



## Measure Information

This document contains the information submitted by measure developers/stewards, but is organized according to NQF's measure evaluation criteria and process. The item numbers refer to those in the submission form but may be in a slightly different order here. In general, the item numbers also reference the related criteria (e.g., item 1b.1 relates to sub criterion 1b).

### Brief Measure Information

**NQF #:** 0514

**Corresponding Measures:**

**De.2. Measure Title:** MRI Lumbar Spine for Low Back Pain

**Co.1.1. Measure Steward:** Centers for Medicare & Medicaid Services

**De.3. Brief Description of Measure:** This measure evaluates the percentage of magnetic resonance imaging (MRI) of the lumbar spine studies for low back pain performed in the outpatient setting where conservative therapy was not attempted prior to the MRI. Antecedent conservative therapy may include claim(s) for physical therapy in the 60 days preceding the lumbar spine MRI, claim(s) for chiropractic evaluation and manipulative treatment in the 60 days preceding the lumbar spine MRI, or claim(s) for evaluation and management at least 28 days but no later than 60 days preceding the lumbar spine MRI. The measure is calculated based on a one-year window of Medicare claims data. The measure has been publicly reported, annually, by the measure steward, the Centers for Medicare & Medicaid Services (CMS), since 2010, as a component of its Hospital Outpatient Quality Reporting (HOQR) Program.

**1b.1. Developer Rationale:** This measure will reduce overuse of imaging for uncomplicated low back pain without prior attempts at antecedent conservative therapy, as overuse in this population can result in detection of incidental findings and reflect poor care coordination. The measure score will guide patient selection of providers, assess quality, and inform quality improvement.

**S.4. Numerator Statement:** MRI of the lumbar spine studies with a diagnosis of low back pain (from the denominator) without the patient having claims-based evidence of prior antecedent conservative therapy.

**S.6. Denominator Statement:** The number of MRI of the lumbar spine studies with a diagnosis of low back pain on the imaging claim performed in a hospital outpatient department on Medicare FFS beneficiaries within a 12-month time window.

**S.8. Denominator Exclusions:** Below, in Section S.11 we provide a detailed list of denominator exclusion conditions. Denominator exclusions are consistent with current guidelines, evidence in literature, and guidance from the measure TEP.

**De.1. Measure Type:** Process

**S.17. Data Source:** Claims

**S.20. Level of Analysis:** Facility, Population : Regional and State

**IF Endorsement Maintenance – Original Endorsement Date:** Oct 28, 2008 **Most Recent Endorsement Date:** Oct 28, 2008

**IF this measure is included in a composite, NQF Composite#/title:**

**IF this measure is paired/grouped, NQF#/title:**

**De.4. IF PAIRED/GROUPED, what is the reason this measure must be reported with other measures to appropriately interpret results?** Not applicable; this is not a paired or grouped measure.

### 1. Evidence, Performance Gap, Priority – Importance to Measure and Report

Extent to which the specific measure focus is evidence-based, important to making significant gains in healthcare quality, and improving health outcomes for a specific high-priority (high-impact) aspect of healthcare where there is variation in or overall less-than-optimal performance. **Measures must be judged to meet all sub criteria to pass this criterion and be evaluated against the remaining criteria.**

**1a. Evidence to Support the Measure Focus – See attached Evidence Submission Form**

NQF\_0514\_MeasureEvidenceForm.docx

**1a.1 For Maintenance of Endorsement: Is there new evidence about the measure since the last update/submission?**

Do not remove any existing information. If there have been any changes to evidence, the Committee will consider the new evidence. Please use the most current version of the evidence attachment (v7.1). Please use red font to indicate updated evidence.

**1b. Performance Gap**

Demonstration of quality problems and opportunity for improvement, i.e., data demonstrating:

- considerable variation, or overall less-than-optimal performance, in the quality of care across providers; and/or
- Disparities in care across population groups.

**1b.1. Briefly explain the rationale for this measure** (e.g., how the measure will improve the quality of care, the benefits or improvements in quality envisioned by use of this measure)

*If a COMPOSITE (e.g., combination of component measure scores, all-or-none, any-or-none), SKIP this question and answer the composite questions.*

This measure will reduce overuse of imaging for uncomplicated low back pain without prior attempts at antecedent conservative therapy, as overuse in this population can result in detection of incidental findings and reflect poor care coordination. The measure score will guide patient selection of providers, assess quality, and inform quality improvement.

**1b.2. Provide performance scores on the measure as specified (current and over time) at the specified level of analysis.** *(This is required for maintenance of endorsement. Include mean, std dev, min, max, interquartile range, scores by decile. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities include.) This information also will be used to address the sub-criterion on improvement (4b1) under Usability and Use.*

Analysis of Medicare fee-for-service (FFS) claims data indicates variation in the use of inappropriate MRI lumbar spine studies. For the period from July 2014 to June 2015, performance rates ranged from 14.9 percent to 64.8 percent, with a weighted mean of 39.5 percent.

The data presented below represent information for the 1,128 facilities whose denominator counts met minimum case count requirements for all years included in the table.

Further details on the descriptive statistics for longitudinal facility performance are included below:

|                                       | 2011*             | 2012* | 2013*       | 2014**            | 2015**      | 2016**            |
|---------------------------------------|-------------------|-------|-------------|-------------------|-------------|-------------------|
| Measurement Period                    | Jan 2009–Dec 2009 |       |             | Jan 2010–Dec 2010 |             | Jan 2011–Dec 2011 |
|                                       | Jul 2013–Jun 2014 |       |             | Jul 2014–Jun 2015 |             | Jul 2012–Jun 2013 |
| Facilities                            | 1,128             | 1,128 | 1,128       | 1,128             | 1,128       | 1,128             |
| Minimum Value                         | 17.9%             | 12.3% | 17.1%       | 17.6%             | 21.8%       | 14.9%             |
| 5th Percentile                        | 23.4%             | 26.7% | 27.0%       | 28.0%             | 30.4%       | 29.1%             |
| 25th Percentile                       | 28.4%             | 32.0% | 32.1%       | 33.1%             | 35.9%       | 35.3%             |
| Median                                | 31.9%             | 35.8% | 35.9%       | 36.8%             | 39.8%       | 39.0%             |
| 75th Percentile                       | 35.9%             | 40.1% | 39.8%       | 41.0%             | 44.4%       | 43.5%             |
| 95th Percentile                       | 43.5%             | 48.6% | 48.5%       | 48.0%             | 51.8%       | 50.6%             |
| Maximum Value                         | 63.5%             | 69.1% | 67.6%       | 67.7%             | 72.5%       | 64.8%             |
| Mean Performance (Standard Deviation) | 32.5% (6.2)       |       | 36.4% (6.8) |                   | 36.5% (6.6) | 37.2% (6.2)       |
|                                       | 39.5% (6.6)       |       |             |                   |             | 40.3% (6.6)       |

\*The measurement period for HOQR data reported from 2011 through 2013 ran from January to December.

\*\*Beginning with 2014 public reporting, the measurement period for HOQR was adjusted to run from July to June; consequently, data are not reported for January through June 2012.

The intentions for reporting this measure is to identify facilities with significant outlying performance and to reduce variation. As shown in the table above, significant outlying performance persists among wide variation, indicating there are facilities for which there is a notable rate of overuse.

**1b.3. If no or limited performance data on the measure as specified is reported in 1b2, then provide a summary of data from the literature that indicates opportunity for improvement or overall less than optimal performance on the specific focus of**

**measurement.**

Data have been included in Section 1b.2; these data represent national performance over time, from 2009 to 2015.

**1b.4. Provide disparities data from the measure as specified (current and over time) by population group, e.g., by race/ethnicity, gender, age, insurance status, socioeconomic status, and/or disability.** *(This is required for maintenance of endorsement. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included.) For measures that show high levels of performance, i.e., “topped out”, disparities data may demonstrate an opportunity for improvement/gap in care for certain sub-populations. This information also will be used to address the sub-criterion on improvement (4b1) under Usability and Use.*

Using 2013 performance data, we evaluated the effect of patient and facility characteristics on the likelihood of each beneficiary having an inappropriate MRI lumbar spine study. Using a logistic regression model, we assessed the relationship between patient and facility characteristics for the 207,573 MRI lumbar spine studies performed in 2013 and found that beneficiary age, gender, and race, as well as facility characteristics, had a significant association with the rate of inappropriate MRI lumbar spine studies.

The regression model indicates that patient race/ethnicity is associated with inappropriate imaging. Asians were less likely to undergo inappropriate imaging compared to White beneficiaries (OR 0.866,  $p=0.020$ ). There was no statistical differences for other racial groups compared to White beneficiaries.

The regression model also shows that gender is associated with inappropriate imaging; women were less likely to undergo inappropriate imaging compared to men (OR 0.848,  $p=0.000$ ).

Patient age also had a statistically significant association with imaging use. When looking at Medicare FFS data, comparing to beneficiaries aged 60 to 69, beneficiaries aged 18 to 29 (OR 0.840,  $p=0.003$ ), 40 to 49 (OR 0.898,  $p=0.000$ ), 50 to 59 (OR 0.894,  $p=0.000$ ), 70 to 79 (OR 0.874,  $p=0.000$ ), and 80 to 89 (OR 0.827,  $p=0.000$ ) were statistically less likely to receive an MRI lumbar spine study without appropriate antecedent conservative therapy. There was no statistical difference in the likelihood that a patient received an inappropriate MRI lumbar spine study for patients aged 30 to 39 or patients aged 90+ compared to beneficiaries aged 60 to 69.

Facility characteristics were also associated with rates of inappropriate imaging. When compared to facilities with fewer than 50 beds (a proxy for facility size), facilities with 101 to 250 beds were less likely to perform inappropriate MRI lumbar spine studies (OR 0.933,  $p=0.001$ ). Similarly, a facility's urbanicity impacted a beneficiary's likelihood of having an inappropriate MRI lumbar spine study – urban facilities were less likely than rural facilities, likely caring for rural beneficiaries, to perform inappropriate MRI lumbar spine studies (OR 0.967,  $p=0.008$ ). Finally, major-teaching facilities were less likely to perform inappropriate MRI lumbar spine studies (OR 0.890,  $p=0.000$ ) compared to non-teaching facilities.

While the regression model identified subpopulations of patients and facilities for which there are statistically significant differences in the rate of inappropriate MRI lumbar spine studies, these disparities do not indicate a need for adjustment of the measure specifications. Adjusting for these differences would mask underlying differences in quality of care. As this is a process measure, there should be no difference in the standard of care for these patients; we believe these statistically significant differences are driven by variation in provider practice. Consequently, we do not believe risk adjustment or stratification is necessary or appropriate for this measure.

**1b.5. If no or limited data on disparities from the measure as specified is reported in 1b.4, then provide a summary of data from the literature that addresses disparities in care on the specific focus of measurement. Include citations. Not necessary if performance data provided in 1b.4**

The following literature excerpts highlight disparities associated with overuse of imaging for low back pain between genders, racial groups, and age bands.

Graves et al. (2012) assessed the association of demographic, job-related, psychosocial, and clinical factors with the use of MRI within six weeks from injury among workers' compensation claimants with acute low back pain for 1,830 workers. A total of 362 (19.8 percent) received an early MRI. Results of a multivariable regression demonstrated that male workers were 43 percent more likely to receive an early MRI than were female workers.

Mathias et al. (2012) used data from the HOQR Program to assess consistency in performance across measures, focusing on whether higher imaging use could be associated with certain hospital characteristics. To do so, the study team examined associations

between hospital characteristics and higher use of imaging, drawing on 2008 HOQR data linked with data from the 2009 American Hospital Association survey. Mathias and his team found that use of imaging varied widely and was weakly correlated across most measures. Of note, hospitals with low volume (<25th percentile) were more likely to report higher imaging use than were hospitals of medium volume (25th to 75th percentile). Of particular interest, rural hospitals were more likely to report highest-decile use of lumbar spine MRI, in addition to several other measures. The study authors concluded that there are significant variations in use of imaging, with some hospitals reporting exceptionally high use.

Pham et al. (2009) analyzed Medicare claims from 2000 through 2002 and from 2004 through 2006, for 35,039 fee-for-service (FFS) Medicare beneficiaries with acute LBP. The research team found that minority beneficiaries received less rapid and less advanced imaging than did white beneficiaries. Beneficiaries covered by Medicaid also received less rapid and less advanced imaging when compared to other patients (22.7 percent versus 29.7 percent [ $p<.001$ ] for imaging within 28 days, and 7.3 percent versus 11.0 percent [ $p<.001$ ] for CT/MRI).

Friedman et al. (2010), using data from the National Hospital Ambulatory Medical Care Survey (NHAMCS), highlighted the frequency of emergency department (ED) visits for the treatment of low back pain and identified the diagnostic and therapeutic strategies used by physicians in a large sample representative of all ED visits throughout the United States. Results showed that, of all patients with LBP, 9.6 percent (95 percent CI: 7.2, 12.6) had a CT or MRI in 2006, compared with 3.2 percent (95 percent CI: 2.0, 5.1) in 2002 ( $p<0.01$ ). Age and type of insurance were associated with advanced imaging, while geographic region was not.

Despite evidence identified in the literature indicating disparities in care for certain facility types and patient populations who present with LBP, these disparities do not indicate a need for adjustment of the measure specifications. Adjusting for these differences would mask underlying differences in quality of care. As this is a process measure that is not currently risk-adjusted, there should be no difference in the standard of care for these patients; we believe these statistically significant differences are driven by variation in provider performance. Consequently, we do not believe risk adjustment or stratification is necessary or appropriate for this measure.

#### REFERENCES

- 1.) Friedman BW, Chilstrom M, Bijur PE, Gallagher EJ. Diagnostic testing and treatment of low back pain in United States emergency departments: a national perspective. *Spine*. 2010; 35(24):E1406-11.
- 2.) Graves JM, Fulton-Kehoe, D, Martin DP, et al. Factors associated with early magnetic resonance imaging utilization for acute occupational low back pain: a population-based study from Washington State Workers' Compensation. *Spine*. 2012; 37(19): 1708-1718.
- 3.) Mathias JS, Feinglass J, Baker DW. Variations in US hospital performance on imaging-use measures. *Med Care*. 2012;50(9):808-14.
- 4.) Pham HH, Landon BE, Reschovsky JD, et al. Rapidity and modality of imaging for acute low back pain in elderly patients. *Archives of Internal Medicine*. 2009; 169(10):972-981.

## 2. Reliability and Validity—Scientific Acceptability of Measure Properties

Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. ***Measures must be judged to meet the sub criteria for both reliability and validity to pass this criterion and be evaluated against the remaining criteria.***

**2a.1. Specifications** The measure is well defined and precisely specified so it can be implemented consistently within and across organizations and allows for comparability. eMeasures should be specified in the Health Quality Measures Format (HQMF) and the Quality Data Model (QDM).

**De.5. Subject/Topic Area** (check all the areas that apply):

Musculoskeletal, Musculoskeletal : Low Back Pain

**De.6. Non-Condition Specific**(check all the areas that apply):

Care Coordination, Safety, Safety : Overuse

**De.7. Target Population Category** (Check all the populations for which the measure is specified and tested if any):

Elderly

**S.1. Measure-specific Web Page** (Provide a URL link to a web page specific for this measure that contains current detailed specifications including code lists, risk model details, and supplemental materials. Do not enter a URL linking to a home page or to general information.)

<https://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic%2FPage%2FQnetTier2&cid=1228695266120>

**S.2a. If this is an eMeasure**, HQMF specifications must be attached. Attach the zipped output from the eMeasure authoring tool (MAT) - if the MAT was not used, contact staff. (Use the specification fields in this online form for the plain-language description of the specifications)

**This is not an eMeasure Attachment:**

**S.2b. Data Dictionary, Code Table, or Value Sets** (and risk model codes and coefficients when applicable) must be attached. (Excel or csv file in the suggested format preferred - if not, contact staff)

**Attachment Attachment:** [NQF\\_0514\\_MeasureCodeList.xlsx](#)

**S.2c.** Is this an instrument-based measure (i.e., data collected via instruments, surveys, tools, questionnaires, scales, etc.)? Attach copy of instrument if available.

**Attachment:**

**S.2d.** Is this an instrument-based measure (i.e., data collected via instruments, surveys, tools, questionnaires, scales, etc.)? Attach copy of instrument if available.

**S.3.1. For maintenance of endorsement:** Are there changes to the specifications since the last updates/submission. If yes, update the specifications for S1-2 and S4-22 and explain reasons for the changes in S3.2.

**S.3.2. For maintenance of endorsement,** please briefly describe any important changes to the measure specifications since last measure update and explain the reasons.

As part of the annual measures maintenance and review process, several modifications were made to the list of excluded procedures and diagnoses, as well as to the look-back periods for certain diagnoses. In 2012, we added lumbar-spine surgery within the 90 days prior to the imaging procedure to the list of excluded procedures; prior lumbar-spine surgery is an appropriate reason for performing an MRI lumbar-spine study based on evidence in the literature (indicated in American College of Radiology Practice Guideline for the Performance of MRI of the Adult Spine). Look-back periods for each of the excluded clinical conditions were also added to the measure's specifications in 2012 (as identified by Lewin clinical staff and evaluated by the Technical Expert Panel [TEP]). Finally, nine additional exclusion categories (congenital spine and spinal-cord malformations, inflammatory and autoimmune disorders, infectious conditions, spinal vascular malformations and/or occult subarachnoid hemorrhage causes, spinal-cord infarction, effects of radiation, spinal abnormalities associated with scoliosis, syringohydromyelia, and postoperative fluid collections and soft-tissue changes) based on the 2012 update to the American College of Radiology Practice Guideline for the Performance of MRI of the Adult Spine and are aligned with the 2015 update to the ACR guideline. When presented to the contractor's TEP, the TEP supported these additions to the measure specifications.

There have been no changes in the measure specifications since the last measure update; however several measure refinements have been made in the past due to literature identified in CMS contractor's annual review of measure's evidence base and feedback from stakeholders.

Since the measure was initially endorsed in 2008, changes to the specifications include updates to the trauma exclusion and to the numerator evaluation and management (E&M) structure (in 2011); addition of the 90-day lookback period for the lumbar spine surgery exclusion (in 2012); and, addition of congenital spine/spinal cord malformations, inflammatory and autoimmune disorders, infectious conditions, spinal vascular malformations, spinal cord infarctions, effects from radiation, spinal abnormalities associated with scoliosis, syringohydromyelia, and postoperative fluid collections/soft tissue changes (in 2014).

**S.4. Numerator Statement** (Brief, narrative description of the measure focus or what is being measured about the target population, i.e., cases from the target population with the target process, condition, event, or outcome) DO NOT include the rationale for the measure.

**IF an OUTCOME MEASURE**, state the outcome being measured. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

MRI of the lumbar spine studies with a diagnosis of low back pain (from the denominator) without the patient having claims-based

evidence of prior antecedent conservative therapy.

**S.5. Numerator Details** (All information required to identify and calculate the cases from the target population with the target process, condition, event, or outcome such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b)

IF an OUTCOME MEASURE, describe how the observed outcome is identified/counted. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

For MRI lumbar-spine studies in the denominator, the numerator is defined by the following categories of antecedent conservative therapy:

- Claim(s) for physical therapy in the 60 days preceding the lumbar spine MRI
- Claim(s) for chiropractic evaluation and manipulative treatment in the 60 days preceding the lumbar spine MRI
- Claim(s) for evaluation and management =28 days and =60 days preceding the lumbar spine MRI

(Specific CPT codes for each type of antecedent conservative therapy are included in the value set for this measure; this detailed list can be found in the Excel workbook provided for criterion S2b.)

Time Period for Data: MRI lumbar spine studies with no evidence of antecedent conservative therapy (chiropractic or physical therapy within 60 days of the MRI study or an evaluation and management visit within 28 days to 60 days of the MRI study), for patients with low back pain, performed within a 12-month time window.

**S.6. Denominator Statement** (Brief, narrative description of the target population being measured)

The number of MRI of the lumbar spine studies with a diagnosis of low back pain on the imaging claim performed in a hospital outpatient department on Medicare FFS beneficiaries within a 12-month time window.

**S.7. Denominator Details** (All information required to identify and calculate the target population/denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b.)

IF an OUTCOME MEASURE, describe how the target population is identified. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

The denominator is defined by the following CPT codes:

MRI Lumbar Spine

CPT 72148, 72149, 72158

MRI lumbar spine CPT codes should be accompanied by a diagnosis of low back pain on the same claim:

ICD-9 codes 721.3, 721.90, 722.10, 722.52, 722.6, 722.93, 724.02, 724.2, 724.3, 724.5, 724.6, 724.70, 724.71, 724.79, 738.5, 739.3, 739.4, 846.1, 846.2, 846.3, 846.4, 846.8, 846.9, 847.2

ICD-10 codes M43.20, M43.25-M43.28, M43.5X5-M43.5X9, M43.8X5-M43.8X9, M43.9, M46.46-M46.47, M47.20, M47.26-M47.28, M47.816-M47.819, M47.896-M47.9, M48.06-M48.07, M51.26-M51.27, M51.34-M51.37, M51.86-M51.87, M53.2X7-M53.2X8, M53.3, M53.86-M53.88, M54.30-M54.32, M54.40-M54.42, M54.5, M54.89, M54.9, M99.03-M99.04, M99.23, M99.33, M99.43, M99.53, M99.63, M99.73, M99.83-M99.84, S33.5XXA-S33.9XXS

The diagnosis of low back pain must be on the MRI lumbar-spine claim (i.e., the lumbar-spine MRI must be billed with a low back pain diagnosis in one of the diagnoses fields on the claim). MRI lumbar spine studies without a diagnosis of low back pain on the claim are not included in the denominator count. If a patient had more than one MRI lumbar spine study for a diagnosis of low back pain on the same day only one study would be counted, but if a patient had multiple MRI lumbar spine studies with a diagnosis of low back pain on a claim during the measurement period each study would be counted (i.e., a patient can be included in the denominator count more than once).

Global and TC claims are considered in order to capture all outpatient volume facility claims, typically paid under the Outpatient Prospective Payment System (OPPS)/Ambulatory Payment Classifications (APC) methodology, and to avoid double counting of professional component claims (i.e., 26 modifier).

A technical unit can be identified by a modifier code of TC. A global unit can be identified by the absence of a TC or 26-modifier code.

MRI lumbar spine studies can be billed separately for the technical and professional components, or billed globally, which includes both the professional and technical components.

Professional component claims will outnumber TC claims due to over-reads.

**S.8. Denominator Exclusions** *(Brief narrative description of exclusions from the target population)*

Below, in Section S.11 we provide a detailed list of denominator exclusion conditions. Denominator exclusions are consistent with current guidelines, evidence in literature, and guidance from the measure TEP.

**S.9. Denominator Exclusion Details** *(All information required to identify and calculate exclusions from the denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b.)*

Indications for measure exclusion include any patients with the following diagnosis code categories:

- Patients with lumbar spine surgery in the 90 days prior to MRI
- Cancer (within twelve months prior to MRI procedure)
- Congenital spine and spinal cord malformations (within five years prior to MRI procedure)
- Inflammatory and autoimmune disorders (within five years prior to MRI procedure)
- Infectious conditions (within one year prior to MRI procedure)
- Spinal vascular malformations and/or the cause of occult subarachnoid hemorrhage (within five years prior to MRI procedure)
- Spinal cord infarction (within one year prior to MRI procedure)
- Neoplastic abnormalities (within five years prior to MRI procedure)
- Treatment fields for radiation therapy (within five years prior to MRI procedure)
- Spinal abnormalities associated with scoliosis (within five years prior to MRI procedure)
- Syringomyelia (within five years prior to MRI procedure)
- Postoperative fluid collections and soft tissue changes (within one year prior to MRI procedure)
- Trauma (within 45 days prior to MRI procedure)
- IV drug abuse (within twelve months prior to MRI procedure)
- Neurologic impairment: (within twelve months prior to MRI procedure)
- HIV (within twelve months prior to MRI procedure)
- Unspecified immune deficiencies (within twelve months prior to MRI procedure)
- Intraspinal abscess (an exclusion diagnosis must be in one of the diagnoses fields on the MRI lumbar spine claim)

(Specific CPT codes, ICD-9 codes, and ICD-10 codes for exclusion are included in the value sets for this measure; this detailed list can be found in the Excel workbook provided for criterion S2b.)

**S.10. Stratification Information** *(Provide all information required to stratify the measure results, if necessary, including the stratification variables, definitions, specific data collection items/responses, code/value sets, and the risk-model covariates and coefficients for the clinically-adjusted version of the measure when appropriate – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format with at S.2b.)*

Not applicable; this measure does not stratify its results.

**S.11. Risk Adjustment Type** (Select type. Provide specifications for risk stratification in measure testing attachment)

No risk adjustment or risk stratification

If other:

**S.12. Type of score:**

Other (specify):

If other: Percentage

**S.13. Interpretation of Score** *(Classifies interpretation of score according to whether better quality is associated with a higher score, a lower score, a score falling within a defined interval, or a passing score)*

Better quality = Lower score

**S.14. Calculation Algorithm/Measure Logic** *(Diagram or describe the calculation of the measure score as an ordered sequence of steps including identifying the target population; exclusions; cases meeting the target process, condition, event, or outcome; time*

*period for data, aggregating data; risk adjustment; etc.)*

This measure calculates the percentage of lumbar-spine MRI studies with a diagnosis of low back pain on the imaging claim for which the patient did not have prior claims-based evidence of antecedent conservative therapy. The measure is calculated based on hospital outpatient claims data, as follows:

1. Select hospital outpatient claims with a CPT code for any MRI lumbar-spine study on a revenue line item
2. Exclude professional component only claims with modifier = '26'
3. Of claims identified in step 2, review relevant look-back periods for claims-based evidence of any procedure or diagnosis excluded from the measure; remove claims for which an exclusion has been identified
4. Set denominator counter = 1
5. Of claims identified in step 4, identify those claims for which there is no evidence of prior conservative therapy (claims for physical therapy in the 60 days preceding the imaging study; claims for chiropractic evaluation in the 60 days preceding the imaging study; or, claims for evaluation and management of at least 28 but equal to or less than 60 days prior to the imaging study). Set numerator count=1 for these claims
6. Aggregate denominator and numerator counts by facility identifier
7. Measure = numerator counts / denominator counts [The value should be recorded as a percentage]

**S.15. Sampling** *(If measure is based on a sample, provide instructions for obtaining the sample and guidance on minimum sample size.)*

IF an instrument-based performance measure (e.g., PRO-PM), identify whether (and how) proxy responses are allowed. This measure relies exclusively on 100% Medicare FFS Standard Analytical File (SAF) data; no sampling of beneficiaries was performed.

**S.16. Survey/Patient-reported data** *(If measure is based on a survey or instrument, provide instructions for data collection and guidance on minimum response rate.)*

Specify calculation of response rates to be reported with performance measure results.

This measure does not use survey data.

**S.17. Data Source** *(Check ONLY the sources for which the measure is SPECIFIED AND TESTED).*

*If other, please describe in S.18.*

Claims

**S.18. Data Source or Collection Instrument** *(Identify the specific data source/data collection instrument (e.g. name of database, clinical registry, collection instrument, etc., and describe how data are collected.)*

IF instrument-based, identify the specific instrument(s) and standard methods, modes, and languages of administration.

This measure is not a PRO-PM measure.

**S.19. Data Source or Collection Instrument** *(available at measure-specific Web page URL identified in S.1 OR in attached appendix at A.1)*

No data collection instrument provided

**S.20. Level of Analysis** *(Check ONLY the levels of analysis for which the measure is SPECIFIED AND TESTED)*

Facility, Population : Regional and State

**S.21. Care Setting** *(Check ONLY the settings for which the measure is SPECIFIED AND TESTED)*

Emergency Department and Services, Inpatient/Hospital, Outpatient Services

If other:

**S.22. COMPOSITE Performance Measure** - Additional Specifications *(Use this section as needed for aggregation and weighting rules, or calculation of individual performance measures if not individually endorsed.)*

Not applicable; this is not a composite measure.

**2. Validity – See attached Measure Testing Submission Form**

NQF\_0514\_MeasureTestingForm.docx

**2.1 For maintenance of endorsement**

*Reliability testing: If testing of reliability of the measure score was not presented in prior submission(s), has reliability testing of the measure score been conducted? If yes, please provide results in the Testing attachment. Please use the most current version of the*

testing attachment (v7.1). Include information on all testing conducted (prior testing as well as any new testing); use red font to indicate updated testing.

## **2.2 For maintenance of endorsement**

Has additional empirical validity testing of the measure score been conducted? If yes, please provide results in the Testing attachment. Please use the most current version of the testing attachment (v7.1). Include information on all testing conducted (prior testing as well as any new testing); use red font to indicate updated testing.

## **2.3 For maintenance of endorsement**

**Risk adjustment:** For outcome, resource use, cost, and some process measures, risk-adjustment that includes social risk factors is not prohibited at present. Please update sections 1.8, 2a2, 2b1,2b4.3 and 2b5 in the Testing attachment and S.140 and S.11 in the online submission form. NOTE: These sections must be updated even if social risk factors are not included in the risk-adjustment strategy. You MUST use the most current version of the Testing Attachment (v7.1) -- older versions of the form will not have all required questions.

## **3. Feasibility**

Extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.

### **3a. Byproduct of Care Processes**

For clinical measures, the required data elements are routinely generated and used during care delivery (e.g., blood pressure, lab test, diagnosis, medication order).

#### **3a.1. Data Elements Generated as Byproduct of Care Processes.**

Coded by someone other than person obtaining original information (e.g., DRG, ICD-9 codes on claims)

If other:

### **3b. Electronic Sources**

The required data elements are available in electronic health records or other electronic sources. If the required data are not in electronic health records or existing electronic sources, a credible, near-term path to electronic collection is specified.

**3b.1. To what extent are the specified data elements available electronically in defined fields (i.e., data elements that are needed to compute the performance measure score are in defined, computer-readable fields)** Update this field for **maintenance of endorsement**.

ALL data elements are in defined fields in electronic claims

**3b.2. If ALL the data elements needed to compute the performance measure score are not from electronic sources, specify a credible, near-term path to electronic capture, OR provide a rationale for using other than electronic sources.** For **maintenance of endorsement**, if this measure is not an eMeasure (eCQM), please describe any efforts to develop an eMeasure (eCQM).

**3b.3. If this is an eMeasure, provide a summary of the feasibility assessment in an attached file or make available at a measure-specific URL. Please also complete and attach the NQF Feasibility Score Card.**

Attachment:

### **3c. Data Collection Strategy**

Demonstration that the data collection strategy (e.g., source, timing, frequency, sampling, patient confidentiality, costs associated with fees/licensing of proprietary measures) can be implemented (e.g., already in operational use, or testing demonstrates that it is ready to put into operational use). For eMeasures, a feasibility assessment addresses the data elements and measure logic and demonstrates the eMeasure can be implemented or feasibility concerns can be adequately addressed.

**3c.1. Required for maintenance of endorsement.** Describe difficulties (as a result of testing and/or operational use of the

measure) regarding data collection, availability of data, missing data, timing and frequency of data collection, sampling, patient confidentiality, time and cost of data collection, other feasibility/implementation issues.

**IF instrument-based,** consider implications for both individuals providing data (patients, service recipients, respondents) and those whose performance is being measured.

This measure is claims-based, and uses CMS hospital outpatient claims as its data source.

Special attention needs to be taken when counting procedures on the Medicare claims files. The biggest issue is how to deal with modifier codes. Modifiers are two digit indicators (alpha or numeric) that represent a service or procedure that has been altered by some specific circumstance, which typically will impact the payment amount.

Procedure modifier code “26” represents the professional component of a procedure and includes the clinician work (i.e., the reading of the image by a physician), associated overhead and professional liability insurance costs. This modifier corresponds to the human involvement in a given service or procedure.

The procedure modifier code “TC” represents the technical component of a service or procedure and includes the cost of equipment and supplies to perform that service or procedure. This modifier corresponds to the equipment/facility part of a given service or procedure.

In most cases, unmodified codes represent a global procedure which includes both the professional and technical components. There are also other modifier codes. All other modifier codes have been counted as a technical code for our purposes. When calculating the measures, we are only concerned with procedures associated with technical and global modifiers, as these modifiers refer to services provided by the facility. This reduces the possibility of double-counting procedures, since a single procedure may result in both a technical and professional record on the claims files. There were very few instances when this occurred as it related to procedures applicable to the measure.

When developing counts of procedures, the objective is to avoid double-counting procedures that may have been billed through multiple revenue centers within a facility. Billing through multiple centers leads to multiple records in the Medicare claims files (i.e., the SAFs). For instance, there may be multiple bills for a single MRI. On one bill, the charges relate to the application of a radiopharmaceutical, which could have a technical modifier code and come from the pharmacy revenue center. On the other bill, the charges relate to the imaging study and may fall under a technical bill from the imaging center revenue center. In this case, we only count the MRI once, since only one MRI was performed. However, if we were summing up the Medicare paid amounts for this procedure, we would include the Medicare paid amounts from both bills, as they each represent payments for services directly related to the particular MRI procedure.

**3c.2. Describe any fees, licensing, or other requirements to use any aspect of the measure as specified (e.g., value/code set, risk model, programming code, algorithm).**

No fees, licensure, or other requirements are necessary to use this measure; however, CPT codes, descriptions, and other data are copyright 2015 American Medical Association. All rights reserved. CPT is a registered trademark of the American Medical Association. Applicable FARS\DFARS Restrictions Apply to Government Use. Fee schedules, relative value units, conversion factors, and/or related components are not assigned by the AMA, are not part of CPT, and the AMA is not recommending their use. The AMA does not directly or indirectly practice medicine or dispense medical services. The AMA assumes no liability for data contained or not contained herein.

## 4. Usability and Use

Extent to which potential audiences (e.g., consumers, purchasers, providers, policy makers) are using or could use performance results for both accountability and performance improvement to achieve the goal of high-quality, efficient healthcare for individuals or populations.

### 4a. Accountability and Transparency

Performance results are used in at least one accountability application within three years after initial endorsement and are publicly reported within six years after initial endorsement (or the data on performance results are available). If not in use at the time of initial endorsement, then a credible plan for implementation within the specified timeframes is provided.

#### 4.1. Current and Planned Use

*NQF-endorsed measures are expected to be used in at least one accountability application within 3 years and publicly reported within 6 years of initial endorsement in addition to performance improvement.*

| Specific Plan for Use | Current Use (for current use provide URL) |
|-----------------------|-------------------------------------------|
|                       |                                           |

**4a1.1 For each CURRENT use, checked above (update for maintenance of endorsement), provide:**

- Name of program and sponsor
- Purpose
- Geographic area and number and percentage of accountable entities and patients included
- Level of measurement and setting

**Public Reporting:**

Name of program and sponsor: The CMS Hospital Outpatient Quality Reporting (HOQR) Program

Purpose: The HOQR Program is a pay for quality data reporting program implemented by CMS for outpatient hospital services. In addition to providing hospitals with a financial incentive to report their quality of care measure data, the HOQR Program provides CMS with data to help Medicare beneficiaries make more informed decisions about their health care. Hospital quality of care information gathered through the HOQR Program is publicly available on the Hospital Compare website.

Accountable entities and patients: The publicly reported values (on Hospital Compare) are calculated for all facilities in the United States that meet minimum case count and other reporting requirements. For the period of 2009 to 2014, 1,189 facilities met the minimum case count each year. Additional facilities met the minimum case count requirements in some, but not all, years. The claims included in the publicly reported calculations are for Medicare FFS patients whose claims are subject to OPSS.

Level of measurement and setting: HOQR measures are measured at the facility, state, and national level. The setting for HOQR measures is the hospital outpatient care setting.

**Quality Improvement with Benchmarking (external benchmarking to multiple organizations):**

Name of program and sponsor: The CMS HOQR Program

Purpose: The HOQR Program is a pay for quality data reporting program implemented by CMS for outpatient hospital services. In addition to providing hospitals with a financial incentive to report their quality of care measure data, the data is publicly reported on the Hospital Compare website. The data reported on Hospital Compare not only shows the hospital's score on the measure, but also provides state and national averages for the measure. This enables consumers to compare the hospital's performance to other facilities and determine if the facility is an outlier.

Accountable entities and patients: The publicly reported values (on Hospital Compare) are calculated for all facilities in the United States that meet minimum case count and other reporting requirements. For the period of 2009 to 2014, 1,189 facilities met the minimum case count each year. Additional facilities met the minimum case count requirements in some, but not all, years. The claims included in the publicly reported calculations are for Medicare FFS patients whose claims are subject to the OPSS.

Level of measurement and setting: HOQR measures are measured at the facility, state, and national level. The setting for HOQR measures is the hospital outpatient care setting.

**4a1.2. If not currently publicly reported OR used in at least one other accountability application (e.g., payment program, certification, licensing) what are the reasons? (e.g., Do policies or actions of the developer/steward or accountable entities restrict access to performance results or impede implementation?)**

*This measure is publicly reported.*

**4a1.3. If not currently publicly reported OR used in at least one other accountability application, provide a credible plan for implementation within the expected timeframes -- any accountability application within 3 years and publicly reported within 6 years of initial endorsement. (Credible plan includes the specific program, purpose, intended audience, and timeline for implementing the measure within the specified timeframes. A plan for accountability applications addresses mechanisms for data aggregation and reporting.)**

*This measure is publicly reported.*

**4a2.1.1. Describe how performance results, data, and assistance with interpretation have been provided to those being measured or other users during development or implementation.**

**How many and which types of measured entities and/or others were included? If only a sample of measured entities were included, describe the full population and how the sample was selected.**

**4a2.1.2. Describe the process(es) involved, including when/how often results were provided, what data were provided, what educational/explanatory efforts were made, etc.**

**4a2.2.1. Summarize the feedback on measure performance and implementation from the measured entities and others described in 4d.1.**

**Describe how feedback was obtained.**

**4a2.2.2. Summarize the feedback obtained from those being measured.**

**4a2.2.3. Summarize the feedback obtained from other users**

**4a2.3. Describe how the feedback described in 4a2.2.1 has been considered when developing or revising the measure specifications or implementation, including whether the measure was modified and why or why not.**

#### **Improvement**

Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated. If not in use for performance improvement at the time of initial endorsement, then a credible rationale describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.

**4b1. Refer to data provided in 1b but do not repeat here. Discuss any progress on improvement (trends in performance results, number and percentage of people receiving high-quality healthcare; Geographic area and number and percentage of accountable entities and patients included.)**

**If no improvement was demonstrated, what are the reasons? If not in use for performance improvement at the time of initial endorsement, provide a credible rationale that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.**

*Though national performance has not improved in the hospital outpatient setting since the inception of public reporting, an improvement in performance over the last two years suggests that facilities are beginning to address this gap in quality of care.*

#### **4b2. Unintended Consequences**

The benefits of the performance measure in facilitating progress toward achieving high-quality, efficient healthcare for individuals or populations outweigh evidence of unintended negative consequences to individuals or populations (if such evidence exists).

**4b2.1. Please explain any unexpected findings (positive or negative) during implementation of this measure including unintended impacts on patients.**

*CMS conducts prepayment claims analysis and post-payment audits that should prevent this factor from having a major impact on the measure calculations performed on claims data.*

**4b2.2. Please explain any unexpected benefits from implementation of this measure.**

### **5. Comparison to Related or Competing Measures**

If a measure meets the above criteria and there are endorsed or new related measures (either the same measure focus or the same target population) or competing measures (both the same measure focus and the same target population), the measures are compared to address harmonization and/or selection of the best measure.

#### **5. Relation to Other NQF-endorsed Measures**

Are there related measures (conceptually, either same measure focus or target population) or competing measures (conceptually

both the same measure focus and same target population)? If yes, list the NQF # and title of all related and/or competing measures.

Yes

**5.1a. List of related or competing measures (selected from NQF-endorsed measures)**

0052 : Use of Imaging Studies for Low Back Pain

**5.1b. If related or competing measures are not NQF endorsed please indicate measure title and steward.**

National Committee for Quality Assurance (NCQA) - Back Pain: Appropriate Imaging for Acute Back Pain

Institute for Clinical Systems Improvement (ICSI) - Adult acute and subacute low back pain: Percentage of patients with a diagnosis of non-specific back pain for whom the physician ordered imaging studies during the six weeks after pain onset, in the absence of red flags

ICSI - Adult acute and subacute low back pain: Percentage of patients with radicular pain for whom the clinician ordered imaging studies during the six weeks after pain onset

**5a. Harmonization of Related Measures**

The measure specifications are harmonized with related measures;

**OR**

The differences in specifications are justified

**5a.1. If this measure conceptually addresses EITHER the same measure focus OR the same target population as NQF-endorsed measure(s):**

**Are the measure specifications harmonized to the extent possible?**

No

**5a.2. If the measure specifications are not completely harmonized, identify the differences, rationale, and impact on interpretability and data collection burden.**

NQF #0514 is similar in construct to NQF measure #0052, Use of Imaging Studies for Low Back Pain (developed by NCQA). Both measures consider the overuse of imaging for patients with a diagnosis of low back pain. However, the measures have key differences in intent and patient population that limit the desirability of complete harmonization. NQF #0052 is a resource utilization measure and does not consider the administration of antecedent conservative therapy prior to imaging (captured in NQF #0514); NQF #0052 includes multiple imaging modalities (i.e., CT, MRI, and X-ray) and several regions of the spine (lumbar, cervical, and thoracic), while NQF #0514 focuses solely on MRI lumbar spine studies; and, NQF #0514 is calculated using Medicare claims data while NQF #0052 is calculated using both administrative claims and electronic clinical data. The stewards for NQF #0514 and #0052 have held a series of harmonization discussions, focusing on the clinical intents of the measures and looking for opportunities for alignment. To better harmonize the measures, the stewards have updated the value sets used to identify low back pain and exclusion conditions. The stewards recommend against complete harmonization due to differences in the clinical focus (only lumbar spine imaging vs. lumbar, cervical, and thoracic spine imaging), intent (inappropriate use vs. resource utilization), age of target population, and structure of the measures. While NQF #0514 is related to the second NCQA measure, significant structural differences make harmonization undesirable: the NCQA measure is calculated for individual physicians; the NCQA measure does not consider the administration of antecedent conservative therapy prior to imaging; the measure includes several imaging modalities, and; it focuses on a different patient population (patients aged 18-80) than NQF #0514. Similarly, while NQF #0514 is related to the ICSI measures, the measures have key differences that limit the desirability of complete harmonization. The first ICSI measure differs from NQF #0514 in several key ways: the measures target different patient populations, and; the ICSI measure does not consider antecedent conservative therapy prior to imaging. There are significant differences between the second ICSI measure and NQF #0514, including: targeted patient populations, and consideration (or not) of antecedent conservative therapy prior to imaging.

**5b. Competing Measures**

The measure is superior to competing measures (e.g., is a more valid or efficient way to measure);

**OR**

Multiple measures are justified.

**5b.1. If this measure conceptually addresses both the same measure focus and the same target population as NQF-endorsed measure(s):**

**Describe why this measure is superior to competing measures (e.g., a more valid or efficient way to measure quality); OR provide a rationale for the additive value of endorsing an additional measure. (Provide analyses when possible.)**

This measure addresses a different target population than does the NCQA measure, and, consequently, the measures are not

viewed as competing measures.

## Appendix

**A.1 Supplemental materials may be provided in an appendix.** All supplemental materials (such as data collection instrument or methodology reports) should be organized in one file with a table of contents or bookmarks. If material pertains to a specific submission form number, that should be indicated. Requested information should be provided in the submission form and required attachments. There is no guarantee that supplemental materials will be reviewed.

Available at measure-specific web page URL identified in S.1 **Attachment:**

## Contact Information

**Co.1 Measure Steward (Intellectual Property Owner):** Centers for Medicare & Medicaid Services

**Co.2 Point of Contact:** Vinitha, Meyyur, [Vinitha.Meyyur@cms.hhs.gov](mailto:Vinitha.Meyyur@cms.hhs.gov), 410-786-7224-

**Co.3 Measure Developer if different from Measure Steward:** The Lewin Group

**Co.4 Point of Contact:** Colleen, McKiernan, [Colleen.McKiernan@lewin.com](mailto:Colleen.McKiernan@lewin.com), 703-269-5595-

## Additional Information

### Ad.1 Workgroup/Expert Panel involved in measure development

**Provide a list of sponsoring organizations and workgroup/panel members' names and organizations. Describe the members' role in measure development.**

The contractor has convened a TEP, which will evaluate and provide feedback on measure-development and maintenance efforts for the imaging efficiency measures. Specifically, the TEP provides direction and feedback through all phases of project activities including updates to the current specifications of the six imaging efficiency measures, review of quantitative testing results, feedback on qualitative testing questions (i.e., results of TEP member questionnaires), and support for endorsement of the measures by NQF.

The following is a list of the contractor's TEP members:

Meenu Arora, MBA  
Quality Improvement Leader, Sequoia Hospital

Brian Baker  
Chief Executive Officer, Carealytics

Peter Benner  
Vice Chair, MNSure

Martha Deed, Ph.D  
Safe Patient Project's Patient Advocacy Network

Lawrence Feinberg, MD  
Attending Physician, University of Colorado Hospital

Elliott Fishman, MD  
Professor of Radiology and Oncology, Johns Hopkins School of Medicine

Marian Hollingsworth  
Patient Advocate

Michael Hutchinson, MD, Ph.D  
Clinical Associate Professor of Neurology, Icahn School of Medicine at Mount Sinai

Gregory M. Kusiak, MBA, FRBMA

President, California Medical Business Services, Inc.

Barbara Landreth, RN, MBA  
Clinical Information Analyst , St. Louis Area Business Health Coalition

Barbara McNeil, MD, Ph.D  
Head Professor of Radiology, Harvard University

Michael J. Pentecost, MD  
Chief Medical Officer, NIA Magellan

David Seidenwurm, MD  
Medical Staff Consultant, Sutter Medical Group

Adam Sharp, MD, MS  
Research Scientist , Kaiser Permanente Southern California

Paul R. Sierzenski, MD, RDMS, FACEP, FAAEM  
Medical Director, Christian Health Care System

C. Todd Staub, MD, FACP  
Chairman, ProHealth Physicians

**Measure Developer/Steward Updates and Ongoing Maintenance**

**Ad.2 Year the measure was first released:** 2011

**Ad.3 Month and Year of most recent revision:** 03, 2016

**Ad.4 What is your frequency for review/update of this measure?** Annually

**Ad.5 When is the next scheduled review/update for this measure?** 03, 2017

**Ad.6 Copyright statement:** This measure does not have a copyright.

**Ad.7 Disclaimers:** CPT codes, descriptions, and other data only are copyright 2015 American Medical Association. All rights reserved. CPT is a registered trademark of the American Medical Association (AMA). Applicable Federal Acquisition Regulation Site (FARS)\Defense Federal Acquisition Regulation Statement (DFARS) Restrictions Apply to Government Use. Fee schedules, relative value units, conversion factors and/or related components are not assigned by the AMA, are not part of CPT, and the AMA is not recommending their use. The AMA does not directly or indirectly practice medicine or dispense medical services. The AMA assumes no liability for data contained or not contained herein.

**Ad.8 Additional Information/Comments:**