number assigned by the US Postal Service to a geographic area for the purposes of efficient mail sorting and delivery. *Only zip codes corresponding to populations of greater than 1,350 individuals per 2010 US Census ZCTA file are allowed for use.*
X	X	Member CBSA Code	MBR_CBSA_CD	Character	Geographic Indicator, US Census Core Based Statistical Area. Only "Metro" codes, representing populations of 50,000+, are included.
	X	Member Hospital Referral Region Code	HRR_CD	Integer	Identifies a member's regional market for tertiary medical care based on the Dartmouth Atlas Hospital Referral Regions.
X	X	Business Line	BUS_LINE	Character	Identifies the book of business (Commercial, Medicare Advantage). MA data 2009-2016.
X	X	Product	PROD	Character	Type of benefit plan commonly used by the health care industry to identify the product.
X	X	Consumer Driven Health Plan Flag	CDHP_CD	Character	Identifies a member enrolled in a High Deductible / Consumer Driven Health Plan.
X	X	Funding	FUNDING	Character	Identifies ASO (self funded) versus fully insured. Used for commercial products only.
X	X	Prescription Coverage Flag	RX_CVG_IND	Character	Identifies a member with pharmacy benefits coverage.
X	X	Mental Health Coverage Flag	MH_COV_IND	Character	Identifies members who have mental health benefits as part of their plan coverage.
X	X	Market Segment Code	MKT_SGMNT_CD	Character	Indicates the relative size of the customer based on the number of covered lives.
X	X	Standard Industry Classification Code	SIC	Character	A federally assigned Standard Industry Classification number that identifies companies by industry. Values have been aggregated into 8 broad categories.
X	X	Dual Eligibility Flag (MA only)	DUAL_ELIG_CD	Character	Medicare Advantage Only -- Identifies member's who have dual eligiblility with Medicare and Medicaid.
X	X	End Stage Renal Disease Flag (MA only)	ESRD_STATUS	Character	Medicare Advantage Only -- Patient diagnosed with End Stage Renal Disease (ESRD).
X	X	Hospice Flag (MA only)	HOSPICE_STATUS	Character	Medicare Advantage Only -- Patient placed in Hospice care.
X	X	Institutional Flag (MA only)	INSTITUTE_STATUS	Character	Medicare Advantage Only -- Patient placed in an institutional setting (excludes confinement stays).
X	X	Group ID (encrypted)	Z_GROUP_ID	Integer	Encrypted, system generated identification number assigned to the member according to which customer segment or employer-specific group plan the member is affiliated with. Close equivalent to Group Number.
X	X	Exchange Indicator	EXCH_IND	Character	ACA indicator Commercial Only: Yes/No indicator of whether plan is offered through an HIE. 2014-2016 data only.
X	X	Metallic Level of Plan	METALLIC_LVL	Character	ACA indicator Commercial Only: Coverage level (Platinum, Gold, Silver, Bronze, Catastrophic). 2014-2016 data only.
X	X	Individual Market Flag	INDV_FLAG	Character	Derived flag for purposes of data set filtering. Value of '1' indicates an Individual Market policy.

Two Statistically De-identified Views are available for selection, Data Set #1 and Data Set #2. Fields available under each view are denoted with an 'X'.

DISCLAIMER: The contents of this document are for informative purposes only and could be subject to change. HCCI cannot warrant how fully populated a variable listed below may be.
Note: Two statistically de-identified data views are available for research, Data Set #1 and Data Set #2. To ensure the data remains de-identified, researchers may only request a single data set, Data Set #1 OR Data Set #2, for a research project. The differences between Data Set #1 and #2 are age and geography.

Member Enrollment Fields

Data Set #1	Data Set #2	Field Name	Field ID	Data Type	Description
X	X	Medicare Advantage/Non Commercial Flag	NONCOM_FLAG	Character	Derived flag for purposes of data set filtering. Value of '1' indicates a Medicare Advantage policy. MA data 2009-2016.
X	X	Age over 65 Flag	OVER65_FLAG	Character	Derived flag for purposes of data set filtering. Value of '1' indicates member age of 65+.

HEALTH CARE COST INSTITUTE

This Data Dictionary applies to the HCCI data in the data Enclave for data years 2008-2016.

Two Statistically De-identified Views are available for selection, Data Set #1 and Data Set #2. Fields available under each view are denoted with an 'X'.

DISCLAIMER: The contents of this document are for informative purposes only and could be subject to change. HCCI cannot warrant how fully populated a variable listed below may be.
Note: Two statistically de-identified data views are available for research, Data Set #1 and Data Set #2. To ensure the data remains de-identified, researchers may only request a single data set, Data Set #1 OR Data Set #2, for a research project. The differences between Data Set #1 and #2 are age and geography.

Medical Claim Fields - Inpatient

Data Set #1	Data Set #2	Field Name	Field ID	Data Type	Description
X	X	Patient Identifier (encrypted)	Z_PATID	Integer	Encrypted, unique identifier for all members in data set. PATID is consistent over time and unique across HCCI data contributors.
X	X	Medical Claim ID (encrypted)	Z_CLMID	Integer	Encrypted Claim ID.
X	X	Claim Sequence Code	CLMSEQ	Character	Number assigned in the source system to the service within the claim. Used with E_CLMID.
X	X	Claim Incurred Year	YR	Character	Incurred year of service in format 'YYYY'.
X	X	Claim Incurred Month	MNTH	Character	Incurred month of service in format 'MM'.
X	X	Claim Form Type	CLM_FRM_TYP	Character	Claim form type.
X	X	Type of Bill	TOB	Character	Type Of Bill indicator for facility claims.
X	X	First Service Date	FST_DT	Date	The beginning date for the service, event, or confinement being billed by the provider.
X	X	Last Service Date	LST_DT	Date	The ending date for the service, event, or confinement being billed by the provider.
X	X	Admit Date	FST_ADMTDT	Date	Admission Date for Inpatient confinement.
X	X	Discharge Date	LAST_DISCHDT	Date	Discharge Date for Inpatient confinement.
X	X	Admit ID	Z_ADMIT_ID	Integer	Encrypted, unique identifier for an inpatient confinement. *Only present on Inpatient claims.*
X	X	Admit Source	ADMIT_SRC	Character	Point of origin for admission.
X	X	Admit Type	ADMIT_TYPE	Character	Source of patients admission.
X	X	Admit Record Flag	ADMITS	Integer	A derived column that flags admissions according to the sum of the allowed dollars. Fractional value that when summed across all service lines in an admission will equal (within rounding) -1, 0, or 1, representing negative, zero, or positive dollars respectively. Use in combination with ADMIT_ID for counting Admissions (utilization count). *Only present on Inpatient claims.*
X	X	Length of Stay	LOS	Integer	Length of Stay for Inpatient confinement. Fractional value that when summed across all service lines in an admission will equal the total length of stay. Use in combination with ADMIT_ID for counting total inpatient days. *Only present on Inpatient claims.*
X	X	Major Diagnostic Category	MDC	Varchar	Major Diagnostic Category. *Only present on Inpatient claims.*
X	X	Claim Paid Date	PAID_DT	Date	The date that appears on the check or EFT for claims payment.
X	X	Charge Amount	CHARGE	Numeric	The submitted charges less any non-covered expenses due to: 1. Ineligible charges 2. Ineligible patients or providers 3. Incomplete information. It is used as the baseline for evaluating the effectiveness of network arrangements.
X	X	Net Paid Amount	AMT_NET_PAID	Numeric	The actual amount paid to the provider for the service performed after all deductions and calculations are performed. This does not include the amount paid fee for service on a capitated service. Values may be positive $ amount or zero or negative $ amount or null.
X	X	Coinsurance Amount	COINS	Numeric	The amount (usually calculated as a percent of the provider's submitted charges) the member pays for a specific service as defined in their benefit plan. For example, 20% of the cost of an outpatient physical therapy visit. Values may be positive $ amount or zero or negative $ amount or null.
X	X	Copayment Amount	COPAY	Numeric	The fixed amount the member pays for a specific service as defined in their benefit plan. For example, $10 for an office visit. Values may be positive $ amount or zero or negative $ amount or null.
X	X	Deductible Amount	DEDUCT	Numeric	The amount applied to the member's deductible.
X	X	Calculated Allowed Amount	CALC_ALLWD	Numeric	The sum of the plan payment plus member cost-share. AMT_NET_PAID + COINS + COPAY + DEDUCT.
X	X	Total Member Cost Share	TOT_MEM_CS	Numeric	The sum of COINS + COPAY + DEDUCT.
X	X	Units	UNITS	Number	The number of units of service/procedure.
X	X	Diagnosis 1	DIAG1	Varchar	First level ICD-9 as entered on the claim.
X	X	Diagnosis 2	DIAG2	Varchar	Second level ICD-9 as entered on the claim.

Two Statistically De-identified Views are available for selection, Data Set #1 and Data Set #2. Fields available under each view are denoted with an 'X'.

DISCLAIMER: The contents of this document are for informative purposes only and could be subject to change. HCCI cannot warrant how fully populated a variable listed below may be.
Note: Two statistically de-identified data views are available for research, Data Set #1 and Data Set #2. To ensure the data remains de-identified, researchers may only request a single data set, Data Set #1 OR Data Set #2, for a research project. The differences between Data Set #1 and #2 are age and geography.

Medical Claim Fields - Inpatient

Data Set #1	Data Set #2	Field Name	Field ID	Data Type	Description
X	X	Diagnosis 3	DIAG3	Varchar	Third level ICD-9 as entered on the claim.
X	X	ICD10_CM Diagnosis Code	ICD10_CM1-25	Varchar	First-twenty fifth level ICD-10-CM diagnosis code as entered on the claim (without decimal point). Effective with incurred DATEs of service on/after 10/01/2015.
X	X	Present on Admission Code	POA1-25	Character	Present on Admission code (for diagnosis 1-25). An indicator that differentiates conditions present at time of admission from those conditions that develop during the inpatient admission.
X	X	Diagnosis Related Group	DRG	Varchar	The Diagnosis Related Group (DRG) Code.
X	X	Diagnosis Related Group Type	DRG_TYPE	Varchar	Type of DRG code used in claims calculation. 'MS' or 'CMS'.
X	X	Discharge Status	DSTATUS	Character	Discharge Status Code. Valid for hospital stays only.
X	X	Procedure Code (CPT/HCPCS)	PROC_CD	Varchar	CPT/HCPCS code.
X	X	Procedure Code 1 (ICD-9)	PROC1	Varchar	ICD-9-CM code. Inpatient claims only.
X	X	Procedure Code 2 (ICD-9)	PROC2	Varchar	ICD-9-CM code. Inpatient claims only.
X	X	Procedure Code 3 (ICD-9)	PROC3	Varchar	ICD-9-CM code. Inpatient claims only.
X	X	ICD10_PCS Procedure Code	ICD10_PCS1-25	Varchar	First-twenty fifth level ICD-10-PCS procedure code as entered on the claim. Effective with incurred DATEs of service on/after 10/01/2015.
X	X	Procedure Code Modifier (CPT/HCPCS)	PROCMOD and PROCMOD_2-4	Varchar	First-fourth procedure code modifier. Clarifies or improves the reporting accuracy of the associated procedure code. PROCMOD2-4 2015 data only.
X	X	Revenue Code	RVNU_CD	Varchar	Identifies a specific accommodation, ancillary service or billing calculation for facility claims.
X	X	Place of Service	POS	Varchar	AMA Place of Service code.
X	X	National Provider Identifier (encrypted)	HNPI	Character	National Provider Identifier (NPI) of the health care provider authorized to prescribe medications. NPI is assigned by NPPES/CMS to a qualified health care provider. This number is encrypted to be consistent across HCCI data contributors using a 32-byte algorithm.
X	X	National Provider Identifier (encrypted) backfill flag	HNPI_FILL_FLG	Character	Derived flag indicating whether the HNPI is a native value as received by payer ('0') or has been backfilled by Optum ('1').
X	X	National Provider Identifier (encrypted) of billing entity	HNPI_BE	Character	National Provider Identifier (NPI) of the health care billing entity delivering the service. NPI is assigned by NPPES/CMS to a qualified health care provider. This number is a one-way hash encrypted value consistent across HCCI data contributors using a 32-byte algorithm.
X	X	National Provider Identifier (encrypted) of billing entity backfill flag	HNPI_BE_FILL_FLG	Character	Derived flag indicating whether the HNPI is a native value as received by payer ('0') or has been backfilled by Optum ('1').
X	X	Provider Category	PROVCAT	Character	Provider category code that indicates the specialty of the health care professional.
	X	Provider Zip Code	PROV_ZIP_5_CD	Character	The number assigned by the US Postal Service to a geographic area for the purposes of efficient mail sorting and delivery.
	X	Provider Zip Code backfill flag	PROV_ZIP5_FILL_FLG	Character	Derived flag indicating whether the PROV_ZIP_5 is a native value as received by payer ('0') or has been backfilled by Optum ('1').
X	X	Provider CBSA Code	PROV_CBSA_CD	Character	Core Based Statistical Area code, a geographic entity defined by the US Census Bureau. Only "Metro" codes, representing populations of 50,000+, are included.
X	X	Network Indicator	NTWRK_IND	Character	Indicates whether a claim was paid in or out of network. 2014-2016 data only.
X	X	Primary Coverage Indicator	PRIMARY_COV_IND	Character	Indicates whether a claim was paid primary, secondary, tertiary, etc. 2014-2016 data only.
X	X	HCCI High Level Service Category	HCCI_HL_CAT	Varchar	Derived "High Level" service category.
X	X	HCCI Detailed Service Category	HCCI_DET_CAT	Varchar	Derived detailed service category.
X	X	Individual Market Flag	INDV_FLAG	Character	Derived flag for purposes of data set filtering. Value of '1' indicates an Individual Market policy.
X	X	Medicare Advantage/Non Commercial Flag	NONCOM_FLAG	Character	Derived flag for purposes of data set filtering. Value of '1' indicates a Medicare Advantage policy. MA data 2009-2016.
X	X	Age over 65 Flag	OVER65_FLAG	Character	Derived flag for purposes of data set filtering. Value of '1' indicates member age of 65+.

HEALTH CARE COST INSTITUTE

This Data Dictionary applies to the HCCI data in the data Enclave for data years 2008-2016.

Two Statistically De-identified Views are available for selection, Data Set #1 and Data Set #2. Fields available under each view are denoted with an 'X'.

DISCLAIMER: The contents of this document are for informative purposes only and could be subject to change. HCCI cannot warrant how fully populated a variable listed below may be.
Note: Two statistically de-identified data views are available for research, Data Set #1 and Data Set #2. To ensure the data remains de-identified, researchers may only request a single data set, Data Set #1 OR Data Set #2, for a research project. The differences between Data Set #1 and #2 are age and geography.

Medical Claim Fields - Outpatient

Data Set #1	Data Set #2	Field Name	Field ID	Data Type	Description
X	X	Patient Identifier (encrypted)	Z_PATID	Integer	Encrypted, unique identifier for all members in data set. PATID is consistent over time and unique across HCCI data contributors.
X	X	Medical Claim ID (encrypted)	Z_CLMID	Integer	Encrypted Claim ID.
X	X	Claim Sequence Code	CLMSEQ	Character	Number assigned in the source system to the service within the claim. Used with E_CLMID.
X	X	Claim Incurred Year	YR	Character	Incurred year of service in format 'YYYY'.
X	X	Claim Incurred Month	MNTH	Character	Incurred month of service in format 'MM'.
X	X	Claim Form Type	CLM_FRM_TYP	Character	Claim form type.
X	X	Type of Bill	TOB	Character	Type Of Bill indicator for facility claims.
X	X	First Service Date	FST_DT	Date	The beginning date for the service, event, or confinement being billed by the provider.
X	X	Last Service Date	LST_DT	Date	The ending date for the service, event, or confinement being billed by the provider.
X	X	Claim First Date	CLM_FST_DT	Date	Minimum FST_DT across all lines of a claim for an Outpatient visit.
X	X	Visit ID	Z_VISITID	Integer	Unique identifier for an outpatient visit. *Only present on Outpatient claims.*
X	X	Visit Record Flag	VISITS	Numeric	A derived column that flags visits according to the sum of the allowed dollars. Fractional value that when summed across all service lines in a visit will equal (within rounding) -1, 0, or 1, representing negative, zero, or positive dollars respectively. Use in combination with VISITID for counting Visits (utilization count). *Only present on Outpatient claims.*
X	X	Procedure Record Flag	PROCS	Numeric	A derived column that flags procedures according to the sum of the allowed dollars. Values of -1, 0, or 1, representing negative, zero, or positive dollars, respectively. Used for counting Procedures (utilization count). *Only present on Outpatient or Physician claims.*
X	X	Claim Paid Date	PAID_DT	Date	The date that appears on the check or EFT for claims payment.
X	X	Charge Amount	CHARGE	Numeric	The submitted charges less any non-covered expenses due to: 1. Ineligible charges 2. Ineligible patients or providers 3. Incomplete information. It is used as the baseline for evaluating the effectiveness of network arrangements.
X	X	Net Paid Amount	AMT_NET_PAID	Numeric	The actual amount paid to the provider for the service performed after all deductions and calculations are performed. This does not include the amount paid fee for service on a capitated service. Values may be positive $ amount or zero or negative $ amount or null.
X	X	Coinsurance Amount	COINS	Numeric	The amount (usually calculated as a percent of the provider's submitted charges) the member pays for a specific service as defined in their benefit plan. For example, 20% of the cost of an outpatient physical therapy visit. Values may be positive $ amount or zero or negative $ amount or null.
X	X	Copayment Amount	COPAY	Numeric	The fixed amount the member pays for a specific service as defined in their benefit plan. For example, $10 for an office visit. Values may be positive $ amount or zero or negative $ amount or null.
X	X	Deductible Amount	DEDUCT	Numeric	The amount applied to the member's deductible.
X	X	Calculated Allowed Amount	CALC_ALLWD	Numeric	The sum of the plan payment plus member cost-share. AMT_NET_PAID + COINS + COPAY + DEDUCT.
X	X	Total Member Cost-Share	TOT_MEM_CS	Numeric	The sum of COINS + COPAY + DEDUCT.
X	X	Units	UNITS	Number	The number of units of service/procedure.
X	X	Diagnosis 1	DIAG1	Varchar	First level ICD-9 as entered on the claim.
X	X	Diagnosis 2	DIAG2	Varchar	Second level ICD-9 as entered on the claim.
X	X	Diagnosis 3	DIAG3	Varchar	Third level ICD-9 as entered on the claim.

Two Statistically De-identified Views are available for selection, Data Set #1 and Data Set #2. Fields available under each view are denoted with an 'X'.

DISCLAIMER: The contents of this document are for informative purposes only and could be subject to change. HCCI cannot warrant how fully populated a variable listed below may be.
Note: Two statistically de-identified data views are available for research, Data Set #1 and Data Set #2. To ensure the data remains de-identified, researchers may only request a single data set, Data Set #1 OR Data Set #2, for a research project. The differences between Data Set #1 and #2 are age and geography.

Medical Claim Fields - Outpatient

Data Set #1	Data Set #2	Field Name	Field ID	Data Type	Description
X	X	ICD10_CM Diagnosis Code	ICD10_CM1-25	Varchar	First-twenty fifth level ICD-10-CM diagnosis code as entered on the claim (without decimal point). Effective with incurred DATEs of service on/after 10/01/2015.
X	X	Present on Admission Code	POA1-25	Character	Present on Admission code (for diagnosis 1-25). An indicator that differentiates conditions present at time of admission from those conditions that develop during the inpatient admission.
X	X	Diagnosis Related Group	DRG	Varchar	The Diagnosis Related Group (DRG) Code.
X	X	Diagnosis Related Group Type	DRG_TYPE	Varchar	Type of DRG code used in claims calculation. 'MS' or 'CMS'.
	X	Discharge Status	DSTATUS	Character	Discharge Status Code. Valid for hospital stays only.
X	X	Procedure Code (CPT/HCPCS)	PROC_CD	Varchar	CPT/HCPCS code.
X	X	Procedure Code 1 (ICD-9)	PROC1	Varchar	ICD-9-CM code. Inpatient claims only.
X	X	Procedure Code 2 (ICD-9)	PROC2	Varchar	ICD-9-CM code. Inpatient claims only.
X	X	Procedure Code 3 (ICD-9)	PROC3	Varchar	ICD-9-CM code. Inpatient claims only.
X	X	Procedure Code Modifier (CPT/HCPCS)	PROCMOD and PROCMOD_2-4	Varchar	First-fourth procedure code modifier. Clarifies or improves the reporting accuracy of the associated procedure code. PROCMOD2-4 2015 data only.
X	X	ICD10_PCS Procedure Code	ICD10_PCS1-25	Varchar	First-twenty fifth level ICD-10-PCS procedure code as entered on the claim. Effective with incurred DATEs of service on/after 10/01/2015.
X	X	Revenue Code	RVNU_CD	Varchar	Identifies a specific accommodation, ancillary service or billing calculation for facility claims.
X	X	Place of Service	POS	Varchar	AMA Place of Service code.
X	X	National Provider Identifier (encrypted)	HNPI	Character	National Provider Identifier (NPI) of the health care provider authorized to prescribe medications. NPI is assigned by NPPES/CMS to a qualified health care provider. This number is encrypted to be consistent across HCCI data contributors using a 32-byte algorithm.
X	X	National Provider Identifier (encrypted) backfill flag	HNPI_FILL_FLG	Character	Derived flag indicating whether the HNPI is a native value as received by payer ('0') or has been backfilled by Optum ('1').
X	X	National Provider Identifier (encrypted) of billing entity	HNPI_BE	Character	National Provider Identifier (NPI) of the health care billing entity delivering the service. NPI is assigned by NPPES/CMS to a qualified health care provider. This number is a one-way hash encrypted value consistent across HCCI data contributors using a 32-byte algorithm.
X	X	National Provider Identifier (encrypted) of billing entity backfill flag	HNPI_BE_FILL_FLG	Character	Derived flag indicating whether the HNPI is a native value as received by payer ('0') or has been backfilled by Optum ('1').
X	X	Provider Category	PROVCAT	Character	Provider category code that indicates the specialty of the health care professional.
	X	Provider Zip Code	PROV_ZIP_5_CD	Character	The number assigned by the US Postal Service to a geographic area for the purposes of efficient mail sorting and delivery.
	X	Provider Zip Code backfill flag	PROV_ZIP5_FILL_FLG	Character	Derived flag indicating whether the PROV_ZIP_5 is a native value as received by payer ('0') or has been backfilled by Optum ('1').
X	X	Provider CBSA Code	PROV_CBSA_CD	Character	Core Based Statistical Area code, a geographic entity defined by the US Census Bureau. Only "Metro" codes, representing populations of 50,000+, are included.
X	X	Network Indicator	NTWRK_IND	Character	Indicates whether a claim was paid in or out of network. 2014-2016 data only.
X	X	Primary Coverage Indicator	PRIMARY_COV_IND	Character	Indicates whether a claim was paid primary, secondary, tertiary, etc. 2014-2016 data only.
X	X	HCCI High Level Service Category	HCCI_HL_CAT	Varchar	Derived "High Level" service category.
X	X	HCCI Detailed Service Category	HCCI_DET_CAT	Varchar	Derived detailed service category.
X	X	Individual Market Flag	INDV_FLAG	Character	Derived flag for purposes of data set filtering. Value of '1' indicates a Medicare Advantage policy. MA data 2009-2016.
X	X	Medicare Advantage/Non Commercial Flag	NONCOM_FLAG	Character	Derived flag for purposes of data set filtering. Value of '1' indicates a Medicare Advantage policy.
X	X	Age over 65 Flag	OVER65_FLAG	Character	Derived flag for purposes of data set filtering. Value of '1' indicates member age of 65+.

HEALTH CARE COST INSTITUTE

This Data Dictionary applies to the HCCI data in the data Enclave for data years 2008-2016.

Two Statistically De-identified Views are available for selection, Data Set #1 and Data Set #2. Fields available under each view are denoted with an 'X'.

DISCLAIMER: The contents of this document are for informative purposes only and could be subject to change. HCCI cannot warrant how fully populated a variable listed below may be.
Note: Two statistically de-identified data views are available for research, Data Set #1 and Data Set #2. To ensure the data remains de-identified, researchers may only request a single data set, Data Set #1 OR Data Set #2, for a research project. The differences between Data Set #1 and #2 are age and geography.

Medical Claim Fields - Physician

Data Set #1	Data Set #2	Field Name	Field ID	Data Type	Description
X	X	Patient Identifier (encrypted)	Z_PATID	Integer	Encrypted, unique identifier for all members in data set. PATID is consistent over time and unique across HCCI data contributors.
X	X	Medical Claim ID (encrypted)	Z_CLMID	Integer	Encrypted Claim ID.
X	X	Claim Sequence Code	CLMSEQ	Character	Number assigned in the source system to the service within the claim. Used with E_CLMID.
X	X	Claim Incurred Year	YR	Character	Incurred year of service in format 'YYYY'.
X	X	Claim Incurred Month	MNTH	Character	Incurred month of service in format 'MM'.
X	X	Claim Form Type	CLM_FRM_TYP	Character	Claim form type.
X	X	Type of Bill	TOB	Character	Type Of Bill indicator for facility claims.
X	X	First Service Date	FST_DT	Date	The beginning date for the service, event, or confinement being billed by the provider.
X	X	Last Service Date	LST_DT	Date	The ending date for the service, event, or confinement being billed by the provider.
X	X	Procedure Record Flag	PROCS	Numeric	A derived column that flags procedures according to the sum of the allowed dollars. Values of -1, 0, or 1, representing negative, zero, or positive dollars, respectively. Used for counting Procedures (utilization count). *Only present on Outpatient or Physician claims.*
X	X	Claim Paid Date	PAID_DT	Date	The date that appears on the check or EFT for claims payment.
X	X	Charge Amount	CHARGE	Numeric	The submitted charges less any non-covered expenses due to: 1. Ineligible charges 2. Ineligible patients or providers 3. Incomplete information. It is used as the baseline for evaluating the effectiveness of network arrangements.
X	X	Net Paid Amount	AMT_NET_PAID	Numeric	The actual amount paid to the provider for the service performed after all deductions and calculations are performed. This does not include the amount paid fee for service on a capitated service. Values may be positive $ amount or zero or negative $ amount or null.
X	X	Coinsurance Amount	COINS	Numeric	The amount (usually calculated as a percent of the provider's submitted charges) the member pays for a specific service as defined in their benefit plan. For example, 20% of the cost of an outpatient physical therapy visit. Values may be positive $ amount or zero or negative $ amount or null.
X	X	Copayment Amount	COPAY	Numeric	The fixed amount the member pays for a specific service as defined in their benefit plan. For example, $10 for an office visit. Values may be positive $ amount or zero or negative $ amount or null.
X	X	Deductible Amount	DEDUCT	Numeric	The amount applied to the member's deductible.
X	X	Calculated Allowed Amount	CALC_ALLWD	Numeric	The sum of the plan payment plus member cost-share. AMT_NET_PAID + COINS + COPAY + DEDUCT.
X	X	Total Member Cost-Share	TOT_MEM_CS	Numeric	The sum of COINS + COPAY + DEDUCT.
X	X	Units	UNITS	Number	The number of units of service/procedure.
X	X	Diagnosis 1	DIAG1	Varchar	First level ICD-9 as entered on the claim.
X	X	Diagnosis 2	DIAG2	Varchar	Second level ICD-9 as entered on the claim.
X	X	Diagnosis 3	DIAG3	Varchar	Third level ICD-9 as entered on the claim.
X	X	ICD10_CM Diagnosis Code	ICD10_CM1-25	Varchar	First-twenty fifth level ICD-10-CM diagnosis code as entered on the claim (without decimal point). Effective with incurred DATEs of service on/after 10/01/2015.
X	X	Present on Admission Code	POA1-25	Character	Present on Admission code (for diagnosis 1-25). An indicator that differentiates conditions present at time of admission from those conditions that develop during the inpatient admission.
X	X	Diagnosis Related Group	DRG	Varchar	The Diagnosis Related Group (DRG) Code.
X	X	Diagnosis Related Group Type	DRG_TYPE	Varchar	Type of DRG code used in claims calculation. 'MS' or 'CMS'.
	X	Discharge Status	DSTATUS	Character	Discharge Status Code. Valid for hospital stays only.
X	X	Procedure Code (CPT/HCPCS)	PROC_CD	Varchar	CPT/HCPCS code.

Two Statistically De-identified Views are available for selection, Data Set #1 and Data Set #2. Fields available under each view are denoted with an 'X'.

DISCLAIMER: The contents of this document are for informative purposes only and could be subject to change. HCCI cannot warrant how fully populated a variable listed below may be.
Note: Two statistically de-identified data views are available for research, Data Set #1 and Data Set #2. To ensure the data remains de-identified, researchers may only request a single data set, Data Set #1 OR Data Set #2, for a research project. The differences between Data Set #1 and #2 are age and geography.

Medical Claim Fields - Physician

Data Set #1	Data Set #2	Field Name	Field ID	Data Type	Description
X	X	Procedure Code 1 (ICD-9)	PROC1	Varchar	ICD-9-CM code. Inpatient claims only.
X	X	Procedure Code 2 (ICD-9)	PROC2	Varchar	ICD-9-CM code. Inpatient claims only.
X	X	Procedure Code 3 (ICD-9)	PROC3	Varchar	ICD-9-CM code. Inpatient claims only.
X	X	ICD10_PCS Procedure Code	ICD10_PCS1-25	Varchar	First-twenty fifth level ICD-10-PCS procedure code as entered on the claim. Effective with incurred DATEs of service on/after 10/01/2015.
X	X	Procedure Code Modifier (CPT/HCPCS)	PROCMOD and PROCMOD_2-4	Varchar	First-fourth procedure code modifier. Clarifies or improves the reporting accuracy of the associated procedure code. PROCMOD2-4 2015 data only.
X	X	Revenue Code	RVNU_CD	Varchar	Identifies a specific accommodation, ancillary service or billing calculation for facility claims.
X	X	Place of Service	POS	Varchar	AMA Place of Service code.
X	X	National Provider Identifier (encrypted)	HNPI	Character	National Provider Identifier (NPI) of the health care provider authorized to prescribe medications. NPI is assigned by NPPES/CMS to a qualified health care provider. This number is encrypted to be consistent across HCCI data contributors using a 32-byte algorithm.
X	X	National Provider Identifier (encrypted) backfill flag	HNPI_FILL_FLG	Character	Derived flag indicating whether the HNPI is a native value as received by payer ('0') or has been backfilled by Optum ('1').
X	X	National Provider Identifier (encrypted) of billing entity	HNPI_BE	Character	National Provider Identifier (NPI) of the health care billing entity delivering the service. NPI is assigned by NPPES/CMS to a qualified health care provider. This number is a one-way hash encrypted value consistent across HCCI data contributors using a 32-byte algorithm.
X	X	National Provider Identifier (encrypted) of billing entity backfill flag	HNPI_BE_FILL_FLG	Character	Derived flag indicating whether the HNPI is a native value as received by payer ('0') or has been backfilled by Optum ('1').
X	X	Provider Category	PROVCAT	Character	Provider category code that indicates the specialty of the health care professional.
X	X	Major Physician Specialty	MAJ_SPEC	Varchar	Derived Major Physician Specialty, based on PROVCAT field.
X	X	Primary Care Physician Flag	PCP	Character	Derived field for Primary Care Physician, based on PROVCAT field.
	X	Provider Zip Code	PROV_ZIP_5_CD	Character	The number assigned by the US Postal Service to a geographic area for the purposes of efficient mail sorting and delivery.
	X	Provider Zip Code backfill flag	PROV_ZIP5_FILL_FLG	Character	Derived flag indicating whether the PROV_ZIP_5 is a native value as received by payer ('0') or has been backfilled by Optum ('1').
X	X	Provider CBSA Code	PROV_CBSA_CD	Character	Core Based Statistical Area code, a geographic entity defined by the US Census Bureau. Only "Metro" codes, representing populations of 50,000+, are included.
X	X	Network Indicator	NTWRK_IND	Character	Indicates whether a claim was paid in or out of network. 2014-2016 data only.
X	X	Primary Coverage Indicator	PRIMARY_COV_IND	Character	Indicates whether a claim was paid primary, secondary, tertiary, etc. 2014-2016 data only.
X	X	HCCI High Level Service Category	HCCI_HL_CAT	Varchar	Derived "High Level" service category.
X	X	HCCI Detailed Service Category	HCCI_DET_CAT	Varchar	Derived detailed service category.
X	X	Individual Market Flag	INDV_FLAG	Character	Derived flag for purposes of data set filtering. Value of '1' indicates an Individual Market policy.
X	X	Medicare Advantage/Non Commercial Flag	NONCOM_FLAG	Character	Derived flag for purposes of data set filtering. Value of '1' indicates a Medicare Advantage policy. MA data 2009-2016.
X	X	Age over 65 Flag	OVER65_FLAG	Character	Derived flag for purposes of data set filtering. Value of '1' indicates member age of 65+.

HEALTH CARE COST INSTITUTE

This Data Dictionary applies to the HCCI data in the data Enclave for data years 2008-2016.

Two Statistically De-identified Views are available for selection, Data Set #1 and Data Set #2. Fields available under each view are denoted with an 'X'.

DISCLAIMER: The contents of this document are for informative purposes only and could be subject to change. HCCI cannot warrant how fully populated a variable listed below may be.
Note: Two statistically de-identified data views are available for research, Data Set #1 and Data Set #2. To ensure the data remains de-identified, researchers may only request a single data set, Data Set #1 OR Data Set #2, for a research project. The differences between Data Set #1 and #2 are age and geography.

Pharmacy Claim Fields

Data Set #1	Data Set #2	Field Name	Field ID	Data Type	Description
X	X	Patient Identifier (encrypted)	Z_PATID	Integer	Encrypted, unique identifier for all members in data set. PATID is consistent over time and unique across HCCI data contributors.
X	X	Pharmacy Claim ID (encrypted)	Z_CLMID	Integer	Encrypted Claim ID.
X	X	Claim Incurred Year	YR	Character	Year the prescription was filled by the pharmacy in format 'YYYY'.
X	X	Claim Incurred Month	MNTH	Character	Month the prescription was filled by the pharmacy in format 'MM'.
X	X	Claim Payment Year and Month	YRMNTH_PD	Character	Year and month the prescription claim was paid in format 'YYYYMM'.
X	X	Prescription Fill Date	FILL_DT	Date	Date the prescription was filled by the pharmacy.
X	X	Claim Paid Date	CHK_DT	Date	The date that appears on the check for claims payment.
X	X	Average Wholesale Price	AVGWHLSL	Numeric	The average price at which wholesalers sell drugs to physicians, pharmacies and other customers. 2014 and 2015 data only.
X	X	Net Paid Amount	AMT_NET_PAID	Numeric	The amount the pharmacy is reimbursed. Also referred to as the net amount.
X	X	Copayment Amount	COPAY	Numeric	The fixed amount the member pays for a specific service as defined in their benefit plan. For example, $10 for a prescription fill.
X	X	Coinsurance Amount	COINS	Numeric	The amount (usually calculated as a percent of the provider's submitted charges) the member pays for a specific service as defined in their benefit plan. For example, 20% of the cost of a prescription.
X	X	Deductible Amount	DEDUCT	Numeric	The amount applied to the member's deductible.
X	X	Calculated Allowed Amount	CALC_ALLWD	Numeric	The sum of the plan payment plus member cost-share. AMT_NET_PAID + COINS + COPAY + DEDUCT.
X	X	Total Member Cost-Share	TOT_MEM_CS	Numeric	The sum of COINS + COPAY + DEDUCT.
X	X	Dispensing Fee	DISPFEE	Numeric	Amount the pharmacy charged to fill the prescription.
X	X	Quantity	QUANTITY	Numeric	Quantity of drug dispensed in metric units.
X	X	National Provider Identifier of Prescriber (encrypted)	HNPI	Character	National Provider Identifier (NPI) of the health care provider authorized to prescribe medications. NPI is assigned by NPPES/CMS to a qualified health care provider. This number is encrypted to be consistent across HCCI data contributors using a 32-byte algorithm.
X	X	National Provider Identifier of Prescriber (encrypted) backfill flag	HNPI_FILL_FLG	Character	Derived flag indicating whether the HNPI is a native value as received by payer ('0') or has been backfilled by Optum ('1').
X	X	Dispense as Written Code	DAW	Character	Identifies if a prescription was filled as written or altered by Pharmacy, Physician or Member.
X	X	First Fill Flag	FST_FILL	Character	Indicates if this is the first time a prescription is being filled.
X	X	Number of Days Supplied	DAYS_SUP	Numeric	Estimated day count the drug supply should last.
X	X	Prescription Refill Number	RFL_NBR	Varchar	Indicates if this is the first, second, or subsequent refill for the prescription.
X	X	Prescription Record Flag	SCRIPTS	Numeric	A derived column that flags prescriptions according to the allowed dollars. Values of -1, 0, or 1, representing negative, zero, or positive dollars, respectively. Used for counting Prescriptions (utilization count).
X	X	National Drug Code	NDC	Character	The unique code that identifies a drug product as defined by the National Drug Data File (all drug products regulated by the FDA must use an NDC).
X	X	AHFS Major Therapeutic Class	MAJ_THRPTC_CL	Character	American Hospital Formulary Service (AHFS) "first tier" classification consisting of 31 categories of drugs sharing similar pharmacologic, therapeutic, and/or chemical characteristics, based on the NDC code.
X	X	Multi-Source Indicator	MULTISRC_IND	Character	Identifier of single-source medication versus multiple-source medication
X	X	Generic Drug Flag	GNRC_IND	Character	Identifier of brand medication versus generic.
X	X	Specialty Pharmacy Flag	SPCLT_IND	Character	Indicates if the pharmacy is a specialty pharmacy.
X	X	Mail Order Pharmacy Flag	MAIL_IND	Character	Indicates if the pharmacy is a mail order pharmacy.

Two Statistically De-identified Views are available for selection, Data Set #1 and Data Set #2. Fields available under each view are denoted with an 'X'.

DISCLAIMER: The contents of this document are for informative purposes only and could be subject to change. HCCI cannot warrant how fully populated a variable listed below may be.
Note: Two statistically de-identified data views are available for research, Data Set #1 and Data Set #2. To ensure the data remains de-identified, researchers may only request a single data set, Data Set #1 OR Data Set #2, for a research project. The differences between Data Set #1 and #2 are age and geography.

Pharmacy Claim Fields

Data Set #1	Data Set #2	Field Name	Field ID	Data Type	Description
X	X	Compound Drug Indicator	CMPD_IND	Character	Indicates if the medication dispensed is a compound drug, a medication mixed/adjusted by a pharmacist to achieve a custom strength, form, or ingredient set. 2014-2016 data only.
X	X	Drug on Formulary	FORM_IND	Character	Indicates if the drug being dispensed is on the formulary list or not. 2014-2016 data only.
X	X	HCCI High Level Service Category	HCCI_HL_CAT	Varchar	Derived "High Level" service category (always 'RX').
X	X	HCCI Detailed Product Category	HCCI_DET_CAT	Varchar	Derived detailed product category.
X	X	Individual Market Flag	INDV_FLAG	Character	Derived flag for purposes of data set filtering. Value of '1' indicates an Individual Market policy.
X	X	Medicare Advantage/Non Commercial Flag	NONCOM_FLAG	Character	Derived flag for purposes of data set filtering. Value of '1' indicates a Medicare Advantage policy. MA data 2009-2016.
X	X	Age over 65 Flag	OVER65_FLAG	Character	Derived flag for purposes of data set filtering. Value of '1' indicates member age of 65+.