INTER-RATER RELIABILITY

What It Is. How to Do It. And Why Your Hospital’s Bottom Line Is at Risk Without It

BY
With the ever-increasing demand for healthcare quality data, it is essential that organizations produce data that is complete, valid and reliable.

Inter-Rater Reliability (IRR) assessments can help your facility evaluate the accuracy and integrity of your data abstractions.
IRR assessments are performed on a sample of abstracted cases to measure the degree of agreement among reviewers.

Agreement can be expressed in the form of a score, most commonly Data Element Agreement Rates (DEAR) and Category Assignment Agreement Rates (CAAR), which are recommended by The Joint Commission and CMS for evaluating data reliability and validity.
The **Data Element Agreement Rate**, or **DEAR**, is a one-to-one comparison of consensus between the original abstractor and the re-abstractor’s findings at the data element level, including all clinical and demographic elements.

To calculate the DEAR for each data element:

1. Count the number of times the original abstractor and re-abstractor agreed on the data element value across all paired records.

2. Divide by the total number of paired records.

3. Convert to a percentage and evaluate the score. DEARs of 80% or better are acceptable.

DEAR results should be used to identify data element mismatches and pinpoint education opportunities for abstractors.

It is also important to analyze the DEAR results for trends among mismatches (within a specific data element or for a particular abstractor) to determine if a more focused review is needed to ensure accuracy across all potentially affected charts.
**Example DEAR Calculation**

*(using a partial AMI data element list)*

<table>
<thead>
<tr>
<th>Abstractor</th>
<th>Record ID</th>
<th>LVSD</th>
<th>ACEI at Discharge</th>
<th>ARB at Discharge</th>
<th>Reason for No ACEI and No ARB at Discharge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original</td>
<td>1001</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>IRR</td>
<td>1001</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Original</td>
<td>1002</td>
<td>N</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>IRR</td>
<td>1002</td>
<td>Y</td>
<td>Y</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Original</td>
<td>1003</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>IRR</td>
<td>1003</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
</tr>
</tbody>
</table>

**DEAR**

1. Add Successfully Matched Answer Values (numerator) \((2+2+2+1)= 7\)
2. Add Total Paired Answer Values (denominator) \((3+3+2+2)= 10\)
3. Divide Numerator by Denominator \((7/10)= 70\%\)

*n/a in the table above represents fields disabled due to skip logic.*
The **Category Assignment Agreement Rate**, or **CAAR**, is the score utilized in the CMS Validation Process which affects Annual Payment Update.

CAAR is a one-to-one comparison of agreement between the original abstractor and the re-abstractor’s record-level results using Measure Category Assignments.

MCAs are algorithm outcomes that determine numerator, denominator and exclusion status and are typically expressed as A, B, C, D, E.

In other words, the same numerator and denominator values reported by the original abstractor should be obtained by the second abstractor.
HOW TO CALCULATE YOUR CAAR

To calculate the CAAR, count the number of times the original abstractor and re-abstractor arrived at the same MCA; then, divide by the total number of paired MCAs.

Again, convert to a percentage for evaluation purposes.

A score of 75% is considered acceptable by CMS, while TJC prefers 85% or above.

CAAR results should be used to identify the overall impact of data element mismatches on the measure outcomes.

CAAR mismatches can then be reviewed in conjunction with associated DEAR mismatches to foster abstractor knowledge.

Remember, CAAR results are also the best predictor of CMS validation results.
**Example of CAAR Calculation Within Measure Set:**

<table>
<thead>
<tr>
<th>Abstractor</th>
<th>Record ID</th>
<th>AMI-1</th>
<th>AMI-2</th>
<th>AMI-3</th>
<th>AMI-5</th>
<th>AMI-7a</th>
<th>AMI-8a</th>
<th>AMI-10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original</td>
<td>1000</td>
<td>E</td>
<td>B</td>
<td>E</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>D</td>
</tr>
<tr>
<td>IRR</td>
<td>1000</td>
<td>E</td>
<td>B</td>
<td>E</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>D</td>
</tr>
<tr>
<td>Original</td>
<td>1001</td>
<td>E</td>
<td>D</td>
<td>B</td>
<td>D</td>
<td>B</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>IRR</td>
<td>1001</td>
<td>E</td>
<td>E</td>
<td>B</td>
<td>D</td>
<td>B</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>Original</td>
<td>1002</td>
<td>E</td>
<td>B</td>
<td>D</td>
<td>B</td>
<td>D</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td>IRR</td>
<td>1002</td>
<td>B</td>
<td>B</td>
<td>D</td>
<td>B</td>
<td>D</td>
<td>E</td>
<td>E</td>
</tr>
</tbody>
</table>

**Example of CAAR Calculation Across Measure Sets:**

<table>
<thead>
<tr>
<th>Population</th>
<th>Number of Measures</th>
<th>Number of Records Reabstracted</th>
<th>Number of Paired MCAs (Denominator)</th>
<th>Number of Successful Matches (Numerator)</th>
<th>CAAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMI</td>
<td>7</td>
<td>3</td>
<td>21</td>
<td>19</td>
<td>90.48%</td>
</tr>
<tr>
<td>HF</td>
<td>3</td>
<td>3</td>
<td>9</td>
<td>9</td>
<td>100%</td>
</tr>
<tr>
<td>PN</td>
<td>3</td>
<td>3</td>
<td>9</td>
<td>8</td>
<td>88.89%</td>
</tr>
<tr>
<td>SCIP</td>
<td>9</td>
<td>3</td>
<td>27</td>
<td>25</td>
<td>92.59%</td>
</tr>
<tr>
<td>Overall</td>
<td>12</td>
<td>66</td>
<td>61</td>
<td>92.42%</td>
<td></td>
</tr>
</tbody>
</table>

**Overall CAAR =**

1. Add Successfully Matched MCAs (numerator) \((19+9+8+25) = 61\)
2. Add Total Paired MCAs (denominator) \((21+9+9+27) = 66\)
3. Divide Numerator by Denominator \((61/66) = 92.42\%\)
Incorporating Inter-Rater Reliability into your routine can reduce data abstraction errors by identifying the need for abstractor education or re-education and give you confidence that your data is not only valid, but reliable.

**When to Use Inter-Rater Reliability**

1. After a specifications manual update
2. For new abstractors
3. When new measure sets are introduced
4. During performance evaluations
RULES OF ENGAGEMENT

1. The IRR sample should be randomly selected from each population using the entire list of cases, not just those with measure failures.

2. Each case should be independently re-abstracted by someone other than the original abstractor.

3. The IRR abstractor then inputs and compares the answer values for each Data Element and the Measure Category Assignments to identify any mismatches.

4. The results are reviewed/discussed with the original abstractor and case is updated with all necessary corrections prior to submission deadlines.

5. If the original and IRR abstractor are unable to reach consensus, we recommend submitting questions to QualityNet for clarification.

6. Lessons learned from mismatches should be applied to all future abstractions.

7. Results should be analyzed for patterns of mismatches to identify the need for additional IRR Reviews and/or targeted education for staff.
While conducting IRR in house is a good practice, it is not always 100% accurate.

Often abstractors correct for physician documentation idiosyncrasies or misinterpret Core Measures guidelines.

American Data Network can provide an unbiased eye to help you ensure your abstractions are accurate.

To get your free Core Measures Data Abstraction Service quote, call Sherry Bird at (501) 225-5533 or click the button below to fill a short form and we’ll contact you at a time that’s convenient.

About American Data Network Core Measures Data Abstraction Service

Our Core Measures Abstraction Service can help your hospital meet the data collection and reporting requirements of The Joint Commission and Centers for Medicare & Medicaid Services. We will work directly with your facility to provide a solution that fits your needs – whether it’s on site, off site, on call, or partial outsourcing. Plus, it is not necessary to use ADN’s data collection tool; our experienced abstraction specialists will work with whatever Core Measures vendor you use.

As a vendor since the inception of Core Measures, ADN has developed a keen understanding of the measure specifications, transmission processes, and improvement initiatives associated with data collection and analytics. In addition, ADN can train your abstractors on the changes to the measure guidelines and conduct follow-up Inter Rater Reliability assessments to ensure their understanding.

Our data abstraction services allow your hospital to reallocate scarce clinical resources to performance improvement, utilization review and case management. Or, use ADN personnel to complement your existing data abstraction staff to provide coverage for employees on temporary leave or to serve as a safety net for abstractor shortages or unplanned employee departures.