

## **Missing and Invalid Data** **(Joint Commission Performance Measurement Systems Only)**

### **Introduction**

Missing data refers to data elements, required for calculating a national hospital quality measure, that have no values present for one or more episodes of care (EOC). Invalid data refers to data element values, required for calculating a national hospital quality measure, that fall outside of the range of allowable values defined by the Joint Commission for that data element. Reducing missing and invalid data minimizes the bias to a measure rate, because episodes of care with missing or invalid data cannot be included in the observed measure rate. A measure's observed rate may not accurately reflect the patient population, if the excluded EOC records differ significantly from the EOCs with no missing data that were included in the measure calculation. Tracking the rate of missing and invalid data also assists in identifying potential areas for process improvement in data abstraction, collection, conversion, downloading, etc.

### **Types of Missing and Invalid Data**

Four data elements are used to summarize the amount of missing and invalid data by a health care organization per month per national hospital quality measure. Three of the data elements represent the actual number of patient level records that could not be evaluated for a specific national hospital quality measure due to missing or invalid data.

The fourth data element represents the number of patient-level records that were initially identified based on the *ICD-9-CM Principal Diagnosis Code*.

The data elements used to assess the initial number of patient level records identified and the amount of missing and invalid data are:

- ***ICD Population Size***

The total number of case-level records identified for a health care organization for a specific national hospital quality measure should be based on one or more required general data elements (e.g., specified *ICD-9-CM Principal Diagnosis Code*), prior to application of data integrity filters, exclusions, and/or sampling methodology or the application of the measure set common logic for the specified time period.

*Note: Required for both proportion (rate-based) and continuous variable type measures. Refer to specific Measure Information Form (MIF) to determine type of measure.*

Data elements used to calculate *ICD Population Size* by national hospital quality measure set

<b>National hospital quality measure set</b>	<b>Data elements used to calculate <i>ICD Population Size</i></b>
Acute Myocardial Infarction	<i>ICD-9-CM Principal Diagnosis Code</i>
Pneumonia	<i>ICD-9-CM Principal Diagnosis Code</i> and <i>ICD-9-CM Other Diagnosis Codes</i>
Heart Failure	<i>ICD-9-CM Principal Diagnosis Code</i>
Pregnancy and Related Conditions	<i>ICD-9-CM Principal Diagnosis Code</i> and <i>ICD-9-CM Other Diagnosis Codes</i> (exception: PR-2 Inpatient Neonatal Mortality uses patient age, calculated using <i>Admission Date</i> and <i>Birthdate</i> )
Surgical Infection Prevention	<i>ICD-9-CM Principal Procedure Code</i> and/or <i>ICD-9-CM Other Procedure Code</i>

- ***Number of Cases with Missing or Invalid Population Data***  
**(Measure Category Assignment=A)**

The total number of cases that contain missing or invalid data which prevent assignment to the measure's population (denominator population for proportion [rate-based] measures; number of cases for continuous variable measures), for the health care organization during the specified time period.

*Note: Required for both proportion (rate-based) and continuous variable type measures, except in stratified measures wherein missing or invalid data counts are reported only for the overall rate measure [SIP-1a or SIP-2a or SIP-3a]. Refer to specific Measure Information Form (MIF) to determine type of measure.*

*Note: See Measure Category Assignment in the Data Dictionary for a description of all measure category assignments.*

- ***Number of Cases with Missing or Invalid Numerator Data***  
**(Measure Category Assignment=C)**

The total number of cases that belong, or might belong, to a measure's numerator population, but contain missing or invalid data which prevent assignment to the numerator population, for the health care organization during the specified time period.

*Note: Required for proportion [rate-based] type measures except in stratified measures wherein missing or invalid data counts are reported only for the overall rate measure [SIP-1a or SIP-2a or SIP-3a]. Refer to the specific MIF to determine type of measure.*

- ***Number of Cases with Missing or Invalid Risk Adjustment Data (Risk Adjustment Category Assignment=F)***

The total number of cases that belong to a measure's population, but contain missing or invalid risk adjustment data, preventing the appropriate risk adjustment of the case, for the health care organization during the specified time period.

*Note: Required for risk-adjusted measures. Refer to the specific MIF to determine if risk adjustment is applicable.*

*Note: Access to patient level/case level data is required by the performance measurement system to calculate the missing/invalid data elements.*

*Note: For risk-adjusted measures, before calculating an estimated value to replace missing risk adjustment data, it is necessary to determine the Number of Cases with Missing or Invalid Risk Adjustment Data and to identify those EOC records.*

### **Steps to Calculate National Hospital Quality Measure Missing/Invalid Data Counts**

The missing/invalid counts are determined by examining the values of the data elements used to calculate the national hospital quality measure. The specific Measure Information Form and algorithm are used to determine which data elements are required for calculating the measure. Refer to the Data Dictionary for an explanation of *Measure Category Assignments*.

#### **1. Identify measure specific data elements.**

Review the Measure Information Form and the associated algorithm for each measure. The documentation identifies which data elements are required for the measure population inclusion, the numerator (for rate-based measures only) and/or risk adjustment (for risk adjusted measures only).

#### **2. Calculate counts for each missing data element category.**

- A. Process episode of care records for a health care organization, for the specified time period, through the measure's algorithm.
- B. Utilizing the *Measure Category Assignments*, assigned to each EOC record by the algorithm, aggregate (count) the number of missing and/or invalid data elements by type (i.e., Missing or Invalid Measure Population Data, Missing or Invalid Numerator Data and Missing or Invalid Risk Adjustment Data)

*Note: The aggregation of the missing data elements can be done at the time the EOCs are processed through the algorithms or at the time the Health Care Organization (HCO)-level data are calculated.*

*Note: For stratified measure sets, Missing or Invalid Measure Population Data and Missing or Invalid Numerator Data is expected only for the overall rate and not for the strata. For example, the HCO should report the missing or invalid*

*data cases (measure category assignment “A” or “C”) for the overall rate measures (SIP-1a, SIP-2a, SIP-3a) but not for the measures in the strata (SIP-1b through 1h, SIP-2b through 2h, SIP-3b through 3h). Refer to the specific algorithms, ‘Steps to Calculate Rates and Measurements’ section and the ORYX® Technical Implementation Guide for more information on calculating rates and transmitting such cases.*

### **3. Transmitting missing/invalid data counts to the Joint Commission.**

The three data elements described above are transmitted to the Joint Commission as part of the HCO-level data for a specific health care organization’s measure.

Refer to the ORYX® Technical Implementation Guide’s Data Dictionary – HCO-level Data Elements Section and the Electronic Data Interchange – HCO-level Transaction Set Sections for more information on how to transmit this data.

## **Use of Missing/Invalid Data Counts**

Missing data counts should be monitored by performance measurement systems and health care organizations to identify data quality issues and problem areas. Follow-up is necessary to determine where the missing and invalid data issues originate and how they can be improved. Lower counts of missing and invalid data increase the likelihood that measure rates accurately represent the population being measured.

## **Front-End Filters and Missing General Data Elements**

Processes to identify and correct missing and invalid data are necessary to improve data quality. One method to identify data quality issues prior to analysis is by implementing front-end filters, which evaluate data elements in an EOC level record for allowable values as well as consistency between data element values compared to pre-defined specifications. Although the national hospital quality measure initiative has no specific requirements for front-end filters, there is an expectation that a measurement system has a data quality component in place to identify and follow-up on missing and invalid data. The process must include a mechanism to monitor missing data rates over time and to work with participating health care organizations to improve the data completeness for specific data elements. The Joint Commission will continue to use missing data thresholds for the ORYX transmission data elements to identify measurement systems and health care organizations that have statistically significant levels of missing data.

Measurement systems that have elaborate data filters in place to examine and reject episode of care (EOC) records with missing and/or invalid general data elements need to retain the counts of the number of records that were rejected. For example, front-end data filters that prohibit potential national hospital quality measure EOCs with missing or invalid general data elements from being included in the *Number of Cases with Missing or Invalid Population Data* category assignment. Two options are available to handle and identify EOC records with missing/invalid data for a subset of the general data elements:

Option 1: Use the Missing/Invalid data category assignments described in this section of the manual to report missing data rates for each national hospital quality measure.

Option 2: For systems with restrictive front-end data filters, EOCs with missing or invalid values for a subset of the general data elements (i.e., *Admission Date, Birthdate, Discharge Date, ICD-9-CM Principal Diagnosis Code, and ICD-9-CM Other Diagnosis Codes*) can be excluded from algorithm processing if the following conditions are met:

- During verification, the system must provide a report of the test cases that were excluded from processing due to data issues for the subset of general data elements. The report should include the *Case Identifier*, measure set and the reason the EOC was rejected.
- Following verification implementation, each system must implement a mechanism to identify the number of EOC records that were unresolved and excluded due to general data element data issues reported by HCO, month and specific national hospital quality measure population. This count must be added to the *Number of Cases with Missing or Invalid Population Data* for every measure in the set and will be transmitted with the other HCO-level data described in the *ORYX Technical Implementation Guide*.
- The method to track missing general data elements not mentioned in Option #2, as well as measure specific data elements must be reported using the data elements: *Number of Cases with Missing or Invalid Population Data, Number of Cases with Missing or Invalid Numerator Data, and/or Number of Cases with Missing or Invalid Risk Adjustment Data*.

### Examples for Calculating Missing and Invalid Data Elements

#### Example 1

AMI-5 data to demonstrate the calculation of the missing data counts.

EOC	Required Data Elements	Data Element Needed for	Data Element Values	Data Element Value Valid and Exists?	Missing Data Type
1	<i>ICD-9-CM Principal Diagnosis Code</i>	Denominator	410	N (Not Valid)	Population
	<i>Discharge Status</i>	Denominator	33	N (Not Valid)	Population
	<i>Contraindication to Beta Blockers at Discharge</i>	Denominator	N	Y	NA
	<i>Beta Blocker Prescribed at Discharge</i>	Numerator	Y	Y	NA
	<i>Admission Date</i>	Measure Population	10/11/2004	Y	NA

E O C	Required Data Elements	Data Element Needed for	Data Element Values	Data Element Value Valid and Exists?	Missing Data Type
	<i>Birthdate</i>	Measure Population	09/12/1925	Y	NA
2	<i>ICD-9-CM Principal Diagnosis Code</i>	Denominator	410.41	Y	NA
	<i>Discharge Status</i>	Denominator	03	Y	NA
	<i>Contraindication to Beta Blockers at Discharge</i>	Denominator	N	Y	NA
	<i>Beta Blocker Prescribed at Discharge</i>	Numerator	Y	Y	NA
	<i>Admission Date</i>	Measure Population	10/25/2004	Y	NA
	<i>Birthdate</i>	Measure Population	07/17/1918	Y	NA
3	<i>ICD-9-CM Principal Diagnosis Code</i>	Denominator	410.51	Y	NA
	<i>Discharge Status</i>	Denominator	01	Y	NA
	<i>Contraindication to Beta Blockers at Discharge</i>	Denominator	N	Y	NA
	<i>Beta Blocker Prescribed at Discharge</i>	Numerator	(null)	N (Missing)	Numerator
	<i>Admission Date</i>	Measure Population	10/26/2004	Y	NA
	<i>Birthdate</i>	Measure Population	07/11/1944	Y	NA
4	<i>ICD-9-CM Principal Diagnosis Code</i>	Denominator	410.81	Y	NA
	<i>Discharge Status</i>	Denominator	88	N (Not Valid)	Population
	<i>Contraindication to Beta Blockers at Discharge</i>	Denominator	N	Y	NA
	<i>Beta Blocker Prescribed at Discharge</i>	Numerator	Y	Y	NA
	<i>Admission Date</i>	Measure Population	10/29/2004	Y	NA

EOC	Required Data Elements	Data Element Needed for	Data Element Values	Data Element Value Valid and Exists?	Missing Data Type
	<i>Birthdate</i>	Measure Population	08/01/1916	Y	NA

Determine the number of EOCs with one or more missing or invalid data elements for each missing data category. In this example, the missing data elements *Number of Cases with Missing or Invalid Population Data* and *Number of Cases with Missing or Invalid Numerator Data* are applicable. The data element *Number of Cases with Missing or Invalid Risk-Adjustment Data* is not applicable since AMI-5 is not a risk-adjusted measure.

Values for missing data elements based on the data presented in AMI-5 example:

***Number of Cases with Missing/Invalid Population Data***  
**=2 (Population) <i.e., EOC # 1 and EOC # 4>**

*Note: In this example, the first EOC contains two missing data elements used to determine AMI-5 population inclusion. Only the first is counted towards the Number of Cases with Missing or Invalid Population Data, since each missing data element should not contain duplicate EOCs. Also, a single episode of care can only exist in one of the missing data elements for a single measure.*

***Number of Cases with Missing/Invalid Numerator Data***  
**=1 (Numerator) <i.e., EOC # 3>**

***Number of Cases with Missing/Invalid Risk Adjustment Data***  
 Not applicable since AMI-5 is not a risk-adjusted measure. This data element would be transmitted as null in the data transmission to the Joint Commission.

#### Example 2

AMI-7 data to demonstrate the calculation of the missing data counts.

EOC	Required Data Elements	Data Element Needed for	Data Element Values	Data Element Value Valid and Exists?	Missing Data Type
1	<i>Admission Date</i>	Measure Population	10/11/2004	Y	NA
	<i>Admission Source</i>	Measure Population	1	Y	NA
	<i>Arrival Date</i>	Measure Population	03/31/2004	Y	NA
	<i>Arrival Time</i>	Measure Population	07:21	Y	NA
	<i>Birthdate</i>	Measure Population	01/25/1942	Y	NA

E O C	Required Data Elements	Data Element Needed for	Data Element Values	Data Element Valid and Exists?	Missing Data Type
	<i>ICD-9-CM Principal Diagnosis Code</i>	Measure Population	410.51	Y	NA
	<i>Initial ECG Interpretation</i>	Measure Population	Y	Y	NA
	<i>Thrombolytic Administration</i>	Measure Population	Y	Y	NA
	<i>Thrombolytic Administration Date</i>	Measure Population	03/31/2004	Y	NA
	<i>Thrombolytic Administration Time</i>	Measure Population	10:40	Y	NA
	<i>Transfer From Another ED</i>	Denominator	N	Y	NA
2	<i>Admission Date</i>	Measure Population	10/15/2004	Y	NA
	<i>Admission Source</i>	Measure Population	2	Y	NA
	<i>Arrival Date</i>	Measure Population	04/02/2004	Y	NA
	<i>Arrival Time</i>	Measure Population	12:40	Y	NA
	<i>Birthdate</i>	Measure Population	09/12/1925	Y	NA
	<i>ICD-9-CM Principal Diagnosis Code</i>	Measure Population	410.11	Y	NA
	<i>Initial ECG Interpretation</i>	Measure Population	a	N (Invalid ECG Interpretation)	Population
	<i>Thrombolytic Administration</i>	Measure Population	Y	Y	NA
	<i>Thrombolytic Administration Date</i>	Measure Population	04/03/2004	Y	NA
	<i>Thrombolytic Administration Time</i>	Measure Population	20:02	Y	NA
	<i>Transfer From Another ED</i>	Denominator	N	Y	NA

EOC	Required Data Elements	Data Element Needed for	Data Element Values	Data Element Valid and Exists?	Missing Data Type
3	<i>Admission Date</i>	Measure Population	10/11/2004	Y	NA
	<i>Admission Source</i>	Measure Population	3	Y	NA
	<i>Arrival Date</i>	Measure Population	04/15/2004	Y	NA
	<i>Arrival Time</i>	Measure Population	1015	Y	NA
	<i>Birthdate</i>	Measure Population	05/29/1950	Y	NA
	<i>ICD-9-CM Principal Diagnosis Code</i>	Measure Population	410.61	Y	NA
	<i>Initial ECG Interpretation</i>	Measure Population	Y	Y	NA
	<i>Thrombolytic Administration</i>	Measure Population	Y	Y	NA
	<i>Thrombolytic Administration Date</i>	Measure Population	04/15/2004	Y	NA
	<i>Thrombolytic Administration Time</i>	Measure Population	(null)	N (Missing Thrombolytic Administration Time)	Population
	<i>Transfer From Another ED</i>	Denominator	N	Y	NA

Values for missing data elements based on the data presented in AMI-7 example:

***Number of Cases with Missing/Invalid Population Data***  
=2 (Population) <i.e., EOC # 2 and EOC # 3>

***Number of Cases with Missing/Invalid Numerator Data***

Not applicable for measures reported as continuous variable measures. This data element would be transmitted as null in the data transmission to the Joint Commission.

***Number of Cases with Missing/Invalid Risk Adjustment Data***

Not applicable, since AMI-7 is not a risk adjusted measure. This data element would be transmitted as null in the data transmission to the Joint Commission.

Example 3

AMI-9 data to demonstrate the calculation of missing data counts

EOC	Required Data Elements	Data Element Needed For	Data Element Values	Data Element Value Valid and Exists?	Missing Data Type
1	<i>Admission Date</i>	Measure Population Risk Adjustment	10/29/2004	Y	NA
	<i>Admission Source</i>	Denominator Risk Adjustment	2	Y	NA
	<i>Birthdate</i>	Measure Population Risk Adjustment	10/23/1920	Y	NA
	<i>Discharge Status</i>	Denominator Numerator	01	Y	NA
	<i>ICD-9-CM Other Diagnosis Codes</i>	Risk Adjustment	162.2	Y	NA
	<i>ICD-9-CM Principal Diagnosis Code</i>	Denominator	410.01	Y	NA
	<i>Sex</i>	Risk Adjustment	M	Y	NA
2	<i>Admission Date</i>	Measure Population Risk Adjustment	10/23/2004	Y	NA
	<i>Admission Source</i>	Denominator Risk Adjustment	2	Y	NA
	<i>Birthdate</i>	Measure Population Risk Adjustment	(null)	N ( <i>Birthdate missing</i> )	Risk Adjustment
	<i>Discharge Status</i>	Denominator Numerator	20	Y	NA
	<i>ICD-9-CM Other Diagnosis Codes</i>	Risk Adjustment	162.2	Y	NA
	<i>ICD-9-CM Principal Diagnosis Code</i>	Denominator	410.91	Y	NA
	<i>Sex</i>	Risk Adjustment	M	Y	NA

EOC	Required Data Elements	Data Element Needed For	Data Element Values	Data Element Value Valid and Exists?	Missing Data Type
3	<i>Admission Date</i>	Measure Population Risk Adjustment	10/12/2004	Y	NA
	<i>Admission Source</i>	Denominator Risk Adjustment	1	Y	NA
	<i>Birthdate</i>	Measure Population Risk Adjustment	03/26/1900	Y	NA
	<i>Discharge Status</i>	Denominator	(null)	N (Missing <i>Discharge Status</i> )	Population
	<i>ICD-9-CM Other Diagnosis Codes</i>	Risk Adjustment	174.3	Y	NA
	<i>ICD-9-CM Principal Diagnosis Code</i>	Denominator	410.51	Y	NA
	<i>Sex</i>	Risk Adjustment	F	Y	NA
4	<i>Admission Date</i>	Measure Population Risk Adjustment	10/02/2004	Y	NA
	<i>Admission Source</i>	Denominator Risk Adjustment	Q	N (Invalid <i>Admission Source</i> )	Population Risk Adjustment
	<i>Birthdate</i>	Measure Population Risk Adjustment	01/01/1965	Y	NA
	<i>Discharge Status</i>	Denominator Numerator	01	Y	NA
	<i>ICD-9-CM Other Diagnosis Codes</i>	Risk Adjustment	311.99	N ( <i>ICD-9-CM Other Diagnosis Code Invalid</i> )	NA
	<i>ICD-9-CM Principal Diagnosis Code</i>	Denominator	410.31	Y	NA
	<i>Sex</i>	Risk Adjustment	M	Y	NA

Values for missing data elements based on the data presented in AMI-9 example:

*Number of Cases with Missing/Invalid Population Data*  
=2 (Population) <i.e., EOC #3 and EOC #4>

*Number of Cases with Missing/Invalid Numerator Data*  
=0

*Number of Cases with Missing/Invalid Risk Adjustment Data*  
=2 (Risk Adjustment) <i.e., EOC #2 and EOC #4>



**Note:**

Retain the values of A and B to report the number of episodes of care with missing/invalid data from each measure/HCO

