

Reliability, Encounter Level Testing and Measure Performance Score Results for the Addressing Social Needs (ASN) electronic Clinical Quality Measure (eCQM) Merit-based Incentive Payment System (MUC2024-072)

Please note due to limitations within the electronic data ecosystem, reliability, encounter level testing, and measure performance score results were not available by the May 2024 Measures Under Consideration (MUC) submission deadline. As noted within the MERIT submission, testing results would be available by August 2024. Therefore, results are included within this attachment for the 2024 MERIT Data Template relevant sections below. We request reviewers reference this attachment for final testing results.

Measure Authoring Tool (MAT) Number

QDM logic for the ASN eCQM for MIPs setting has been drafted. Four domains have been fully tested with 100% coverage and 100% pass in MADiE. The housing domain logic is complete but will be refined further and finalized prior to the PRMR. Due to the complexities of the logic, the measure was split into four QDM accounts, as having all 460+ test cases in one QDM causes technical glitches within the MADiE tool. Therefore, the measure purposely has no CMS ID yet until the eCQM specifications have been fully tested.

Data Source

One dataset was used for testing:

- **Dataset A** used calendar year 2023 with 12 hospitals with 13,989 providers seeing 1,553,331 patients.

As explained in [Patient/Encounter Level \(Data Element Level\) Testing section](#), coded screening data were infrequent at this test site and coded intervention data were not being captured. Prioritization of social needs data elements varied across datasets. Results show the housing domain was most frequently screened of the 4 domains.

Reliability: Measure Score Level (Accountable Entity Level) Testing

The measure is specified at the clinician or clinician group level. Reliability testing was completed using a Signal-to-Noise calculation at the provider ID level.

13,989 accountable entities (providers) included in results. Median in Table 1.

Table 1. Measure Score Level Reliability Analysis: Signal-to-Noise, 13,989 providers

Reliability	Housing	Food	Transportation	Utilities
Mean (SD)	0.850 (0.131)	0.851 (0.131)	0.854 (0.129)	0.829 (0.148)
Median	0.868	0.871	0.867	0.845

Min-Max	(0.687-1.000)	(0.693-1.000)	(0.686 -1.000)	(0.645-1.000)
----------------	---------------	---------------	----------------	---------------

Interpretation of Measure Score Reliability Signal-to Noise results:

Overall results show high reliability for all measured entities.

Patient/Encounter Level (Data Element Level) Testing

Patient-level testing of the individual data elements in the final performance measure (i.e., measure of agreement such as kappa or correlation coefficient) were not completed.

There was limited social needs screening captured in structured fields in 2023, shown in Table 2. Our testing partners noted that their hospitals began systematically screening, collecting, and documenting social needs data on more patients beginning in 2024. There were also almost no follow-up interventions being documented in structured and encoded fields in the 2023 testing data.

Table 2. Percentage of patients with indication of a social needs screen, by domain

% of encounters screened or had z-code indicating social need			
Housing	Food	Transportation	Utilities
4.7%	4.7%	1.9%	4.7%

Data element validity requires high rates of data capture and low rates of missing data. Although screening patients for social needs is standard practice according to data from the 2022 American Hospital Association (AHA) Information Technology (IT) survey, our tested entities are not capturing screening data in encoded terminology as required by this measure or even in structure fields (Richwine C, Meklir S. Hospitals' collection and use of data to address social needs and social determinants of health. *Health Serv Res.* 2024; 1-7. doi:[10.1111/1475-6773.14341](https://doi.org/10.1111/1475-6773.14341)). As an eQOM this measure requires that data are captured in structured fields and using USCDI-aligned standard terminology. To support testing in the absence of aligned terminology with testing partners, we relied on exact match data approximation from structured fields.

Measure Performance

ASN eQOM enables a summary score of quality for each domain. 13,989 accountable entities (providers) seeing 1,553,331 patients were included in the analysis of the distribution of performance scores.

Measure performance scores are the domain-specific summary scores, calculated as the (numbers of encounters screened negative + number of encounters positive with a follow-up) divided by all encounters. Due to the lack of standardized intervention data, summary scores are skewed towards those screened negative divided by all encounters.

Mean, standard deviation, median, minimum, and maximum are included in Table 3 below. 10th and 90th percentiles are not included due to few accountable entities this would not provide more detailed

information. Due to the low number of accountable entities, this would not provide more detailed information.

We were unable to test the impact of social risk factors measure scores.

Table 3. Measure Performance: Proportion

Summary scores	Housing	Food	Transportation	Utilities
Mean (SD*)	0.0046 (0.0337)	0.0046 (0.0334)	0.0053 (0.0374)	0.0023 (0.0232)
Median	0	0	0	0
Min-Max	(0-1)	(0-1)	(0-1)	(0-1)

Interpretation of Measure Performance results:

Meaningful measure performance data requires robust screening and intervention data. Other results are based on minimal data; thus, it is difficult to derive meaningful comparisons between measure scores (topics in data availability are detailed above). Even with limited data, results are presented and demonstrate outcomes for hospitals or providers that do not have consistent processes for documenting social needs screening and interventions. Social needs screening and intervention may be occurring, but the data documentation is incomplete. These findings support the intent of the measure to drive toward interoperable and discrete terminology to support health equity. Data capture uptake continues to expand at measured entities and can be influenced by CMS and other stakeholders.

Measure Performance, Narrative Measure Rates

This measure output uses narrative measure scores for stakeholders to better understand where improvement opportunities lie. Output shown below, Table 4. Each row is mutually exclusive; each column is a domain, and the columns will total to the number of patients, being 1,553,331.

Table 4. Narrative Measure Rates, Dataset A (1,553,331 patients)

Narrative Food, Transportation, Utilities	Food	Transportation	Utilities	Narrative Housing	Housing
Encounters "No Documented Assessment for Disparity Related Need"	1,480,157 (95.3%)	1,479,995 (95.3%)	1,524,424 (98.1%)	Encounters "No documented assessment for BOTH housing instability AND homelessness Disparity Related Needs"	1,479,983 (95.3%)
Encounters "Declined Assessment for a Disparity Documented"	6,800 (0.4%)	5,024 (0.3%)	211 (0%)	Encounters "Declined Assessment for a Disparity Documented"	168 (0%)

Encounters “Assessed, Disparity Related Need identified, but No Intervention Documented”	6,668 (0.4%)	2,113 (0.1%)	2,599 (0.2%)	Encounters “Assessed, Disparity Related Need identified of EITHER housing instability or homelessness, but No intervention Documented”	15,390 (1%)
Encounters “Assessed and Disparity Related Need Identified, and Intervention Documented”	10 (0%)	3 (0%)	2 (0%)	Encounters “Assessed and Disparity Related Need Identified of EITHER housing instability or homelessness and Intervention Documented”	9 (0%)
Encounters “Documented assessment for Disparity Related Need, but none Identified”	59,696 (3.8%)	66,196 (4.3%)	26,094 (1.7%)	Encounters “Documented assessment for Disparity Related Need, but none Identified for BOTH housing instability and homelessness”	57,781 (3.7%)

Interpretation Measure Performance, Narrative Measure Rates:

Likely because of the absence of data to support screening and intervention activities, the highest percent is in “No documented assessment for disparity related need.” Due to no intervention data stored in structured and encoded fields for either dataset, Encounters Assessed and Disparity Related Need Identified, and Intervention Documented are zero. Drivers, including this measure, are moving the ecosystem toward standardized processes supported by interoperable terminology. As these drivers support best practice, the percent of encounters with measurable findings for assessment and intervention on identified needs will increase, aligned with the mission of the measure.

TIN/NPI data were not available to complete testing of the exact measure specifications. This was approximated using provider ID.