NATIONAL QUALITY FORUM

MEASURE WORKSHEET

This document summarizes the evaluation of the measure as it progresses through NQF's Consensus Development Process (CDP). The information submitted by measure developers/stewards is included after the Brief Measure Information, Preliminary Analysis, and Pre-meeting Public and Member Comments sections.

To navigate the links in the worksheet: Ctrl + click link to go to the link; ALT + LEFT ARROW to return

Brief Measure Information

NQF #: 2720

De.2. Measure Title: National Healthcare Safety Network (NHSN) Antimicrobial Use Measure

Co.1.1. Measure Steward: Centers for Disease Control and Prevention

De.3. Brief Description of Measure: This measure assesses antimicrobial use in hospitals based on medication administration data that hospitals collect electronically at the point of care and report via electronic file submissions to CDC's National Healthcare Safety Network (NHSN). The antimicrobial use data that are in scope for this measure are antibacterial agents administered to adult and pediatric patients in a specified set of ward and intensive care unit locations: medical ICU, medical-surgical ICU, surgical ICU (adult only), medical ward, medical-surgical ward, surgical ward, general hematology-oncology ward (adult only), and step-down unit (adult only). The measure compares antimicrobial use that the hospitals report with antimicrobial use that is predicted on the basis of nationally aggregated data. The measure is comprised of a discrete set of ratios, Standardized Antimicrobial Administration Ratios (SAARs), each of which summarizes observed-to-predicted antimicrobial use for one of 40 antimicrobial agent-patient care location combinations. The SAARs are designed to serve as high value targets or high level indicators for antimicrobial stewardship programs (ASPs). SAAR values that are outliers are intended to prompt analysis of possible overuse, underuse, or inappropriate use of antimicrobials, subsequent actions aimed at improving the quality of antimicrobial prescribing, and impact evaluations of ASP interventions.

1b.1. Developer Rationale: The measure provides summary results that hospital and health system antimicrobial stewardship programs (ASPs) can use as quantitative aids in their efforts to evaluate and improve antibiotic prescribing. The Standardized Antimicrobial Administration Ratios (SAARs) that comprise the measure focus on high value targets and high level indicators of antimicrobial use for ASPs. The SAARs can be used by ASPs to benchmark antimicrobial use in multiple patient care locations, identify opportunities for improvement, and gauge the impact of stewardship efforts. At the outset, the SAARs provide a set of signals that often warrant further analysis, such as an evaluation of the extent to which a specific antibiotic or group of antibiotics accounts for a high or low SAAR value and the extent to which an antibiotic or group of antibiotics were used appropriately for specific indications. While the SAARs do not provide a definitive indication that antibiotics are overused or underused, they provide an important starting place for further analysis and possible action. Some of the analytic follow up can be completed with hospital- and patient care location-specific data reported to CDC's National Healthcare Safety Network (NHSN) Antimicrobial Use and Resistance (AUR) Module, using analytic features built into the NHSN application. However, additional analyses to determine the appropriateness of antibiotic use in individual instances are likely to require access to detailed, patient-level data that is beyond the scope of data collection and analysis using the NHSN module, e.g., clinical indications for specific antibiotics and dose and duration decisions.

S.4. Numerator Statement: Days of antimicrobial therapy for antimicrobial agents administered to adult and pediatric patients in medical ICU, medical-surgical ICU, surgical ICU (adult only), medical ward, medical-surgical ward, surgical ward, general hematology-oncology ward (adult only), and step-down unit (adult only).

S.6. Denominator Statement: Days present for each patient care location—adult and pediatric medical ICU, medical-surgical ICU, surgical ICU (adult only), medical ward, medical-surgical ward, surgical ward, general hematology-oncology ward (adult only), and step-down unit (adult only) is defined as the number of patients who were present for any portion of each day of a calendar month for each location. The day of admission, discharge, and transfer to and from locations are included in days present. All days present are summed for each location and month, and the aggregate sums for each location-month combination comprise the denominator data for the measure.

S.8. Denominator Exclusions: Hospital patient care locations other than adult and pediatric medical ICU, medical-surgical ICU, surgical ICU (adult only), medical ward, medical-surgical ward, surgical ward, general hematology-oncology ward (adult only), and step-down unit (adult only) are excluded from this measure.

De.1. Measure Type: Process

S.17. Data Source: Electronic health data, electronic format Admission Discharge Transfer

S.20. Level of Analysis: Facility

IF Endorsement Maintenance – Original Endorsement Date: Dec 10, 2015 **Most Recent Endorsement Date:** Dec 10, 2015

Preliminary Analysis: Maintenance of Endorsement

To maintain NQF endorsement endorsed measures are evaluated periodically to ensure that the measures still meets the NQF endorsement criteria ("maintenance"). The emphasis for maintaining endorsement is focused on how effective the measure is for promoting improvements in quality. Endorsed measures should have some experience from the field to inform the evaluation. The emphasis for maintaining endorsement is noted for each criterion.

Criteria 1: Importance to Measure and Report

1a. Evidence

Maintenance measures – less emphasis on evidence unless there is new information or change in evidence since the prior evaluation.

1a. Evidence. The evidence requirements for a *structure, process or intermediate outcome* measure is that it is based on a systematic review (SR) and grading of the body of empirical evidence where the specific focus of the evidence matches what is being measured. For measures derived from patient report, evidence also should demonstrate that the target population values the measured process or structure and finds it meaningful.

The developer provides the following evidence for this measure:

- **Systematic Review of the evidence specific to this measure?** ☒ **Yes** ☐ **No**
- **Quality, Quantity and Consistency of evidence provided?** ☒ **Yes** ☐ **No**
- **Evidence graded?** ☒ **Yes** ☐ **No**

Summary of prior review in 2015

- Data from the ISDA/SHEA guidelines for developing an institutional program to enhance antimicrobial stewardship (2007) was presented along with four other systematic reviews.
    - Guideline Recommendations include:
        - Using proactive strategies that form the basis of an antimicrobial stewardship program: prospective audit with intervention and feedback (A-I), formulary restriction and preauthorization (A-II), and using preauthorization, monitoring overall trends in antimicrobial use is necessary (B-III).
        - Using computer-based surveillance to efficiently target interventions, track resistance patterns, and identify nosocomial infections and ADEs (B-II)
        - Process measures and outcomes measures are useful to determine the impact of antimicrobial stewardship (B-III)
- Four additional systematic reviews support the evidence base. Main findings:
    - (Feazel, 2014)…implementation of ASPs had an overall protective benefit (pooled risk ratio: 0.48; 95% CI: 0.38, 0.62), indicating a risk reduction for CDI of 52%.
    - (Davey et al., 2013) Interventions intended to decrease excessive prescribing were associated with reduction in CDI and colonization or infection with aminoglycoside- or cephalosporin-resistant gram-negative bacteria, methicillin-resistant Staphylococcus aureus and vancomycin-resistant Enterococcus faecalis. Meta-analysis showed that four interventions intended to increase effective prescribing for pneumonia were associated with significant reduction in mortality (risk ratio 0.89, 95% CI 0.82 to 0.97), whereas nine interventions intended to decrease excessive prescribing were not associated with significant increase in mortality (risk ratio 0.92, 95% CI 0.81 to 1.06).
    - (Kaki et al., 2011) Most stewardship interventions are associated with a decrease in either targeted or overall antibiotic use in critical care patients.
    - (Davey et al., 2006) The evidence supports the theory that limiting the use of specific antimicrobial drugs will reduce the prevalence of resistant gram-negative bacteria and C. diff infection. For gram-positive bacteria, there is a lack of evidence rather than evidence of no effect.

Changes to evidence from last review

☐ **The developer attests that there have been no changes in the evidence since the measure was last evaluated.**

☒ **The developer provided updated evidence for this measure:**

Updates:

- The developer provides a 2017 systematic review, *Interventions to improve antibiotic prescribing practices for hospital inpatients*, which updates the Davey 2006 and 2013 reviews previously submitted. Evidence included 221 studies. Main findings include:
    - Enablement (increasing means/reducing barriers to increase capability or opportunity), and restriction (using rules to reduce the opportunity to engage in the target behavior), are effective in increasing compliance with antibiotic policy and reducing duration of antibiotic treatment (high certainty).
    - Lower use of antibiotics has not been associated with a difference in mortality and likely reduces length of stay (moderate certainty)
    - Impact on microbial outcomes (very low certainty)

- Three recommendations from *Implementing an Antibiotic Stewardship Program: Guidelines by the Infectious Diseases Society of America and the Society for Healthcare Epidemiology of America (2016)* are included.
    - We recommend preauthorization and/or prospective audit and feedback over no such interventions (Moderate-quality evidence; strong recommendation).

        - Findings from 23 studies consistently showed that preauthorization was associated with significant reduction in the use of the restricted agents and of associated costs, decreased antibiotic use, and decreased antibiotic resistance particularly among gram-negative pathogens.
    - We recommend antibiotic stewardship interventions designed to reduce the use of antibiotics associated with a high risk of CDI compared with no such intervention (Moderate quality evidence; strong recommendation).
        - Nine studies consistently found significant associations between ASP and decreased CDI incidence.
    - We suggest monitoring antibiotic use as measured by days of therapy (DOTs) in preference to defined daily dose (DDD) (low-quality evidence; weak recommendation).
        - Five studies identified a significant impact of ASP interventions in reduction of antibiotic prescribing, increased antibiotic susceptibility, decreased CDI incidence, however CDI incidence is affected by other practices and observed effects on resistance are unpredictable due to confounders.

Questions for the Committee:

- The evidence provided by the developer is updated, directionally the same, and stronger compared to that for the previous NQF review. Does the Committee agree there is no need for repeat discussion and vote on Evidence?

Guidance from the Evidence Algorithm

Process measure based on systematic review (Box 3) → QQC presented (Box 4) → Quantity: high; Quality: high; Consistency: moderate/high (Box 5) → High (Box 5a) → High

Preliminary rating for evidence: ☒ **High** ☐ **Moderate** ☐ **Low** ☐ **Insufficient**

RATIONALE:

1b. Gap in Care/Opportunity for Improvement and 1b. Disparities

Maintenance measures – increased emphasis on gap and variation

1b. Performance Gap. The performance gap requirements include demonstrating quality problems and opportunity for improvement.

- Standardized Antimicrobial Administration Ratios (SAARs) represent observed to predicted antimicrobial use.
- Mean and medican SAARs as well as the percent of SAARs statistically higher or lower than 1 are provided for both the adult and pediatric populations. SAARs are broken down by antibacterial agent grouping and unit (i.e., ICUs, wards, step down units and oncology units) (Table 3, attachment).
- For all agents and units for the adult population, 44% of SAARs are lower than 1, while 45% of SAARs are greater than 1.
- For all agents and units for the pediatric population, 43% of SAARs are lower than 1, while 40% are greater than 1.
- Scores above or below one may represent over or under use of antibiotics. There appears to be variation in performance compared to the predicted targets, representing potential opportunity for improvement for facilities.
- The developer also referenced that literature supports opportunities to improve the appropriateness of antibiotic use.

Disparities

- The developer states their data source does not include disparity variables, but there is no compelling literature that supports that variation in facility antimicrobial use is related to social risk factors.
- A retrospective analysis (1996-2007) of all surgical patients treated for sepsis at a tertiary care center demonstrated no differences in demographic and comorbidities between inappropriately and appropriately treated groups. (Davies et al, 2014)

Questions for the Committee:

- Is there a gap in care that warrants a national performance measure?
- Are you aware of evidence that disparities exist in this area of healthcare?

Preliminary rating for opportunity for improvement: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

RATIONALE:

Committee Pre-evaluation Comments:
Criteria 1: Importance to Measure and Report (including 1a, 1b, 1c)

1a. Evidence

Comments:

**The evidence for the problem this measure addresses is high.

**updated evidence provided and continues to be strong

**Developer provided updated evidence

**robust

1b. Performance Gap

Comments:

**Performance gap is moderate due to SAAR measure outside of expected does not definitely tell you whether care is better or worse. There is evidence that there is a large variation in antibiotic use, but whether that variance reflect effective care is in question.

**gap noted no disparities presented

**There is still a significant performance gap

**variability and oppty for improvement; my concern is with how truly representative these data are; my guess is situation is much more dire/variable

Criteria 2: Scientific Acceptability of Measure Properties

2a. Reliability: Specifications and Testing

2b. Validity: Testing; Exclusions; Risk-Adjustment; Meaningful Differences; Comparability; Missing Data

Reliability

2a1. Specifications requires the measure, as specified, to produce consistent (reliable) and credible (valid) results about the quality of care when implemented. For maintenance measures – no change in emphasis – specifications should be evaluated the same as with new measures.

2a2. Reliability testing demonstrates if the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise enough to distinguish differences in performance across providers. For maintenance measures – less emphasis if no new testing data provided.

Validity

2b2. Validity testing should demonstrate the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality. For maintenance measures – less emphasis if no new testing data provided.

2b2-2b6. Potential threats to validity should be assessed/addressed.

Complex measure evaluated by Scientific Methods Panel? ☐ **Yes** ☒ **No**

Evaluators: Patient Safety project team staff

Evaluation of Reliability and Validity: Link A (Project Team staff)

- Developer conducted validity testing of the numerator and denominator data elements.
    - Antimicrobial days numerator: percent agreement 60-80% (at the outset of validation) and Days present denominator: percent agreement 70-80% (at the outset of validation). By design the process led to >99% agreement for all required data elements prior to data submission to CDC.
    - Since data element validity testing was provided additional reliability testing is not required.
- Face validity was also tested by an expert panel of infectious disease physicians and clinical pharmacists by means of consensus development using a Delphi process. The panel reviewed the validity of core data elements and construct of antimicrobial groupings in 2018. Individual voting results are not provided.
- The measure is risk adjusted. Each group of SAAR antimicrobial agents is modeled separately. Models are also different for adults and pediatrics. Factors considered include hospital teaching status, hospital bedsize, hospital ICU bedsize, percentage of ICU beds among total beds, average length of hospital stay, and patient care location.

Questions for the Committee regarding reliability:

- Do you have any concerns that the measure can be consistently implemented (i.e., are measure specifications adequate)?

Questions for the Committee regarding validity:

- Do you have any concerns regarding the validity of the measure (e.g., risk-adjustment approach and variables included or excluded, ability to detect meaningful differences)?

Preliminary rating for reliability: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

Preliminary rating for validity: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

Evaluation A: Scientific Acceptability

Scientific Acceptability: Preliminary Analysis Form

Measure Number: 2720

Measure Title: National Healthcare Safety Network (NHSN) Antimicrobial Use Measure

Type of measure:

☒ **Process** ☐ **Process: Appropriate Use** ☐ **Structure** ☐ **Efficiency** ☐ **Cost/Resource Use**

☐ **Outcome** ☐ **Outcome: PRO-PM** ☐ **Outcome: Intermediate Clinical Outcome** ☐ **Composite**

Data Source:

☐ **Claims** ☒ **Electronic Health Data** ☐ **Electronic Health Records** ☐ **Management Data**

☐ **Assessment Data** ☐ **Paper Medical Records** ☐ **Instrument-Based Data** ☐ **Registry Data**

☐ **Enrollment Data** ☒ **Other:** electronic format Admission Discharge Transfer

Level of Analysis:

☐ **Clinician: Group/Practice** ☐ **Clinician: Individual** ☒ **Facility** ☐ **Health Plan**
☐ **Population: Community, County or City** ☐ **Population: Regional and State**
☐ **Integrated Delivery System** ☐ **Other**

Measure is:

☐ **New** ☒ **Previously endorsed (**NOTE: Empirical validity testing is expected at time of maintenance review; if not possible, justification is required.)

RELIABILITY: SPECIFICATIONS

1) **Are submitted specifications precise, unambiguous, and complete so that they can be consistently implemented?** ☒ **Yes** ☐ **No**

   Submission document: "MIF_xxxx" document, items S.1-S.22

   NOTE*: NQF staff will conduct a separate, more technical, check of eCQM specifications, value sets, logic, and feasibility, so no need to consider these in your evaluation.*

2) **Briefly summarize any concerns about the measure specifications.**

   No concerns. Updates from the previous evaluation: adult and pediatric patient populations were modeled separately, adult general hematology-oncology wards and adult step-down units were added, and antimicrobial groupings were added and re-categorized.

RELIABILITY: TESTING

Submission document: "MIF_xxxx" document for specifications, testing attachment questions 1.1-1.4 and section 2a2

3) **Reliability testing level** ☐ **Measure score** ☐ **Data element** ☒ **Neither**

4) **Reliability testing was conducted with the data source and level of analysis indicated for this measure** ☐ **Yes** ☐ **No**

5) If score-level and/or data element reliability testing was NOT conducted or if the methods used were NOT appropriate, was **empirical <u>VALIDITY</u> testing** of <u>patient-level data</u> conducted?

   ☒ **Yes** ☐ **No**

   - Data element testing was previously reviewed as reliability testing but meets the criteria to be considered data element VALIDITY testing. Per NQF requirements, for this measure, additional reliability testing is not required IF data element validity is demonstrated. Data element testing has been updated since the previous evaluation and will be evaluated in the validity section with results/ratings also applicable to reliability.

6) **Assess the method(s) used for reliability testing**

   Submission document: Testing attachment, section 2a2.2

7) **Assess the results of reliability testing**

   Submission document: Testing attachment, section 2a2.3

8) Was the method described and appropriate for assessing the proportion of variability due to real differences among measured entities? NOTE: If multiple methods used, at least one must be appropriate.

   Submission document: Testing attachment, section 2a2.2

   ☐ **Yes**

☐ **No**

☒ **Not applicable** (score-level testing was not performed)

9) Was the method described and appropriate for assessing the reliability of ALL critical data elements?

Submission document: Testing attachment, section 2a2.2

☐ **Yes**

☐ **No**

☒ **Not applicable** (data element testing was not performed)

10) **OVERALL RATING OF RELIABILITY** (taking into account precision of specifications and all testing results):

☐ **High** (NOTE: Can be HIGH only if score-level testing has been conducted)

☒ **Moderate** (NOTE: Moderate is the highest eligible rating if score-level testing has not been conducted)

☐ **Low** (NOTE: Should rate LOW if you believe specifications are NOT precise, unambiguous, and complete or if testing methods/results are not adequate)

☐ **Insufficient** (NOTE: Should rate INSUFFICIENT if you believe you do not have the information you need to make a rating decision)

11) **Briefly explain rationale for the rating of OVERALL RATING OF RELIABILITY and any concerns you may have with the approach to demonstrating reliability.**

- In this case, the validity rating may also apply to the reliability rating.

VALIDITY: ASSESSMENT OF THREATS TO VALIDITY

12) **Please describe any concerns you have with measure exclusions.**

Submission document: Testing attachment, section 2b2.

- Select patient care locations in hospitals (e.g., psychiatric wars, OB/GYN units) are excluded.
- No concerns.

13) Please describe any concerns you have regarding the ability to identify meaningful differences in performance.

Submission document: Testing attachment, section 2b4.

- Meaningful differences are analyzed by identifying the percent of facilities in which SAARs are statistically lower or higher than 1. (Attachment, Table 3). The distribution shows many facilities SAARs are statistically higher and lower than 1 for all antimicrobial groupings and patient care locations.

14) Please describe any concerns you have regarding comparability of results if multiple data sources or methods are specified.

Submission document: Testing attachment, section 2b5.

N/A

15) Please describe any concerns you have regarding missing data.

Submission document: Testing attachment, section 2b6.

- None.

16) Risk Adjustment

16a. **Risk-adjustment method** ☐ **None** ☒ **Statistical model** ☐ **Stratification**

16b. **If not risk-adjusted, is this supported by either a conceptual rationale or empirical analyses?**

☐ Yes ☐ No ☒ Not applicable

16c. **Social risk adjustment:**

16c.1 Are social risk factors included in risk model? ☐ Yes ☒ No ☐ Not applicable

16c.2 Conceptual rationale for social risk factors included? ☐ Yes ☒ No

16c.3 Is there a conceptual relationship between potential social risk factor variables and the measure focus? ☐ Yes ☒ No

16d.**Risk adjustment summary:**

16d.1 All of the risk-adjustment variables present at the start of care? ☒ Yes ☐ No

16d.2 If factors not present at the start of care, do you agree with the rationale provided for inclusion? ☐ Yes ☐ No

16d.3 Is the risk adjustment approach appropriately developed and assessed? ☒ Yes ☐ No

16d.4 Do analyses indicate acceptable results (e.g., acceptable discrimination and calibration) ☒ Yes ☐ No

16d.5.Appropriate risk-adjustment strategy included in the measure? ☒ Yes ☐ No

16e**. Assess the risk-adjustment approach**

- The data sample used for risk adjustment for the adult models, included 2156 patient care locations from 449 hospitals. For the pediatric models, data from 170 patient care locations from 109 hospitals was used.
- Predictive models were constructed using forward stagewise Negative Binomial regression assessing Wald and likelihood ratio Chi-square tests at a 0.05 significance level and model fit improvement using Akaike and Bayesian Information Criterion.
- Hospital-specific and patient care location-specific variables considered were hospital teaching status, hospital bedsize, hospital ICU bedsize, percentage of ICU beds among total beds, average length of hospital stay, and patient care location.
- Each group of SAAR antimicrobial agents is modeled separately. Models are different for adults and pediatrics. Bootstrap validation was conducted to validate model performance.
- All Likelihood Ratio Tests for the best models for each SAAR antimicrobial agent group indicated significant improvement as well as the lowest Akaike and/or Bayesian Information Criterion values.
- Social risk factors were not considered for adjustment due to lack of literature of a conceptual relationship between these factors and inpatient antimicrobial use.

VALIDITY: TESTING

17) Validity testing level: ☒ Measure score ☒ Data element ☐ Both

18) Method of establishing validity of the measure score:

☒ Face validity

☐ Empirical validity testing of the measure score

☐ N/A (score-level testing not conducted)

19) **Assess the method(s) for establishing validity**

Submission document: Testing attachment, section 2b2.2

- Testing at the data element level using a sample of 48 facilities that have reported antimicrobial use and days present to the CDC's NHSN Antimicrobial Use and Resistance (AUR) Module during 2016 through 2018.

- Hospitals compared antimicrobial use data in local eMAR/BCMA to the antimicrobial use data that were collected using the third party software vendor system or the healthcare system's corporate data warehouse. Local eMAR or BCMA were considered the gold standard. Hospitals and third party software vendors compared the numerator and denominator data in the third party software system pre- and post-aggregation. The gold standard in this process was the data obtained from eMAR, BCMA, and/or ADT systems.
- Data elements tested were antimicrobial administration from point of care systems (eMAR or BCMA) and patient care location data collected from ADT systems. Data elements used in the adjustment model were not tested.
- This testing builds upon previous data-element testing submitted from 2015, which included 13 hospitals that tested data collected and 24 hospitals that tested data aggregation.

- Face validity was established through an expert panel of infectious disease physicians and clinical pharmacists using a Delphi process. The panel reviewed the validity of core data elements and construct of antimicrobial groupings in 2018. (This was an update to previously established face validity of the core elements by an expert panel).

20) Assess the results(s) for establishing validity

Submission document: Testing attachment, section 2b2.3

- Antimicrobial days numerator: percent agreement 60-80% (at the outset of validation)
- Days present denominator: percent agreement 70-80% (at the outset of validation)
- By design the process led to >99% agreement for all required data elements prior to data submission to CDC.

21) Was the method described and appropriate for assessing conceptually and theoretically sound hypothesized relationships?

Submission document: Testing attachment, section 2b1.

☒ **Yes**

☐ **No**

☐ **Not applicable** (score-level testing was not performed)

22) **Was the method described and appropriate for assessing the accuracy of ALL critical data elements?**

NOTE that data element validation from the literature is acceptable.

Submission document*: Testing attachment, section 2b1.*

☒ **Yes**

☐ **No**

☐ **Not applicable** (data element testing was not performed)

23) OVERALL RATING OF VALIDITY taking into account the results and scope of all testing and analysis of potential threats.

☐ **High** (NOTE: Can be HIGH only if score-level testing has been conducted)

☒ **Moderate** (NOTE: Moderate is the highest eligible rating if score-level testing has NOT been conducted)

☐ **Low** (NOTE: Should rate LOW if you believe that there <u>are</u> threats to validity and/or relevant threats to validity were <u>not assessed OR</u> if testing methods/results are not adequate)

☐ **Insufficient** (NOTE: For instrument-based measures and some composite measures, testing at both the score level and the data element level is required; if not conducted, should rate as INSUFFICIENT.)

24) **Briefly explain rationale for rating of OVERALL RATING OF VALIDITY and any concerns you may have with the developers' approach to demonstrating validity.**

- Face validity testing was conducted.
- Data element validity testing was conducted, and results indicated >99% agreement for both the numerator and denominator prior to data submission to the CDC.
- Is there any concern about the numerator agreement of 60-80% or denominator agreement of 70-80% at the outset of the validation process?

ADDITIONAL RECOMMENDATIONS

25) If you have listed any concerns in this form, do you believe these concerns warrant further discussion by the multi-stakeholder Standing Committee? If so, please list those concerns below.

Committee Pre-evaluation Comments:
Criteria 2: Scientific Acceptability of Measure Properties (including all 2a, 2b, and 2c)

2a1. Reliability – Specifications

Comments:

**Data element specifications are well defined and should be highly reliable once mapped to the EHR in each facility. The validity of the specifications are moderate to high.

**No concerns

**No concerns

**acceptable; completely dependent upon technology/data feeds

2a2. Reliability – Testing

Comments:

**Not required per NQF requirements as data element validity was demonstrated.

**No concerns

**No

**no; given the limitations of these data are reliable for sites who have the technology/infrastructure to capture all the requisite elements

2b1. Validity –Testing

Comments:

**Testing not done

**none

**no

**Yes, external validity is my concern; these data may not represent the universe of hospitals; this is a very select subset/"sample" even in this sample there are sites who may only have ever submitted one month of data

2b4-7. Threats to Validity

2b4. Meaningful Differences

Comments:

**I have concerns over what the differences from 1 actually mean for an institution. As ~90 of institutions either were statistically below or above 1, what does this mean? How do we interpret the score in relationship to quality?

**No concerns

**No concerns

**The construction of the measure eliminates all hospitals who do not have the technology required to capture the requisite elements; there is no proxy for sites who could report DOT by location and antibx type using administrative data

2b2-3. Other Threats to Validity
2b2. Exclusions
2b3. Risk Adjustment

Comments:

**No concerns in this area

**appropriate risk adj

**Risk adjustment strategy is appropriate

**just important to keep in mind the "risk adjustment/stratification" is based on organizational characteristic - not pt variables. That is perfectly fine; we need to get our arms around "use"..you manage what you measure

Criterion 3. Feasibility

Maintenance measures – no change in emphasis – implementation issues may be more prominent

3. Feasibility is the 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.

- All data elements are in defined fields in electronic sources and routinely generated as the byproduct of electronic medication administration record keeping at the point of care.
- Data collection of the electronic components has proven feasible across hospitals.
- Upfront technical costs and implementation challenges are the main feasibility considerations. Technical assistance is provided by the CDC.
- There are no fees associated with this measure.

Questions for the Committee:

- Is the data collection strategy ready to be widely operationalized?
- Is the measure logic understandable and can the measure be calculated without undue burden?

Preliminary rating for feasibility: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

RATIONALE:

Committee Pre-evaluation Comments:
Criteria 3: Feasibility

3. Feasibility

Comments:

**Organizations require a considerable investment to accurately provide the required data elements the first time; however once employed, the reporting is good

**No concerns

**No concerns

**Remain concerned with the discounting of any possibility of a proxy using more readily available administrative data . Agree this measure would be a "gold standard" and encouraged/incentivized..however, missing a lot of "good enough" and "good to know" by ignoring proxies.

Criterion 4: Usability and Use

Maintenance measures – increased emphasis – much greater focus on measure use and usefulness, including both impact/improvement and unintended consequences

4a. Use (4a1. Accountability and Transparency; 4a2. Feedback on measure)

4a. Use evaluate the extent to which audiences (e.g., consumers, purchasers, providers, policymakers) use or could use performance results for both accountability and performance improvement activities.

4a.1. 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.

Current uses of the measure

Publicly reported? ☐ **Yes** ☒ **No**

Current use in an accountability program? ☐ **Yes** ☒ **No** ☐ **UNCLEAR**

OR

Planned use in an accountability program? ☒ **Yes** ☐ **No**

Accountability program details

- National Healthcare Safety Network (NHSN), Centers for Disease Control and Prevention
    - Used for public health/disease surveillance, internal quality improvement, and quality improvement with benchmarking.
    - In almost all states, at least some hospitals (range 4-49% across states) are reporting data to the NHSN and gaining access to benchmark data.
- Voluntary reporting has increased to over 1000 hospitals nationwide (>5 fold increase since endorsement). Increase reporting has allowed for the refinement of predictive models.
- Lessons learned through increasing use of the measure coupled with further development of the predictive models will enable use for accountability purposes. NHSN seeks to add data about infectious disease burden and use of antimicrobials for prophylaxis to the predictive models, which will allow for improved value for assess intra- and inter-organization variation.

4a.2. Feedback on the measure by those being measured or others. Three criteria demonstrate feedback: 1) those being measured have been given performance results or data, as well as assistance with interpreting the measure results and data; 2) those being measured and other users have been given an opportunity to provide feedback on the measure performance or implementation; 3) this feedback has been considered when changes are incorporated into the measure

Feedback on the measure by those being measured or others

- All hospitals participating in NHSN AU surveillance have access to submitted data and tools for analyzing their data. Feedback from reporting hospitals has been positive.

Additional Feedback:

N/A

***Questions for the Committee**:*

- The measure is not currently used for accountability or publicly reported (though an increasing number of hospitals across the nation are reporting data to NHSN). Is the developer's plan for use in an accountability program acceptable?
- Can the performance results be used to further the goal of appropriate antibiotic use and improved patient outcomes?

Preliminary rating for Use: ☒ **Pass** ☐ **No Pass**

RATIONALE:

4b. Usability (4a1. Improvement; 4a2. Benefits of measure)

4b. Usability evaluate the extent to which audiences (e.g., consumers, purchasers, providers, policymakers) use or could use performance results for both accountability and performance improvement activities.

4b.1 Improvement. Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated.

Improvement results

- Reporting hospitals have used their SAARs to inform priorities for antimicrobial stewardship.

4b2. Benefits vs. harms. 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).

Unexpected findings (positive or negative) during implementation

- Benefits include: Better understanding of antimicrobial use data and how to use EHR or 3rd party vendor software responsible for collecting/reporting AU data, improvement in hospital workflow based on data quality issues (e.g., changes in order sets to fix free text entry; allowing off label administrations to be entered into EHR correctly), and improvement in the quality of data captured (e.g., correcting errors in how routes are mapped).

Potential harms

- None identified.

Additional Feedback:

N/A

***Questions for the Committee**:*

- Have there been improvements in the appropriate use of antimicrobials based on the implementation of this measure?
- Do the benefits of the measure outweigh any potential unintended consequences?

Preliminary rating for Usability and use: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

RATIONALE:

Committee Pre-evaluation Comments:
Criteria 4: Usability and Use

4a1. Use - Accountability and Transparency

Comments:

**The measure is rapidly being adopted by many institutions as a means to identify where to dive deeper into antibiotic prescibing patterns and behaviors. The measure is not currently in use in performance based programs and I have a difficult time understanding how it could be used in a performance program.

**not curently publicly reported - voluntary , would liketo see public reporting now that measure has been refined

**Not currently used in accountability program, but CMS plans to consider this. Feedback has been considered when changes have been made

**Has great potential..however, right now we understand hospitals are using these data for internal PI efforts and they are (and rightly so) not being used in accountability. My concern is CMS will prematurely introduce into payment w/o accounting for the vast majority of hospital who cannot report the elements to NHSN

4b1. Usability – Improvement

Comments:

**The measure is a high-level assessment intended to highlight and help facilities to know where to look deeper for areas of improvement. Given there is nothing else, an imperfect measure is preferable to nothing. The only harm I see is the mis-use of the measure into a performance program when the results of the benchmark observed vs expected do not clearly relate to quality.

**No concerns

**I believe this measure is still usable to improve quality

**These data can be invaluable...however need to radically expand reporting sites and months of reporting to really understand what the data are telling us

Criterion 5: Related and Competing Measures

Related or competing measures
None identified.

Harmonization
N/A

Committee Pre-evaluation Comments: Criterion 5: Related and Competing Measures

5. Related and Competing

Comments:

**There are currently no related or competing measures

**None

**No concerns at this time

**No

Public and Member Comments

Comments and Member Support/Non-Support Submitted as of: 6/5/2019

- **No NQF Members have submitted support/non-support choices as of this date.**

Brief Measure Information

NQF #: 2720

Corresponding Measures:

De.2. Measure Title: National Healthcare Safety Network (NHSN) Antimicrobial Use Measure

Co.1.1. Measure Steward: Centers for Disease Control and Prevention

De.3. Brief Description of Measure: This measure assesses antimicrobial use in hospitals based on medication administration data that hospitals collect electronically at the point of care and report via electronic file submissions to CDC's National Healthcare Safety Network (NHSN). The antimicrobial use data that are in scope for this measure are antibacterial agents administered to adult and pediatric patients in a specified set of ward and intensive care unit locations: medical ICU, medical-surgical ICU, surgical ICU (adult only), medical ward, medical-surgical ward, surgical ward, general hematology-oncology ward (adult only), and step-down unit (adult only). The measure compares antimicrobial use that the hospitals report with antimicrobial use that is predicted on the basis of nationally aggregated data. The measure is comprised of a discrete set of ratios, Standardized Antimicrobial Administration Ratios (SAARs), each of which summarizes observed-to-predicted antimicrobial use for one of 40 antimicrobial agent-patient care location combinations. The SAARs are designed to serve as high value targets or high level indicators for antimicrobial stewardship programs (ASPs). SAAR values that are outliers are intended to prompt analysis of possible overuse, underuse, or inappropriate use of antimicrobials, subsequent actions aimed at improving the quality of antimicrobial prescribing, and impact evaluations of ASP interventions.

1b.1. Developer Rationale: The measure provides summary results that hospital and health system antimicrobial stewardship programs (ASPs) can use as quantitative aids in their efforts to evaluate and improve antibiotic prescribing. The Standardized Antimicrobial Administration Ratios (SAARs) that comprise the measure focus on high value targets and high level indicators of antimicrobial use for ASPs. The SAARs can be used by ASPs to benchmark antimicrobial use in multiple patient care locations, identify opportunities for improvement, and gauge the impact of stewardship efforts. At the outset, the SAARs provide a set of signals that often warrant further analysis, such as an evaluation of the extent to which a specific antibiotic or group of antibiotics accounts for a high or low SAAR value and the extent to which an antibiotic or group of antibiotics were used appropriately for specific indications. While the SAARs do not provide a definitive indication that antibiotics are overused or underused, they provide an important starting place for further analysis and possible action. Some of the analytic follow up can be completed with hospital- and patient care location-specific data reported to CDC's National Healthcare Safety Network (NHSN) Antimicrobial Use and Resistance (AUR) Module, using analytic features built into the NHSN application. However, additional analyses to determine the appropriateness of antibiotic use in individual instances are likely to require access to detailed, patient-level data that is beyond the scope of data collection and analysis using the NHSN module, e.g., clinical indications for specific antibiotics and dose and duration decisions.

S.4. Numerator Statement: Days of antimicrobial therapy for antimicrobial agents administered to adult and pediatric patients in medical ICU, medical-surgical ICU, surgical ICU (adult only), medical ward, medical-surgical ward, surgical ward, general hematology-oncology ward (adult only), and step-down unit (adult only).

S.6. Denominator Statement: Days present for each patient care location—adult and pediatric medical ICU, medical-surgical ICU, surgical ICU (adult only), medical ward, medical-surgical ward, surgical ward, general hematology-oncology ward (adult only), and step-down unit (adult only) is defined as the number of patients who were present for any portion of each day of a calendar month for each location. The day of admission, discharge, and transfer to and from locations are included in days present. All days present are summed for

each location and month, and the aggregate sums for each location-month combination comprise the denominator data for the measure.

S.8. Denominator Exclusions: Hospital patient care locations other than adult and pediatric medical ICU, medical-surgical ICU, surgical ICU (adult only), medical ward, medical-surgical ward, surgical ward, general hematology-oncology ward (adult only), and step-down unit (adult only) are excluded from this measure.

De.1. Measure Type: Process

S.17. Data Source: Electronic health data; electronic format Admission Discharge Transfer

S.20. Level of Analysis: Facility

IF Endorsement Maintenance – Original Endorsement Date: Dec 10, 2015 **Most Recent Endorsement Date:** Dec 10, 2015

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?

1. Evidence and Performance Gap – 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_evidence_attachment-AU_review_final.docx

1a.1 <u>For Maintenance of Endorsement:</u> 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.

Yes

1a. Evidence (subcriterion 1a)

NATIONAL QUALITY FORUM—Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:** 2720

Measure Title: National Healthcare Safety Network (NHSN) Antimicrobial Use Measure

IF the measure is a component in a composite performance measure, provide the title of the Composite Measure here: Click here to enter composite measure #/ title

Date of Submission: Click here to enter a date

Instructions

- ***Complete 1a.1 and 1a.2 for all measures. If instrument-based measure, complete 1a.3.***
- ***Complete EITHER 1a.2, 1a.3 or 1a.4 as applicable for the type of measure and evidence.***
- ***For composite performance measures:***

- ***A separate evidence form is required for each component measure unless several components were studied together.***
- ***If a component measure is submitted as an individual performance measure, attach the evidence form to the individual measure submission.***

- **All information needed to demonstrate meeting the evidence subcriterion (1a) must be in this form. An appendix of *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.**
- **If you are unable to check a box, please highlight or shade the box for your response.**
- **Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.**

Note: The information provided in this form is intended to aid the Standing Committee and other stakeholders in understanding to what degree the evidence for this measure meets NQF's evaluation criteria.

1a. Evidence to Support the Measure Focus

The measure focus is evidence-based, demonstrated as follows:

- **Outcome: [3] Empirical data demonstrate a relationship between the outcome and at least one healthcare structure, process, intervention, or service. If not available, wide variation in performance can be used as evidence, assuming the data are from a robust number of providers and results are not subject to systematic bias.**
- **Intermediate clinical outcome: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured intermediate clinical outcome leads to a desired health outcome.**
- **Process: [5] a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured process leads to a desired health outcome.**
- **Structure: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured structure leads to a desired health outcome.**
- **Efficiency: [6] evidence not required for the resource use component.**
- **For measures derived from patient reports, evidence should demonstrate that the target population values the measured outcome, process, or structure and finds it meaningful.**
- **Process measures incorporating Appropriate Use Criteria: See NQF's guidance for evidence for measures, in general; guidance for measures specifically based on clinical practice guidelines apply as well.**

Notes

3. Generally, rare event outcomes do not provide adequate information for improvement or discrimination; however, serious reportable events that are compared to zero are appropriate outcomes for public reporting and quality improvement.

4. The preferred systems for grading the evidence are the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) guidelines and/or modified GRADE.

5. Clinical care processes typically include multiple steps: assess → identify problem/potential problem → choose/plan intervention (with patient input) → provide intervention → evaluate impact on health status. If the measure focus is one step in such a multistep process, the step with the strongest evidence for the link to the desired outcome should be selected as the focus of measurement. Note: A measure focused only on collecting PROM data is not a PRO-PM.

6. Measures of efficiency combine the concepts of resource use and quality (see NQF's Measurement Framework: Evaluating Efficiency Across Episodes of Care; AQA Principles of Efficiency Measures).

1a.1.This is a measure of: (*should be consistent with type of measure entered in De.1*)

Outcome

☐ Outcome: Click here to name the health outcome

☐ Patient-reported outcome (PRO): Click here to name the PRO

PROs include HRQoL/functional status, symptom/symptom burden, experience with care, health-related behaviors. (A PRO-based performance measure is not a survey instrument. Data may be collected using a survey instrument to construct a PRO measure.)

☐ Intermediate clinical outcome (*e.g., lab value*): Click here to name the intermediate outcome

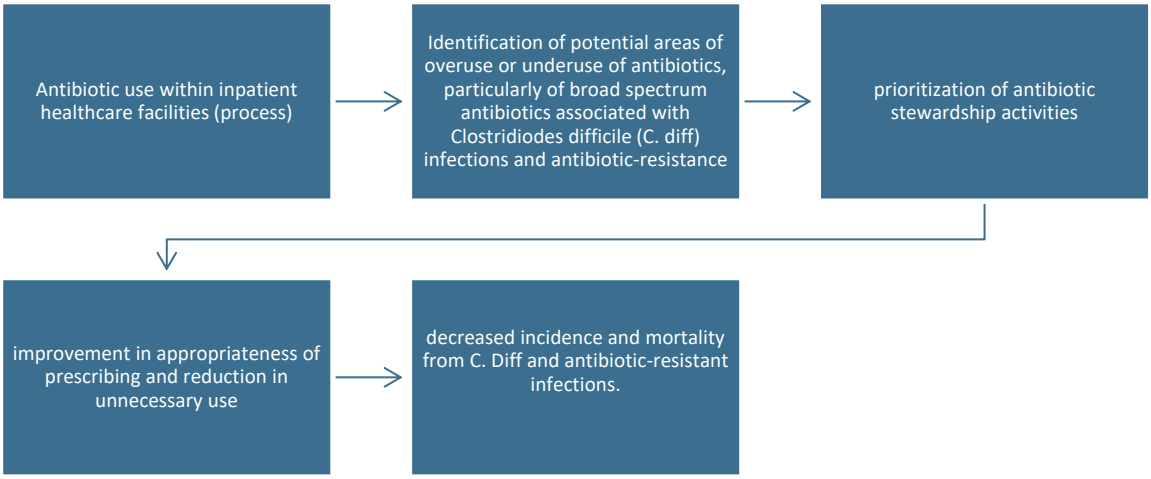
☒ Process: Click here to name what is being measured

☐ Appropriate use measure: Click here to name what is being measured

☐ Structure: Click here to name the structure

☐ Composite: Click here to name what is being measured

1a.2 LOGIC MODEL Diagram or briefly describe the steps between the healthcare structures and processes (e.g., interventions, or services) and the patient's health outcome(s). The relationships in the diagram should be easily understood by general, non-technical audiences. Indicate the structure, process or outcome being measured.



Antibiotic use within inpatient healthcare facilities (process) >> Identification of potential areas of overuse or underuse of antibiotics, particularly of broad-spectrum antibiotics associated with Clostridioides difficile (C. diff) infections and antibiotic resistance >> Prioritization of antibiotic stewardship activities >> Improvement in appropriateness of prescribing and reduction in unnecessary use >> Decreased incidence and mortality from C. Diff and antibiotic-resistant infections.

1a.3 Value and Meaningfulness: IF this measure is derived from patient report, provide evidence that the target population values the measured ***outcome, process, or structure*** and finds it meaningful. (Describe how and from whom their input was obtained.)

****RESPOND TO ONLY ONE SECTION BELOW -EITHER 1a.2, 1a.3 or 1a.4) ****

1a.2 FOR OUTCOME MEASURES including PATIENT REPORTED OUTCOMES - Provide empirical data demonstrating the relationship between the outcome (or PRO) to at least one healthcare structure, process, intervention, or service.

1a.3. SYSTEMATIC REVIEW(SR) OF THE EVIDENCE (for INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURES, INCLUDING THOSE THAT ARE INSTRUMENT-BASED) If the evidence is not based on a systematic review go to section 1a.4) If you wish to include more than one systematic review, add additional tables.

What is the source of the systematic review of the body of evidence that supports the performance measure? A systematic review is a scientific investigation that focuses on a specific question and uses explicit, prespecified scientific methods to identify, select, assess, and summarize the findings of similar but separate studies. It may include a quantitative synthesis (meta-analysis), depending on the available data. (IOM)

X Clinical Practice Guideline recommendation (with evidence review)

☐ US Preventive Services Task Force Recommendation

X Other systematic review and grading of the body of evidence (*e.g., Cochrane Collaboration, AHRQ Evidence Practice Center*)

☐ Other

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	**This review is an update from the 2006 and 2013 versions previously cited on our submissions. See section 1.4 below for previously submitted evidence.** **Interventions to improve antibiotic prescribing practices for hospital inpatients.** **Davey P, Marwick CA, Scott CL, Charani E, McNeil K, Brown E, Gould IM, Ramsay CR, Michie S.** **Cochrane Database of Systematic Reviews 2017, Issue 2. Art. No.: CD003543. DOI: 10.1002/14651858.CD003543.pub4** **https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD003543.pub4/epdf/full**
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	The authors found high-certainty evidence that interventions such as enablement (defined as 'increasing means/reducing barriers to increase capability or opportunity'), and restriction (defined as 'using rules to reduce the opportunity to engage in the target behavior'), are effective in increasing compliance with antibiotic policy and reducing duration of antibiotic treatment. Lower use of antibiotics has not been associated with a difference in mortality and likely reduces length of stay. Additional trials comparing antibiotic stewardship with no intervention are unlikely to change the authors' conclusions. Both enablement and restriction were

	independently associated with increased compliance with antibiotic policies, and enablement consistently increased the effect of restrictive interventions. Interventions were successful in safely reducing unnecessary antibiotic use in hospitals without adversely affecting mortality and probably reducing length of stay. Consequently, effective dissemination of these findings could have considerable health service and policy impact. Future research should instead focus on targeting treatment and assessing other measures of patient safety, assess different stewardship interventions, and explore the barriers and facilitators to implementation. More research is required on unintended consequences of restrictive interventions.
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	**Effectiveness: prescribing outcomes from randomized controlled trials (RCTs)** High certainty: Proportion of participants who were treated according to antibiotic prescribing guidelines High certainty: Duration of all antibiotic treatment Moderate certainty: Mortality Moderate certainty: Mean length of hospital stay per patient Low certainty: Delay in treatment Low certainty: Negative professional culture **Effect modifiers: impact of behavior change functions and additional impact of feedback, RCTs and interrupted time series (ITS) studies** High certainty: Enablement High certainty: Restriction Moderate certainty: Addition of feedback to enablement Low certainty: Addition of enablement to restriction
Provide all other grades and definitions from the evidence grading system	GRADE Working Group grades of evidence **High certainty:** Further research is very unlikely to change our confidence in the estimate of effect. **Moderate certainty:** Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate. **Low certainty:** Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

	Very low certainty: We are very uncertain about the estimate
Grade assigned to the **recommendation** with definition of the grade	There was high-certainty evidence from 14 RCTs with 3318 participants to show that the duration of antibiotic treatment decreased by 1.95 days (95% CI 2.22 to 1.67) Information from non-randomized studies showed interventions to be associated with improvement in prescribing according to antibiotic policy in routine clinical practice, with 70% of interventions being hospital-wide compared with 31% for RCTs. The risk of death was similar between intervention and control groups (11% in both arms), indicating that antibiotic use can likely be reduced without adversely affecting mortality (RD 0%, 95% CI -1% to 0%; 28 RCTs; 15,827 participants; moderate-certainty evidence). Antibiotic stewardship interventions probably reduce length of stay by 1.12 days (95% CI 0.7 to 1.54 days; 15 RCTs; 3834 participants; moderate-certainty evidence). Both enablement and restriction were independently associated with increased compliance with antibiotic policies, and enablement enhanced the effect of restrictive interventions (high-certainty evidence). Enabling interventions that included feedback were probably more effective than those that did not (moderate-certainty evidence). There was very low-certainty evidence about the effect of the interventions on reducing *Clostridioides* (formerly *Clostridium*) *difficile* infections (median -48.6%, interquartile range -80.7% to -19.2%; 7 studies). This was also the case for resistant gram-negative bacteria (median -12.9%, interquartile range -35.3% to 25.2%; 11 studies) and resistant gram-positive bacteria (median -19.3%, interquartile range -50.1% to +23.1%; 9 studies). There was too much variance in microbial outcomes to reliably assess the effect of change in antibiotic use.
Provide all other grades and definitions from the recommendation grading system	GRADE Working Group grades of evidence **High certainty:** Further research is very unlikely to change our confidence in the estimate of effect. **Moderate certainty:** Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate. **Low certainty:** Further research is very likely to have an important impact on our confidence in the estimate

	of effect and is likely to change the estimate. **Very low certainty:** We are very uncertain about the estimate
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	This review includes 221 relevant studies: 58 randomized controlled trials (RCT) and 163 non-randomized studies (NRS). The authors included three non-randomized study designs to measure behavioral and clinical outcomes and analyze variation in the effects: non- randomized trials (NRT), controlled before-after (CBA) studies and interrupted time series (ITS) studies. For this update they also included three additional NRS designs (case control, cohort, and qualitative studies) to identify unintended consequences.
Estimates of benefit and consistency across studies	Heterogeneity was quantified among the studies using the I^2 statistic and Cochran's Q test. The I^2 statistic quantifies the percentage of the total variation across studies that is due to heterogeneity rather than chance; smaller percentages suggest less observed heterogeneity. **Outcome** prescribing, compliance with guideline Inconsistency, I^2=92%; not serious, effect size rather than direction. Direction of effect consistent despite high levels of heterogeneity. **Outcome** prescribing, reduction in duration of antibiotic treatment Inconsistency, I^2=89%; Not serious, most variation ins effect size rather than direction **Outcome** mortality Inconsistency, I^2=0%; not serious **Outcome** length of hospital stay Inconsistency, I^2=19%; not serious, effect size rather than direction
What harms were identified?	One RCT and six NRS raised concerns that restrictive interventions may lead to delay in treatment and negative professional culture because of breakdown in communication and trust between infection specialists and clinical teams (low-certainty evidence).
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	• **Implementing an Antibiotic Stewardship Program: Guidelines by the Infectious Diseases Society of America and the Society for Healthcare Epidemiology of America** • **Barlam TF, Cosgrove SE, Abbo LM, MacDougall C, Schuetz AN, Septimus EJ, Srinivasan A, Dellit TH, Falck-Ytter YT, Fishman NO, Hamilton CW, Jenkins TC, Lipsett PA, Malani PN, May LS, Moran GJ, Neuhauser MM, Newland JG, Ohl CA, Samore MH, Seo SK, Trivedi KK.** • **2016** • **Barlam TF, Cosgrove SE, Abbo LM, MacDougall C, Schuetz AN, Septimus EJ, Srinivasan A, Dellit TH, Falck-Ytter YT, Fishman NO, Hamilton CW, Jenkins TC, Lipsett PA, Malani PN, May LS, Moran GJ, Neuhauser MM, Newland JG, Ohl CA, Samore MH, Seo SK, Trivedi KK. Executive Summary: Implementing an Antibiotic Stewardship Program: Guidelines by the Infectious Diseases Society of America and the Society for Healthcare Epidemiology of America. Clin Infect Dis. 2016 May 15;62(10):e51-e77.** • **https://academic.oup.com/cid/article/62/10/e51/2462846**
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	I. We recommend preauthorization and/or prospective audit and feedback over no such interventions. Pre-authorization and/or prospective audit and feedback (PAF) improve antibiotic use and are a core component of any stewardship program. Preauthorization has been associated with a significant reduction in the use of the restricted agents and of associated costs. Outcome studies with preauthorization have shown decreased antibiotic use and decreased antibiotic resistance, particularly among gram-negative pathogens. While implementing a preauthorization program, the skills of the person providing the approval are important. Antibiotic approval by an antibiotic stewardship team consisting of a clinical pharmacist and an infectious disease attending physician is more effective than an off-hour approval by infectious disease fellows in recommendation appropriateness, cure rate and treatment failures. PAF have also shown to improve antibiotic resistance, and reduce *Clostridioides* (formerly *Clostridium*) *difficile* infections (CDI) rates, without a negative impact on patient outcomes. PAF can be very labor intensive, and identification of appropriate patients for intervention can be challenging and require computerized surveillance systems.

	PAF effectiveness may depend on the infrastructure in place at an institution. Programs should decide whether to include one strategy or a combination of both strategies based on the availability of facility-specific resources for consistent implementation, but some implementation is essential.
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	I. Moderate-quality evidence
Provide all other grades and definitions from the evidence grading system	Working Group grades of evidence **High certainty:** Further research is very unlikely to change our confidence in the estimate of effect. **Moderate certainty:** Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate. **Low certainty:** Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate. **Very low certainty:** We are very uncertain about the estimate
Grade assigned to the **recommendation** with definition of the grade	I. Strong recommendation
Provide all other grades and definitions from the recommendation grading system	1. Strong recommendation - Population: Most people in this situation would want the recommended course of action and only a small proportion would not - Healthcare workers : Most people should receive the recommended course of action - Policy makers: The recommendation can be adapted as a policy in most situations 2. Weak recommendation - Population: The majority of people in this situation would want the recommended course of actions, but many would not - Healthcare workers: Be prepared to help people to make a decision that is consistent with their own values/decisions aids and shared decision making. - Policy makers: There is a need for substantial debate and involvement of stakeholders
Body of evidence: • Quantity – how many studies?	A total of 23 studies were reviewed for this evidence. A majority of the studies included quasi-experimental study designs, quasi-

• Quality – what type of studies?	experimental interrupted time series analyses and two systematic reviews with meta-analyses.
Estimates of benefit and consistency across studies	Across the studies reviewed that provided the body of evidence, there were consistent findings that preauthorization was associated with significant reduction in the use of the restricted agents and of associated costs. Outcome studies with preauthorization consistently demonstrated decreased antibiotic use and decreased antibiotic resistance particularly among gram-negative pathogens. The benefit of preauthorization compared with PAF has had limited study. Restrictive measures such as preauthorization were compared with persuasive measures such as PAF, and authors concluded that restrictive interventions were preferred when the need is urgent.
What harms were identified?	None identified
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	• **Implementing an Antibiotic Stewardship Program: Guidelines by the Infectious Diseases Society of America and the Society for Healthcare Epidemiology of America** • **Barlam TF, Cosgrove SE, Abbo LM, MacDougall C, Schuetz AN, Septimus EJ, Srinivasan A, Dellit TH, Falck-Ytter YT, Fishman NO, Hamilton CW, Jenkins TC, Lipsett PA, Malani PN, May LS, Moran GJ, Neuhauser MM, Newland JG, Ohl CA, Samore MH, Seo SK, Trivedi KK.** • **2016** • **Barlam TF, Cosgrove SE, Abbo LM, MacDougall C, Schuetz AN, Septimus EJ, Srinivasan A, Dellit TH, Falck-Ytter YT, Fishman NO, Hamilton CW, Jenkins TC, Lipsett PA, Malani PN, May LS, Moran GJ, Neuhauser MM, Newland JG, Ohl CA, Samore MH, Seo SK, Trivedi KK. Executive Summary: Implementing an Antibiotic Stewardship Program: Guidelines by the Infectious Diseases Society of America and the Society for Healthcare Epidemiology of America. Clin Infect Dis. 2016 May 15;62(10):e51-e77.** • **https://academic.oup.com/cid/article/62/10/e51/2462846**
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a	V. We recommend antibiotic stewardship interventions designed to reduce the use of antibiotics associated with a high risk of CDI compared with no such intervention.

guideline, summarize the conclusions from the SR.	The goal of reducing CDI is a high priority for all antibiotic stewardship programs (ASPs). Primary ASP interventions included restrictions of high-risk antibiotics and/or broad-spectrum antibiotics and both these methods demonstrated associated decrease in antibiotic use , decreased CDI , increase antibiotic susceptibility, and overall cost savings attributable to fewer cases of CDI. Implementation of ASPs have been associated with statistically significant sudden or linear-trend decreases in nosocomial CDI rates, which have been sustained for up to 7 years.
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	V. Moderate-quality evidence
Provide all other grades and definitions from the evidence grading system	Working Group grades of evidence **High certainty:** Further research is very unlikely to change our confidence in the estimate of effect. **Moderate certainty:** Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate. **Low certainty:** Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate. **Very low certainty:** We are very uncertain about the estimate
Grade assigned to the **recommendation** with definition of the grade	Strong recommendation
Provide all other grades and definitions from the recommendation grading system	Strong recommendation - Population: Most people in this situation would want the recommended course of action and only a small proportion would not - Healthcare workers : Most people should receive the recommended course of action - Policy makers: The recommendation can be adapted as a policy in most situations Weak recommendation - Population: The majority of people in this situation would want the recommended course of actions, but many would not - Healthcare workers: Be prepared to help people to make a decision that is consistent with their own values/decisions aids and shared decision making. - Policy makers: There is a need for substantial debate and involvement of stakeholders

Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	The recommendation and evidence was supported by 9 studies. Most of these studies (n=8) were quasi-experimental study designs, quasi-experimental interrupted time series analyses and one investigation was a case-control study design.
Estimates of benefit and consistency across studies	All investigations consistently found statistically significant associations between ASP and decreased incidence of CDI. Benefits of ASP included decline in use of restrictive and broad-spectrum antibiotics, increased antibiotic susceptibility and overall cost savings attributable to fewer cases of CDI.
What harms were identified?	None identified
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	• **Implementing an Antibiotic Stewardship Program: Guidelines by the Infectious Diseases Society of America and the Society for Healthcare Epidemiology of America** • **Barlam TF, Cosgrove SE, Abbo LM, MacDougall C, Schuetz AN, Septimus EJ, Srinivasan A, Dellit TH, Falck-Ytter YT, Fishman NO, Hamilton CW, Jenkins TC, Lipsett PA, Malani PN, May LS, Moran GJ, Neuhauser MM, Newland JG, Ohl CA, Samore MH, Seo SK, Trivedi KK.** • **2016** • **Barlam TF, Cosgrove SE, Abbo LM, MacDougall C, Schuetz AN, Septimus EJ, Srinivasan A, Dellit TH, Falck-Ytter YT, Fishman NO, Hamilton CW, Jenkins TC, Lipsett PA, Malani PN, May LS, Moran GJ, Neuhauser MM, Newland JG, Ohl CA, Samore MH, Seo SK, Trivedi KK. Executive Summary: Implementing an Antibiotic Stewardship Program: Guidelines by the Infectious Diseases Society of America and the Society for Healthcare Epidemiology of America. Clin Infect Dis. 2016 May 15;62(10):e51-e77.** • **https://academic.oup.com/cid/article/62/10/e51/2462846**
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	XX. 21. We suggest monitoring antibiotic use as measured by days of therapy (DOTs) in preference to defined daily dose (DDD) . DOTs and DDD are standardized methods for measurement of antibiotic use and both are useful at facility-level monitoring and interfacility comparisons. DOTs have been recommended as a robust measure since the metric is not impacted by dose

	adjustments and can be used both in adult and pediatric populations. DDDs have more limited use in pediatric population due to weight-based dosing.
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	Low-quality evidence
Provide all other grades and definitions from the evidence grading system	Working Group grades of evidence **High certainty:** Further research is very unlikely to change our confidence in the estimate of effect. **Moderate certainty:** Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate. **Low certainty:** Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate. **Very low certainty:** We are very uncertain about the estimate
Grade assigned to the **recommendation** with definition of the grade	Weak recommendation
Provide all other grades and definitions from the recommendation grading system	Strong recommendation - Population: Most people in this situation would want the recommended course of action and only a small proportion would not - Healthcare workers : Most people should receive the recommended course of action - Policy makers: The recommendation can be adapted as a policy in most situations Weak recommendation - Population: The majority of people in this situation would want the recommended course of actions, but many would not - Healthcare workers: Be prepared to help people to make a decision that is consistent with their own values/decisions aids and shared decision making. - Policy makers: There is a need for substantial debate and involvement of stakeholders
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	This recommendation was based on 5 studies. These included one quasi-experimental study, two cross-sectional studies, one systematic review and one systematic review with meta-analyses.

Estimates of benefit and consistency across studies	Current literature has identified a significant impact of ASP interventions in overall reduction of antibiotic prescription, increased antibiotic susceptibility in both gram-positive and gram-negative bacteria, and decreased CDI incidence rates, however reduction in CDI incidence are affected by practices besides antibiotic use such as compliance with infection control measures. Further, observed effects on resistance are unpredictable due to confounding variables which are difficult to adjust for, including many pathogen and host factors.
What harms were identified?	None identified.
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	

1a.4 OTHER SOURCE OF EVIDENCE

If source of evidence is NOT from a clinical practice guideline, USPSTF, or systematic review, please describe the evidence on which you are basing the performance measure.

1a.4.1 Briefly SYNTHESIZE the evidence that supports the measure. A list of references without a summary is not acceptable.

1a.4.2 What process was used to identify the evidence?

1a.4.3. Provide the citation(s) for the evidence.

PREVIOUSLY SUBMITTED EVIDENCE:

Guideline citation Dellit TH, Owens RC, McGowan JE, Jr., Gerding DN, Weinstein RA, Burke JP, et al. Infectious Diseases Society of America and the Society for Healthcare Epidemiology of America guidelines for developing an institutional program to enhance antimicrobial stewardship. Clin Infect Dis. 2007;44(2):159-77.

and **URL for guideline** (*if available online*): http://cid.oxfordjournals.org/content/44/2/159.full.pdf

Identify guideline recommendation number and/or page number and **quote verbatim, the specific guideline recommendation**.

"6. There are 2 core strategies, both proactive, that provide the foundation for an antimicrobial stewardship program. These strategies are not mutually exclusive.

A. Prospective audit with intervention and feedback. Prospective audit of antimicrobial use with direct interaction and feedback to the prescriber, performed by either an infectious diseases

physician or a clinical pharmacist with infectious diseases training, can result in reduced inappropriate use of antimicrobials (A-I).

B. Formulary restriction and preauthorization. Formulary restriction and preauthorization requirements can lead to immediate and significant reductions in antimicrobial use and cost (A-II)… In institutions that use preauthorization to limit the use of selected antimicrobials, monitoring overall trends in antimicrobial use is necessary to assess and respond to such shifts in use (B-III).

"9. Computer-based surveillance can facilitate good stewardship by more efficient targeting of antimicrobial interventions, tracking of antimicrobial resistance patterns, and identification of nosocomial infections and adverse drug events (B-II)."

"11. Both process measures (did the intervention result in the desired change in antimicrobial use?) and outcome measures (did the process implemented reduce or prevent resistance or other unintended consequences of antimicrobial use?) are useful in determining the impact of antimicrobial stewardship on antimicrobial use and resistance patterns (B-III)"

1a.4.3. Grade assigned to the quoted recommendation with definition of the grade:

6A: A-I; 6B1: A-II; 6B2: B-III

9: B-II

11. B-III

Infectious Diseases Society of America–United States Public Health Service grading system for ranking recommendations in clinical guidelines.

Strength of recommendation

A. Good evidence to support a recommendation for use
B. Moderate evidence to support a recommendation for use
C. Poor evidence to support a recommendation for use

Quality of evidence

I. Evidence from _1 properly randomized, controlled trial
II. Evidence from _1 well-designed clinical trial, without randomization; from cohort or case-controlled analytic studies (preferably from 11 center); from multiple time-series; or from dramatic results from uncontrolled experiments
III. Evidence from opinions of respected authorities, based on clinical experience, descriptive studies, or reports of expert committees

OTHER SYSTEMATIC REVIEW OF THE BODY OF EVIDENCE

1a.6.1a. Feazel LM, Malhotra A, Perencevich EN, Kaboli P, Diekema DJ, Schweizer ML**.** Effect of antibiotic stewardship programmes on Clostridium difficile incidence: a systematic review and meta-analysis. J Antimicrob Chemother. 2014;69(7):1748-54. http://jac.oxfordjournals.org/content/69/7/1748.full.pdf

1a.6.1b. Davey P, Brown E, Charani E, Fenelon L, Gould IM, Holmes A, et al. Interventions to improve antibiotic prescribing practices for hospital inpatients. Cochrane Database Syst Rev. 2013;4:CD003543. http://onlinelibrary.wiley.com/store/10.1002/14651858.CD003543.pub3/asset/CD003543.pdf?v=1&t=hvxzajv5&s=a6f3c724ce051d8acba5866a07e3c5ac8c818e83

1a.6.1c. Kaki R, Elligsen M, Walker S, Simor A, Palmay L, Daneman N. Impact of antimicrobial stewardship in critical care: a systematic review. J Antimicrob Chemother. 2011;66(6):1223-30. http://jac.oxfordjournals.org/content/66/6/1223.full.pdf

1a.6.1d. Davey P, Brown E, Fenelon L, Finch R, Gould I, Holmes A, et al. Systematic review of antimicrobial drug prescribing in hospitals. Emerg Infect Dis. 2006;12(2):211-6. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3373108/

1a.6.2. Citation and URL for methodology for evidence review and grading (*if different from 1a.6.1*)**:**

1a.6.1a. (Feazel, 2014)

Criteria for inclusion included the following types of observational design: experimental (e.g. randomized trials), quasi-experimental (e.g. interrupted time series and before–after), case–control and cohort studies. The natural log of the risk ratio and variance was calculated for each study. Both fixed effects and random effects models to obtain pooled risk ratio estimates were constructed using Microsoft Excel 2007 and Cochrane Review Manager (RevMan) version 5.2.

1a.6.1b (Davey et al., 2013)

GRADE Working Group grades of evidence were used. Criteria for inclusion included all randomized and nonrandomized controlled trials (RCTs and CCTs), controlled before-after studies (CBAs) and interrupted time series studies (ITSs) with at least three data points before and after implementation of the intervention in accordance with Effective Practice and Organisation of Care Group (EPOC) criteria. Effect size was estimated with 95% confidence interval (CI) for at least one relevant outcome after the intervention. 'Relevant data' was defined as an intervention that included a change in antibiotic treatment for hospital inpatients and at least one of the study's reported outcomes was directly attributable to change in antibiotic treatment.

1a.6.1c. (Kaki et al., 2011)

Criteria for inclusion were based on the Cochrane Effective Practice and Organization of Care (EPOC) Review Group inclusion criteria for randomized controlled trials, interrupted time series and controlled before-and-after studies, modified to allow inclusion of uncontrolled before-and-after studies, as long as they met the following criteria: (i) measurement and reporting of potential confounding variables from the before-and-after periods; and (ii) either no statistically significant differences (P,0.05) were identified among the measured confounders, or if significant differences were identified, they were adjusted for by multivariate regression

1a.7. FINDINGS FROM SYSTEMATIC REVIEW OF BODY OF THE EVIDENCE SUPPORTING THE MEASURE

1a.7.1. What was the specific structure, treatment, intervention, service, or intermediate outcome addressed in the evidence review?

1a.7.1a (Feazel, 2014)

Outcomes addressed: Effect of antimicrobial stewardship programs on the risk of CDI in hospitalized adult patients (through meta-analysis)

Main Results: When the results of all studies were pooled in a random effects model, implementation of ASPs had an overall protective benefit (pooled risk ratio: 0.48; 95% CI: 0.38, 0.62), indicating a risk reduction for CDI of 52%.

1a.7.1b (Davey et al., 2013)

Outcomes addressed: Impact of hospital antibiotic stewardship interventions on reducing the incidence of antimicrobial-resistant pathogens or Clostridium difficile infection and their impact on clinical outcome.

Main Results: Reliable data about impact on antibiotic prescribing data were available for 76 interventions (44 persuasive, 24 restrictive and 8 structural). For the persuasive interventions, the median change in antibiotic prescribing was 42.3% for the ITSs, 31.6% for the controlled ITSs, 17.7% for the CBAs, 3.5% for the cluster-RCTs and 24.7% for the RCTs. Interventions intended to decrease excessive prescribing were associated with reduction in Clostridium difficile infections and colonization or infection with aminoglycoside- or cephalosporin-resistant gram-negative bacteria, methicillin-resistant Staphylococcus aureus and vancomycin-resistant Enterococcus faecalis. Meta-analysis of clinical outcomes showed that four interventions intended to increase effective prescribing for pneumonia were associated with significant reduction in mortality (risk ratio 0.89, 95% CI 0.82 to 0.97), whereas nine interventions intended to decrease excessive prescribing were not associated with significant increase in mortality (risk ratio 0.92, 95% CI 0.81 to 1.06).

1a.7.1c. (Kaki et al., 2011)

Outcomes addressed: Antibiotic drug utilization, antibiotic costs, antibiotic appropriateness, antibiotic duration, Clostridium difficile colitis, other antibiotic adverse effects, antibiotic resistance, nosocomial infection rates, length of stay and mortality.

Conclusion: "Most stewardship interventions are associated with a decrease in either targeted or overall antibiotic use in critical care patients. A variety of stewardship interventions have been associated with reduced antimicrobial durations of therapy, but impacts of antibiotic appropriateness have only been narrowly studied (and documented) in programmes based on computer-assisted decision support. Similarly, adverse events have only been evaluated with computer-assisted decision support programmes, and arguably the most important antimicrobial adverse event (C. difficile colitis) has yet to be assessed for any stewardship intervention in the ICU context. Importantly, the reductions in antimicrobial utilization associated with stewardship interventions have not been associated with any worsening in nosocomial infection rates, length of stay or mortality among intensive care patients."

1a.7.1d (Davey et al., 2006)

Outcomes addressed: Microbiologic outcomes: gram-negative resistant bacteria, 10 studies; *Clostridium difficile*–associated diarrhea, 5 studies; vancomycin-resistant enterococci, 3 studies; and methicillin-resistant *Staphylococcus aureus*, 2 studies.

Primary conclusion: Four of the 16 studies provided strong evidence that changes in prescribing antimicrobial drugs to hospital inpatients can improve microbial outcomes. Eight of the remaining studies provided some evidence that antimicrobial drug–prescribing interventions can improve microbial outcomes, but flaws in their design indicated that there were plausible alternative explanations for the results. The remaining 4 studies were unequivocally negative.

Implications for Practice: The evidence supports the theory that limiting the use of specific antimicrobial drugs will reduce the prevalences of resistant gram-negative bacteria and C. diff infection. For gram-positive bacteria, there is a lack of evidence rather than evidence of no effect.

1a.7.2. Grade assigned for the quality of the quoted evidence with definition of the grade:

Summary of findings for the main comparison.

Interventions focused on healthcare professionals to improve antibiotic prescribing compared with usual care none to improve antibiotic prescribing in acute inpatient settings.

Effect size for intervention versus control	**Number of studies**	**Quality of the evidence (GRADE)**
Risk of mortality 0.92 (95% CI 0.81 to 1.06)	11 comparisons from 11 studies (7 RCT, 3 cluster-RCT, 1 cluster-CCT) in 20 hospitals with 9,817 patients	**Moderate** High risk of bias especially around study design
Difference in length of stay -0.04 days (95% CI - 0.34 to 0.25)	6 comparisons from 6 studies (4 RCT, 2 cluster-RCT) in 8 hospitals with 8,071 patients	**Very Low** Studies are very heterogeneous and have high risk of bias
Risk of readmission 1.26 (95%CI 1.02 to 1.57)	5 comparisons from 5 studies (4 RCT, 1 Cluster-RCT) in 12 hospitals with 5,856 patients	**Very Low** Studies are very heterogeneous and have high risk of bias

GRADE Working Group grades of evidence

High quality: Further research is very unlikely to change our confidence in the estimate of effect.

Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

Very low quality: We are very uncertain about the estimate.

Abbreviations: CBA: controlled before and after; CCT: controlled clinical trial; CI: confidence interval; ITS: interrupted time series; RCT: randomized controlled trial.

1a.7.4. What is the time period covered by the body of evidence? (*provide the date range, e.g., 1990-2010*).

Date range: no date limit-2013 (Feazel, 2014)

Date range: 1980-2006 (Davey et al., 2013)

Date range: 1996 to 2010 (Kaki et al., 2011)

Date range: 1980 to 2003 (Davey et al., 2006)

QUANTITY AND QUALITY OF BODY OF EVIDENCE

1a.7.5. How many and what type of study designs are included in the body of evidence? (*e.g., 3 randomized controlled trials and 1 observational study*)

1a.7.5a (Feazel, 2014)

Of 78 full articles review, 16 studies the meta-analysis inclusion criteria reflecting 442 873 patients, 13 were quasi-experimental (6 ITS and 7 were before–after), 1 was a retrospective case–control and 1 was a retrospective cohort studies; none were randomized.

1a.7.5a (Davey et al., 2013)

89 studies that reported 95 interventions. Of the 89 studies, 56 were ITSs (of which 4 were controlled ITSs), 25 were RCT (of which 5 were cluster-RCTs), 5 were CBAs and 3 were CCTs (of which 1 was a cluster-CCT).

1a.7.5b. (Kaki et al., 2011)

24 studies that met quality inclusion criteria, 3 were randomized controlled trials; 3 were interrupted time series and 18 (75%) uncontrolled before-and-after studies.

1a.7.5c (Davey et al., 2006)

16 of 66 intervention studies that met the inclusion criteria reported reliable data about microbiologic outcomes. Two studies were RCTs; 1 study was a CCT; 13 studies used an ITS design.

1a.7.6. What is the overall quality of evidence <u>across studies</u> in the body of evidence? (*discuss the certainty or confidence in the estimates of effect particularly in relation to study factors such as design flaws, imprecision due to small numbers, indirectness of studies to the measure focus or target population*)

Both the Cochrane review and the systematic review focused on the ICU setting demonstrated a decrease in use of targeted or overall antibiotics. Decreases in excessive prescribing were associated in a reduction of C. diff infections for Cochrane reviews that included a larger number of studies, but were not assessed in studies included in the ICU-focused systematic review by Kaki et al. The meta-analysis by Feazel et al. showed a significant risk reduction for C. diff with stewardship programs, but did not report quantitative changes of the antibiotic use for the studies included.

The Cochrane review reported that 18 (20%) of the studies had low risk of bias, 31 (35%) studies had medium risk of bias and 40 (45%) had high risk of bias. High risk of bias was more common in CCTs, RCTs or CRCTs (22/28, 79%) than in ITS or CITS (13/56, 23%). Most of the RCTs had high risk of selection bias. The only RCTs that had low risk of selection bias were either cluster-RCTs or structural interventions, for which concealment of allocation is relatively straightforward. The systematic review by Kaki et al. was limited by the quality of studies available for analysis as well as limitations inherent in the relaxation of Cochrane quality criteria led to the inclusion of many before-and-after analyses, which are prone to temporal confounding. The meta-analysis relied on observational quasi-experimental studies, a design that is subject to many biases and can suffer from lower internal validity. In addition, there was high heterogeneity among the studies included, which was addressed with stratified analysis.

ESTIMATES OF BENEFIT AND CONSISTENCY ACROSS STUDIES IN BODY OF EVIDENCE

1a.7.7. What are the estimates of benefit—magnitude and direction of effect on outcome(s) <u>across studies</u> in the body of evidence? (*e.g., ranges of percentages or odds ratios for improvement/ decline across studies, results of meta-analysis, and statistical significance*)

(Davey et al., 2013 – Cochrane Review)

Impact of interventions on prescribing outcomes (i.e., reduction in targeted or overall antibiotic use)

- Overall, for the persuasive interventions, the median (interquartile range) change in antibiotic prescribing was 42.3% for the interrupted time series studies (ITSs), 31.6% for the controlled interrupted time series studies (CITSs), 17.7% for the controlled before-after studies (CBAs), 3.5% for the cluster-randomized controlled trials (CRCTs) and 24.7% for the randomized controlled trials (RCTs).
- Overall, the restrictive interventions had a median effect size of 34.7% for the interrupted time series designs, 17.1% for the controlled before-after designs and 40.5% for the randomized controlled trials.
- Meta-analysis of persuasive versus restrictive interventions (52 studies) showed a consistent impact on prescribing and microbial outcomes with at least 25% of studies showing an effect in the intended direction at each time point.Clostridium difficile infections: (Five studies) showing a median effect size of a 68% reduction.

(Feazel, 2014 – Systematic review and meta-analysis)

- Overall, the pooled risk ratio for the effect of stewardship programs on the risk of C. diff infections was 0.48 (95% CI 0.38, 0.62. p< 0.00001). By setting, the risk ratio was most effective in geriatrics population [0.44 (95% CI 0.35, 0.56; p<,0.00001)
- Stewardship activities that focused on specific antibiotics were effective in reducing risk for these classes known to increase C. Diff risk: cephalosporins 0.50 (95% CI 0.39, 0.64; p< 0.00001) and fluoroquinolones 0.45 (95% CI 0.30, 0.67; p<0.00001)

1a.7.8. What harms were studied and how do they affect the net benefit (benefits over harms)?

(Davey et al., 2013– Cochrane Review)

- Mortality: There was no increase in total mortality associated with interventions that intended to reduce unnecessary antibiotic treatment. Of note, indicators of mortality were not limited to patients with sepsis or defined infections.
- Readmission: Five studies that examined readmission showed a significant increase in readmissions associated with the interventions. Infection-related readmissions accounted for 39% of readmissions within 30 days, and there was no significant difference between intervention and control groups for infection-related readmissions. Challenges of reliable measurement of infection-related readmissions raised doubts about the validity of readmission as a health outcome measure for interventions to reduce excessive antibiotic prescribing.

UPDATE TO THE SYSTEMATIC REVIEW(S) OF THE BODY OF EVIDENCE

1a.7.9. If new studies have been conducted since the systematic review of the body of evidence, provide for <u>each</u> new study: 1) citation, 2) description, 3) results, 4) impact on conclusions of systematic review.

The Cochrane review published in 2013 included evidence current to 2006. The published work on antimicroibal stewardship has continued to increse since 2006 and this review is in the process of being updated. It is anticipated that a greater number of studies will reinforce the finding and increase the confidence in estimates of the effect.

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.

The measure provides summary results that hospital and health system antimicrobial stewardship programs (ASPs) can use as quantitative aids in their efforts to evaluate and improve antibiotic prescribing. The Standardized Antimicrobial Administration Ratios (SAARs) that comprise the measure focus on high value targets and high level indicators of antimicrobial use for ASPs. The SAARs can be used by ASPs to benchmark antimicrobial use in multiple patient care locations, identify opportunities for improvement, and gauge the impact of stewardship efforts. At the outset, the SAARs provide a set of signals that often warrant further analysis, such as an evaluation of the extent to which a specific antibiotic or group of antibiotics accounts for a high or low SAAR value and the extent to which an antibiotic or group of antibiotics were used appropriately for specific indications. While the SAARs do not provide a definitive indication that antibiotics are overused or underused, they provide an important starting place for further analysis and possible action. Some of the analytic follow up can be completed with hospital- and patient care location-specific data reported to CDC's National Healthcare Safety Network (NHSN) Antimicrobial Use and Resistance (AUR) Module, using analytic features built into the NHSN application. However, additional analyses to determine the appropriateness of antibiotic use in individual instances are likely to require access to detailed, patient-level data that is beyond the scope of data collection and analysis using the NHSN module, e.g., clinical indications for specific antibiotics and dose and duration decisions.

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.*

See Table 3 - NHSN SAAR Distribution and statistical comparison by reporting measure

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.

Numerous individual studies and systematic reviews provide strong evidence that measurement of antimicrobial use and data-driven interventions by antimicrobial stewardship programs (ASPs) lead to more judicious use of antibiotics, reduced antimicrobial resistance, and other favorable healthcare outcomes (Davey 2017, Feazel 2014; Davey 2006; Davey 2013; Kaki 2011).

Antimicrobial use measurement enables ASPs to understand prescribing practices, focus efforts on improvement, and determine the impact of their activities (Pollack, 2014). Although standardized metrics have been developed to measure antibiotic use, differences in measurement, limited uptake, and variation among facilities has impeded the ability to compare antibiotic use among hospitals.

The measure serves as a quantitative guide for hospital and health system ASPs, enabling them to benchmark antibiotic use in their facilities and patient care locations against nationally aggregated data. The measure focuses on antibiotic agents that have been shown to be high value targets for antimicrobial stewardship programs activities such as protocols for use or post-prescription reviews to determine need for de-escalation,

dose-optimization or oral conversion. Knowledge about antibiotic use patterns of these agents is a primary means to prioritize and evaluate antimicrobial stewardship efforts.

Citations:

Feazel LM, Malhotra A, Perencevich EN, Kaboli P, Diekema DJ, Schweizer ML. Effect of antibiotic stewardship programmes on Clostridium difficile incidence: a systematic review and meta-analysis. J Antimicrob Chemother. 2014;69(7):1748-54. http://jac.oxfordjournals.org/content/69/7/1748.full.pdf

Davey P, Brown E, Fenelon L, Finch R, Gould I, Holmes A, et al. Systematic review of antimicrobial drug prescribing in hospitals. Emerg Infect Dis. 2006;12(2):211-6. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3373108/

Davey P, Brown E, Charani E, Fenelon L, Gould IM, Holmes A, et al. Interventions to improve antibiotic prescribing practices for hospital inpatients. Cochrane Database Syst Rev. 2013;4:CD003543. http://onlinelibrary.wiley.com/store/10.1002/14651858.CD003543.pub3/asset/CD003543.pdf?v=1&t=hvxzajv5&s=a6f3c724ce051d8acba5866a07e3c5ac8c818e83

Kaki R, Elligsen M, Walker S, Simor A, Palmay L, Daneman N. Impact of antimicrobial stewardship in critical care: a systematic review. J Antimicrob Chemother. 2011;66(6):1223-30. http://jac.oxfordjournals.org/content/66/6/1223.full.pdf

Pollack LA, Srinivasan A. Core Elements of Hospital Antibiotic Stewardship Programs from the Centers for Disease Control and Prevention. Clinical Infectious Diseases. 2014;59(suppl 3):S97-S100.

Barlam TF, Cosgrove SE, Abbo LM, MacDougall C, Schuetz AN, Septimus EJ, Srinivasan A, Dellit TH, Falck-Ytter YT, Fishman NO, Hamilton CW, Jenkins TC, Lipsett PA, Malani PN, May LS, Moran GJ, Neuhauser MM, Newland JG, Ohl CA, Samore MH, Seo SK, Trivedi KK. Executive Summary: Implementing an Antibiotic Stewardship Program: Guidelines by the Infectious Diseases Society of America and the Society for Healthcare Epidemiology of America. Clin Infect Dis. 2016 May 15;62(10):e51-e77.

Davey P, Marwick CA, Scott CL, Charani E, McNeil K, Brown E, Gould IM, Ramsay CR, Michie S. Cochrane Database of Systematic Reviews 2017, Issue 2. Art. No.: CD003543. DOI: 10.1002/14651858.CD003543.pub4

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.*

Data source includes 2,156 adult patient care locations from 449 acute care hospitals and 170 pediatric patient care locations from 109 acute care hospitals from which AU data reported for 2017 were used to update predictive models. These data do not include disparity descriptors such as race, ethnicity, gender and age. However there is no compelling external data or analytic work suggesting that variation in hospital antimicrobial use is associated with social risk factors.

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

Sparse data are available on disparities in appropriateness of antibiotic use in hospitals. A retrospective analysis (1996-2007) of prospective data on all surgical patients treated for sepsis at to a tertiary care center demonstrated no differences in demographic and comorbidities between inappropriately and appropriately treated groups. (Davies et al, 2014)

Davies SW, Efird JT, Guidry CA, Hranjec T, Metzger R, Swenson BR, et al. Does it Matter if we get it right? Impact of appropriateness of empiric antimicrobial therapy among surgical patients. Shock. 2014;42(3):185-91.

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):*

Infectious Diseases (ID)

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

Safety : Overuse

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

Populations at Risk

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.)*

http://www.cdc.gov/nhsn/PDFs/pscManual/11pscAURcurrent.pdf

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:** Copy_of_TABLE_1_NHSN_Antimicrobial_Use_Measure_-S.2-_S_15._Detailed_risk_model_specifications-.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.

No, this is not an instrument-based measure **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.

Yes

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.

Adult and pediatric patient populations were modeled separately, adult general hematology-oncology wards and adult step-down units were added, and antimicrobial groupings were added and re-categorized as part of measure maintenance to take into account new agents and new insights from data that were reported to NHSN and antimicrobial stewardship SMEs whose input was sought in a systematic review. New antimicrobial and patient care location groupings, with new predictive models for each antimicrobial –patient care location grouping.

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).

Days of antimicrobial therapy for antimicrobial agents administered to adult and pediatric patients in medical ICU, medical-surgical ICU, surgical ICU (adult only), medical ward, medical-surgical ward, surgical ward, general hematology-oncology ward (adult only), and step-down unit (adult only).

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).

An antimicrobial day (also known as a day of therapy) is defined by any amount of a specific antimicrobial agent administered in a calendar day to a particular patient as documented in an electronic medication administration record (eMAR) and/or bar coding medication record (BCMA). All antimicrobial days for specified categories of antibacterial agents administered in specified patient care locations—adult and pediatric medical ICU, medical-surgical ICU, surgical ICU (adult only), medical ward, medical-surgical ward, surgical ward, general hematology-oncology ward (adult only), and step-down unit (adult only)—are summed for each location across months and comprise the numerator data for the measure. The specified categories of antimicrobial agents are: 1) Broad spectrum antibacterial agents predominantly used for hospital-onset infections, 2) Broad spectrum antibacterial agents predominantly used for community-acquired infections, 3) Antibacterial agents predominately used for resistant Gram-positive infections, 4) Narrow spectrum beta-lactam agents, 5) Antifungal agents predominantly used for invasive candidiasis, 6) Antibacterial agents posing the highest risk for CDI, 7) Azithromycin (pediatrics only), 8) All antibacterial agents.

See attached Table 1. NHSN Antimicrobial Use Measure proposal for lists and descriptions of patient care locations and antibacterial agent categories

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

Days present for each patient care location—adult and pediatric medical ICU, medical-surgical ICU, surgical ICU (adult only), medical ward, medical-surgical ward, surgical ward, general hematology-oncology ward (adult only), and step-down unit (adult only) is defined as the number of patients who were present for any portion of each day of a calendar month for each location. The day of admission, discharge, and transfer to and from locations are included in days present. All days present are summed for each location and month, and the aggregate sums for each location-month combination comprise the denominator data for the measure.

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).

See attached Table 1b. NHSN Antimicrobial Use Measure proposal for list and description of patient care locations included in the measure.

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

Hospital patient care locations other than adult and pediatric medical ICU, medical-surgical ICU, surgical ICU (adult only), medical ward, medical-surgical ward, surgical ward, general hematology-oncology ward (adult only), and step-down unit (adult only) are excluded from this measure.

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.)*

See Table 1b. NHSN Antimicrobial Use Measure for description of patient care locations. Listed locations are included in the measure; all other locations are excluded.

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.)*

Antimicrobial use data is stratified by hospital-specific and patient care location-specific variables: hospital teaching status (major [medical school and post-graduate training], graduate only [residents and/or fellows], undergraduate only [medical students], not a teaching hospital); hospital bedsize; hospital ICU bedsize; percentage of ICU beds among total beds (number ICU beds/total number hospital beds); average length of hospital stay (number annual admissions/ number annual patient days); patient care location.

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

Statistical risk model

If other:

S.12. Type of score:

Ratio

If other:

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 = Score within a defined interval

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.)*

The Standardized Antimicrobial Administration Ratio (SAAR), the ratio of observed to predicted antimicrobial use, is a score that can be above, equal to, or below 1.0. A high score (above 1.0) that achieves statistical significance may indicate excessive antimicrobial use. A score that is not significantly different than 1.0 indicates antimicrobial use that is equivalent to the referent population's antimicrobial use. A low score (below 1.0) that achieves statistical significance may indicate antimicrobial under use.

Each SAAR is calculated as follows:

1. Identify the antimicrobial days reported for each patient care location included in the SAAR for the measurement period

2. Total each of these numbers for an observed number of antimicrobial days

3. Obtain the predicted antimicrobial days in the same patient care locations by multiplying the observed days present by the corresponding antimicrobial use rate in the standard population obtained from the relevant regression model

4. Sum the predicted antimicrobial days for the patient care locations included in the SAAR

5. Divide the total number of antimicrobial days by the predicted number of antimicrobial days

6. Result = SAAR

A discrete set of SAARs comprise the antimicrobial use measure: SAARs that are intended to serve as high value targets for antimicrobial stewardship programs and SAARs that are intended to serve as high level indicators of all antimicrobial use across multiple patient care locations.

High value targets – SAARs for 38 different antibacterial agent-patient care location combinations (24 adult, 14 pediatric)

Adult

1. Broad spectrum antibacterial agents predominantly used for hospital-onset infections – adult medical, medical-surgical, and surgical intensive care units

2. Broad spectrum antibacterial agents predominantly used for hospital-onset infections – adult medical, medical-surgical, and surgical wards

3. Broad spectrum antibacterial agents predominantly used for hospital-onset infections – adult general hematology-oncology wards

4. Broad spectrum antibacterial agents predominantly used for hospital-onset infections – adult step-down units

5. Broad spectrum antibacterial agents predominantly used for community-acquired infections – adult medical, medical-surgical, and surgical intensive care units

6. Broad spectrum antibacterial agents predominantly used for community-acquired infections – adult medical, medical-surgical, and surgical wards

7. Broad spectrum antibacterial agents predominantly used for community-acquired infections – adult general hematology-oncology wards

8. Broad spectrum antibacterial agents predominantly used for community-acquired infections – adult step-down units

9. Antibacterial agents predominantly used for resistant Gram-positive infections – adult medical, medical-surgical, and surgical intensive care units

10. Antibacterial agents predominantly used for resistant Gram-positive infections – adult medical, medical-surgical, and surgical wards

11. Antibacterial agents predominantly used for resistant Gram-positive infections – adult general hematology-oncology wards

12. Antibacterial agents predominantly used for resistant Gram-positive infections – adults step-down units

13. Narrow spectrum beta-lactam agents – adult medical, medical-surgical, and surgical intensive care units

14. Narrow spectrum beta-lactam agents – adult medical, medical-surgical, and surgical wards

15. Narrow spectrum beta-lactam agents – adult general hematology-oncology wards

16. Narrow spectrum beta-lactam agents – adult step-down units

17. Antibacterial agents posing highest risk for CDI – adult medical, medical-surgical, and surgical intensive care units

18. Antibacterial agents posing highest risk for CDI – adult medical, medical-surgical, and surgical wards

19. Antibacterial agents posing highest risk for CDI – adult general hematology-oncology wards

20. Antibacterial agents posing highest risk for CDI – adult step-down units

21. Antifungal agents predominantly used for invasive candidiasis – adult medical, medical-surgical, and surgical intensive care units

22. Antifungal agents predominantly used for invasive candidiasis – adult medical, medical-surgical, and surgical wards

23. Antifungal agents predominantly used for invasive candidiasis – adult general hematology-oncology wards

24. Antifungal agents predominantly used for invasive candidiasis – adult step-down units

Pediatric

1. Broad spectrum antibacterial agents predominantly used for hospital-onset infections – pediatric medical and medical-surgical intensive care units

2. Broad spectrum antibacterial agents predominantly used for hospital-onset infections – pediatric medical, medical-surgical, and surgical wards

3. Broad spectrum antibacterial agents predominantly used for community-acquired infections – pediatric medical and medical-surgical intensive care units

4. Broad spectrum antibacterial agents predominantly used for community-acquired infections – pediatric medical, medical-surgical, and surgical wards

5. Antibacterial agents predominantly used for resistant Gram-positive infections – pediatric medical and medical-surgical intensive care units

6. Antibacterial agents predominantly used for resistant Gram-positive infections – pediatric medical, medical-surgical, and surgical wards

7. Narrow spectrum beta-lactam agents – pediatric medical and medical-surgical intensive care units

8. Narrow spectrum beta-lactam agents – pediatric medical, medical-surgical, and surgical wards

9. Azithromycin – pediatric medical and medical-surgical intensive care units

10. Azithromycin – pediatric medical, medical-surgical, and surgical wards

11. Antibacterial agents posing highest risk for CDI – pediatric medical and medical-surgical intensive care units

12. Antibacterial agents posing highest risk for CDI – pediatric medical, medical-surgical, and surgical wards

13. Antifungal agents predominantly used for invasive candidiasis – pediatric medical and medical-surgical intensive care units

14. Antifungal agents predominantly used for invasive candidiasis – pediatric medical, medical-surgical, and surgical wards

High level indicators – SAARs for 2 different antibacterial agent-patient care location combinations

Adult

1. All antibacterial agents – adult medical, medical-surgical, and surgical intensive care units and wards, general hematology-oncology wards, step-down units

Pediatric

1. All antibacterial agents – pediatric medical intensive care units and wards, medical-surgical intensive care units and wards, and surgical wards

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.

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.

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

If other, please describe in S.18.

Electronic health data; electronic format Admission Discharge Transfer

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.

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)*

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

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

Facility

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

Inpatient/Hospital

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.)*

2. Validity – See attached Measure Testing Submission Form

Table_3_NQF_testing_document_Final_-Submit_to_NQF_12-17-18.xlsx,AU_measure_NQF_testing_attachment_2019_v.7.1_Final-Submitted_to_NQF_12-17-18-636906577271743387-revised4-25-19.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.

Yes

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.

No

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.

Yes - Updated information is included

Measure Testing (subcriteria 2a2, 2b1-2b6)

Measure Number (*if previously endorsed*)**:** 2720
Measure Title: National Healthcare Safety Network (NHSN) Antimicrobial Use Measure
Date of Submission: 12/17/2018

Type of Measure:

☐ **Outcome (*including PRO-PM*)**	☐ **Composite – *STOP – use composite testing form***
☐ Intermediate Clinical Outcome	☐ Cost/resource
☒ Process *(including Appropriate Use)*	☐ Efficiency
☐ Structure	

Instructions

1. **Measures must be tested for all the data sources and levels of analyses that are specified. *If there is more than one set of data specifications or more than one level of analysis, contact NQF staff* about how to present all the testing information in one form.**
2. **For all measures, sections 1, 2a2, 2b1, 2b2, and 2b4 must be completed.**
3. **For outcome and resource use measures, section 2b3 also must be completed.**
4. **If specified for multiple data sources/sets of specificaitons (e.g., claims and EHRs), section 2b5 also must be completed.**
5. **Respond to all questions as instructed with answers immediately following the question. All information on testing to demonstrate meeting the subcriteria for reliability (2a2) and validity (2b1-2b6) must be in this form. An appendix for *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.**
6. **If you are unable to check a box, please highlight or shade the box for your response.**
7. **Maximum of 25 pages (*incuding questions/instructions;* minimum font size 11 pt; do not change margins). *Contact NQF staff if more pages are needed.***
8. **Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.**
9. **For information on the most updated guidance on how to address social risk factors variables and testing in this form refer to the release notes for version 7.1 of the Measure Testing Attachment.**

Note: The information provided in this form is intended to aid the Standing Committee and other stakeholders in understanding to what degree the testing results for this measure meet NQF's evaluation criteria for testing.

2a2. Reliability testing [10] demonstrates the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise. For instrument-based measures (including PRO-PMs) and composite performance measures, reliability should be demonstrated for the computed performance score.

2b1. Validity testing [11] demonstrates that the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality. For instrument-based measures (including PRO-PMs) and composite performance measures, validity should be demonstrated for the computed performance score.

2b2. Exclusions are supported by the clinical evidence and are of sufficient frequency to warrant inclusion in the specifications of the measure; [12]

AND

If patient preference (e.g., informed decisionmaking) is a basis for exclusion, there must be evidence that the exclusion impacts performance on the measure; in such cases, the measure must be specified so that the information about patient preference and the effect on the measure is transparent (e.g., numerator category computed separately, denominator exclusion category computed separately). [13]

2b3. For outcome measures and other measures when indicated (e.g., resource use):

- **an evidence-based risk-adjustment strategy (e.g., risk models, risk stratification) is specified; is based on patient factors (including clinical and social risk factors) that influence the measured outcome and are present at start of care; [14,15] and has demonstrated adequate discrimination and calibration**

OR

- **rationale/data support no risk adjustment/ stratification.**

2b4. Data analysis of computed measure scores demonstrates that methods for scoring and analysis of the specified measure allow for identification of statistically significant and practically/clinically meaningful [16] differences in performance;

OR

there is evidence of overall less-than-optimal performance.

2b5. If multiple data sources/methods are specified, there is demonstration they produce comparable results.

2b6. Analyses identify the extent and distribution of missing data (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias.

Notes

10. Reliability testing applies to both the data elements and computed measure score. Examples of reliability testing for data elements include, but are not limited to: inter-rater/abstractor or intra-rater/abstractor studies; internal consistency for multi-item scales; test-retest for survey items. Reliability testing of the measure score addresses precision of measurement (e.g., signal-to-noise).

11. Validity testing applies to both the data elements and computed measure score. Validity testing of data elements typically analyzes agreement with another authoritative source of the same information. Examples of validity testing of the measure score include, but are not limited to: testing hypotheses that the measures scores indicate quality of care, e.g., measure scores are different for groups known to have differences in quality assessed by another valid quality measure or method; correlation of measure scores with another valid indicator of quality for the specific topic; or relationship to conceptually related measures (e.g., scores on process measures to scores on outcome measures). Face validity of the measure score as a quality indicator may be adequate if accomplished through a systematic and transparent process, by identified experts, and explicitly addresses whether performance scores resulting from the measure as specified can be used to distinguish good from poor quality. The degree of consensus and any areas of disagreement must be provided/discussed.

12. Examples of evidence that an exclusion distorts measure results include, but are not limited to: frequency of occurrence, variability of exclusions across providers, and sensitivity analyses with and without the exclusion.

13. Patient preference is not a clinical exception to eligibility and can be influenced by provider interventions.

14. Risk factors that influence outcomes should not be specified as exclusions.

15. With large enough sample sizes, small differences that are statistically significant may or may not be practically or clinically meaningful. The substantive question may be, for example, whether a statistically significant difference of one percentage point in the percentage of patients who received smoking cessation counseling (e.g., 74 percent v. 75 percent) is clinically meaningful;

or whether a statistically significant difference of $25 in cost for an episode of care (e.g., $5,000 v. $5,025) is practically meaningful. Measures with overall less-than-optimal performance may not demonstrate much variability across providers.

1. DATA/SAMPLE USED FOR ALL TESTING OF THIS MEASURE

Often the same data are used for all aspects of measure testing. In an effort to eliminate duplication, the first five questions apply to all measure testing. If there are differences by aspect of testing,(e.g., reliability vs. validity) be sure to indicate the specific differences in question 1.7.

1.1. What type of data was used for testing? (*Check all the sources of data identified in the measure specifications and data used for testing the measure. Testing must be provided for all the sources of data specified and intended for measure implementation.* ***If different data sources are used for the numerator and denominator, indicate N [numerator] or D [denominator] after the checkbox.***)

Measure Specified to Use Data From: (***must be consistent with data sources entered in S.17***)	**Measure Tested with Data From:**
☐ abstracted from paper record	☐ abstracted from paper record
☐ claims	☐ claims
☐ registry	☐ registry
☒ abstracted from electronic health record	☒ abstracted from electronic health record
☐ eMeasure (HQMF) implemented in EHRs	☐ eMeasure (HQMF) implemented in EHRs
☒ other: electronic format Admission Discharge Transfer	☒ other: electronic format Admission Discharge Transfer

1.2. If an existing dataset was used, identify the specific dataset (*the dataset used for testing must be consistent with the measure specifications for target population and healthcare entities being measured; e.g., Medicare Part A claims, Medicaid claims, other commercial insurance, nursing home MDS, home health OASIS, clinical registry*).

Not Applicable

1.3. What are the dates of the data used in testing? 2015-2018

Reliability testing at the data element level using a sample of hospitals that have reported antimicrobial use and days present to the CDC's National Healthcare Safety Network (NHSN) Antimicrobial Use and Resistance (AUR) Module during a 2-year time period, 2016 through 2018.

Validity testing of the core measure data elements, antimicrobial days (i.e., days of therapy) and days present, and the core measure construct of antimicrobial groupings was completed by means of consensus development using a Delphi process during a series of meetings conducted from April to July 2018 among outside expert panel of Infectious Disease physicians and clinical pharmacists who are leaders in their domain.

1.4. What levels of analysis were tested? (*testing must be provided for all the levels specified and intended for measure implementation, e.g., individual clinician, hospital, health plan*)

Measure Specified to Measure Performance of: (***must be consistent with levels entered in item S.20***)	**Measure Tested at Level of:**
☐ individual clinician	☐ individual clinician
☐ group/practice	☐ group/practice

☒ hospital/facility/agency	☒ hospital/facility/agency
☐ health plan	☐ health plan
☐ other: Click here to describe	☐ other:

1.5. How many and which measured entities were included in the testing and analysis (by level of analysis and data source)? (*identify the number and descriptive characteristics of measured entities included in the analysis (e.g., size, location, type); if a sample was used, describe how entities were selected for inclusion in the sample*)

Reliability testing at the data element level using a sample of hospitals that have reported antimicrobial use and days present to the CDC's National Healthcare Safety Network (NHSN) Antimicrobial Use and Resistance (AUR) Module during a 3-year time period, 2011 through 2014. Thirteen facilities completed reliability testing of data elements collection and 24 facilities completed reliability testing of data aggregation. Reliability testing was expanded during the 3-year time period, 2016 through 2018, and included 48 facilities, and these facilities completed reliability testing for data elements collection and data aggregation.

A.) Data elements collection - included reliability testing of antimicrobial administration data elements collected from point of care systems (i.e., electronic medication administration record [eMAR] or bar coding medication administration [BCMA]) and patient care location data collected from admission/discharge/transfer (ADT) systems.

B.) Data aggregation - included reliability testing of the aggregation of numerator data (i.e., antimicrobial days by patient care location) and denominator data (i.e., days present by patient care location) in accordance with NHSN Antimicrobial Use and Resistance (AUR) protocol requirements.

1.6. How many and which patients were included in the testing and analysis (by level of analysis and data source)? (*identify the number and descriptive characteristics of patients included in the analysis (e.g., age, sex, race, diagnosis); if a sample was used, describe how patients were selected for inclusion in the sample*)

Data collection for this measure does not include counts of individual patients or descriptive characteristics of individual patients because patient level data are aggregated prior to submission to NHSN.

1.7. If there are differences in the data or sample used for different aspects of testing (e.g., reliability, validity, exclusions, risk adjustment), identify how the data or sample are different for each aspect of testing reported below.

Yes different samples were used for reliability testing and risk adjustment testing.

The size of the hospital samples used for reliability testing of A.) Data collection and B.) Data aggregation differed because of resource constraints. The hospital samples used for A.) and B.) were convenience samples comprised of teaching and non-teaching hospitals and hospitals with a range of bedsizes (42-919 beds).

The hospital samples used for reliability testing from 2016-2018 were a convenience sample comprised of teaching and non-teaching hospitals and with a range of bedsizes (12-619 beds).

The sample used for risk-adjustment included all eligible patient care locations reporting at least 9 months of data to the AU Option in 2017. For adult predictive models, this sample was comprised of 2,156 patient care locations (131 medical ICUs, 318 medical/surgical ICUs, 73 surgical ICUs, 472 medical wards, 554

medical/surgical wards, 247 surgical wards, 68 general hematology-oncology wards, 293 step-down units) from 449 hospitals. For pediatric predictive models, the sample was comprised of 170 patient care locations (4 medical ICUs, 46 medical/surgical ICUs, 21 medical wards, 94 medical/surgical wards, 5 surgical wards) from 109 hospitals.

1.8 What were the social risk factors that were available and analyzed? For example, patient-reported data (e.g., income, education, language), proxy variables when social risk data are not collected from each patient (e.g. census tract), or patient community characteristics (e.g. percent vacant housing, crime rate) which do not have to be a proxy for patient-level data.

No patient-level sociodemographic variables are used in the measure and none were available for analysis.

2a2. RELIABILITY TESTING

***Note**: If accuracy/correctness (validity) of data elements was empirically tested, separate reliability testing of data elements is not required – in 2a2.1 check critical data elements; in 2a2.2 enter "see section 2b2 for validity testing of data elements"; and skip 2a2.3 and 2a2.4.*

2a2.1. What level of reliability testing was conducted? (*may be one or both levels*)
☒ **Critical data elements used in the measure** (*e.g., inter-abstractor reliability; data element reliability must address ALL critical data elements*)
☐ **Performance measure score** (e.g., *signal-to-noise analysis*)

2a2.2. For each level checked above, describe the method of reliability testing and what it tests (*describe the steps—do not just name a method; what type of error does it test; what statistical analysis was used*)

A.) Data collection

Through this multi-step process, hospitals compared the antimicrobial use data in the local eMAR or BCMA system to the antimicrobial use data that were collected using the third party software vendor system or the healthcare system's corporate data warehouse to ensure the 'data feed' between the two systems, from eMAR or BCMA system to vendor system or data warehouse, was accurate. The data collected via the third party software system or corporate data warehouse are the data transferred to NHSN. Note: While the antimicrobial use data in the local eMAR or BCMA system was considered the gold standard, some hospitals did identify inconsistencies in their local systems that were addressed during the process. The steps used for reliability testing are as follows:

1. Hospitals manually review antimicrobial use line list generated from the eMAR or BCMA system to verify that all required antimicrobials are present on the line list or appropriately absent.
2. Hospitals manually review routes of antimicrobial administration collected to verify that all route data are accurately collected from eMAR or BCMA systems.
3. Hospitals include patient care location data from ADT systems in manual verification of antimicrobial administration and route of administration per patient per calendar day in three separate patient locations.

B.) Data Aggregation

Through this process, hospitals and third party software vendors compared the numerator and denominator data in the third party software system pre- and post-aggregation completed by the software

to ensure that the software was correctly aggregating the patient-level information. The gold standard in this process was the data obtained from eMAR, BCMA, and/or ADT systems.

1. Hospitals verify that day of discharge and transfer from a patient care location are included in aggregations of antimicrobial days and days present.
2. Hospitals confirm that facility-wide counts of days present conform to the NHSN protocol requirements.

2a2.3. For each level of testing checked above, what were the statistical results from reliability testing? (e.*g., percent agreement and kappa for the critical data elements; distribution of reliability statistics from a signal-to-noise analysis*)

As per NQF recommendation, this section was moved to validity.

2a2.4 What is your interpretation of the results in terms of demonstrating reliability? (i.*e., what do the results mean and what are the norms for the test conducted?*)

Highly reliable data.

2b1. VALIDITY TESTING

2b1.1. What level of validity testing was conducted? (*may be one or both levels*)

☒ **Critical data elements** (*data element validity must address ALL critical data elements*)

☐ **Performance measure score**

☐ **Empirical validity testing**

☐ **Systematic assessment of face validity of performance measure score as an indicator** of quality or resource use (*i.e., is an accurate reflection of performance on quality or resource use and can distinguish good from poor performance*) **NOTE**: Empirical validity testing is expected at time of maintenance review; if not possible, justification is required.

CDC conducted systematic review and a Delphi process with a national panel of antimicrobial stewardship experts including Infectious diseases physicians and clinical pharmacists who are leaders in their domain.

2b1.2. For each level of testing checked above, describe the method of validity testing and what it tests (*describe the steps—do not just name a method; what was tested, e.g., accuracy of data elements compared to authoritative source, relationship to another measure as expected; what statistical analysis was used)*

As per NQF recommendation, reliability of data elements was moved to the validity section.

Percent agreement was used as the summary measure for reliability testing during the initial testing, 2011 through 2014, and the more recent testing, 2016 through 2018. NHSN's rules for what constitutes an antimicrobial day or a day present were invoked to confirm (or refute) that the third party vendor solution or corporate data warehouse conformed to the requirements. The CDC protocol for testing validity of antimicrobial days and days present—and the percent agreement--was used to identify and correct systematic data omissions or erroneous data transformations before the reportable data were submitted to NHSN. The gold standard used to measure the percent agreement was the data in the local eMAR or BCMA system, which provided the basis for comparison with the measure data to be submitted from either a third party software

system or a data warehouse. Percent agreement for the measure data elements ranged from 60-80% for the antimicrobial days numerator and 70-80% for the days present denominator at the outset of validation process. By design the process led to >99% agreement for all required data elements prior to data submission to CDC. All patients in eligible patient care locations are included in the data submission therefore exclusions were not validated and there are no validity results to report.

Validity testing of the core measure data elements, antimicrobial days (i.e., days of therapy) and days present, and the core measure construct of antimicrobial groupings was completed by means of consensus development using a Delphi process during a series of meetings conducted from April to July 2018 among outside expert panel of Infectious Disease physicians and clinical pharmacists who are leaders in their domain. The full list of panelists can be found below.

Expert panel informing Adult Standardized Antimicrobial Administration Ratio (SAARs):

Name	Title	Affiliation
Christopher Graber, MD	Infectious Disease Specialist, Clinical Director, VA Greater Los Angeles Antimicrobial Stewardship Program Director	VA Greater Los Angeles Healthcare System
Clark Force, RPh	Antimicrobial Stewardship Pharmacist	Tucson Medical Center
Eddie Stenehjem, MD	Medical Director of Intermountain Healthcare's Urban Central Region Antimicrobial Stewardship Program	Intermountain Healthcare
Edward Septimus, MD	Medical Director Infection Prevention and Epidemiology	Clinical Services Group HCA Healthcare System
Elizabeth Dodds-Ashley, PharmD	Associate Professor of Medicine, Clinical Pharmacist	Duke Antimicrobial Stewardship Outreach Network
Florian Daragjati, PharmD	Pharmacy Coordinator, Ascension Center for Excellence for Antimicrobial Stewardship and Infection Prevention	Ascension
Haley Burgess, PharmD	AVP, Clinical Pharmacy and Medication Safety	Clinical Services Group HCA Healthcare System
Kalvin Yu, MD	Regional Chief of Infectious Diseases, Southern California Permanente Medical Group	Kaiser Permanente West Los Angeles Medical Center
Kevin Hsueh, MD	Medical Director of Antimicrobial Stewardship, Associate Medical Director for Infection Prevention	Washington University School of Medicine
Makoto Jones, MD	Infectious Disease Physician, Associate Professor at University of Utah	VA Salt Lake City Healthcare System
Marc Meyer, BPharm, RPh	Clinical Pharmacist for Southwest Health System	Southwest Memorial Hospital
Marc Scheetz, PharmD	Infectious Disease Pharmacist	Northwestern Memorial Hospital
Matthew Goetz, MD	Chief, Infectious Diseases	VA Greater Los Angeles Healthcare System
Mohamad Fakih, MD	Senior Medical Director, Ascension Center for Excellence for Antimicrobial Stewardship and Infection Prevention	Ascension
Rebecca Battjes, MPH	Infection Preventionist	Ascension St. John Hospital
Rebekah Moehring, MD	Medical Director, Duke Antimicrobial Stewardship and Evaluation Team	Duke Center for Antimicrobial Stewardship and Infection Prevention, Duke Medical Center
Stanley Deresinski, MD	Clinical Professor of Medicine, Division of Infectious Diseases and Geographic Medicine	Stanford University School of Medicine
Whitney Buckel, PharmD	Antimicrobial Stewardship Pharmacist Manager	Intermountain Healthcare

Expert panel informing Pediatric SAARs:

Name	Title	Affiliation
Adam Hersh, MD, PhD	Associate Professor of Pediatrics, Division of Pediatric Infectious Diseases	University of Utah
Andrea Benin, MD	Senior Vice President, Quality and Patient Safety	Connecticut Children's Medical Center
Jared Olson, PharmD	Clinical Pharmacist, Infectious Diseases, Co-Director of Antimicrobial Stewardship Program	Intermountain Healthcare, Primary Children's Medical Center
Jason Child, PharmD	HIV/Infectious Diseases Clinical Specialist, Co-Director of Antimicrobial Stewardship	Children's Hospital Colorado
Jason Newland, MD	Professor of Pediatrics, Diector of Antimicrobial Stewardship Program	Washington University, St. Louis Children's Hospital
Jeffrey Gerber, MD, PhD	Medical Director, Antimicrobial Stewardship Program, Associate Director for Pediatric Clinical Effectiveness	Children's Hospital of Philadelphia
Jennifer Girotto, PharmD	Associate Clinical Professor of Pharmacy Practice, Department of Pharmacy Practice	Connecticut Children's Medical Center
Matthew Kronman, MD	Associate Professor, Division of Infectious Diseases	Seattle Children's Hospital
Michael Smith, MD	Associate Professor of Pediatrics, Pediatric Infectious Diseases, Medical Director Pediatric Antimicrobial Stewardship	Duke University Medical Center
Pranita Tamma, MD	Director, Pediatric Antimicrobial Stewardship Program, Assistant Professor, Pediatrics	John Hopkins University School of Medicine
Ritu Banerjee, MD, PhD	Associate Professor of Pediatrics, Pediatric Infectious Diseases Specialist	Vanderbilt University School of Medicine

2b1.3. What were the statistical results from validity testing? (*e.g., correlation; t-test*)

Not tested statistically

2b1.4. What is your interpretation of the results in terms of demonstrating validity? (i.*e., what do the results mean and what are the norms for the test conducted?*)

Face validity of core measure data elements and measure constructs (i.e. antimicrobial groupings) demonstrated through a Delphi process at series of meetings with SME stakeholders beginning in April 2018.

2b2. EXCLUSIONS ANALYSIS.

NA ☐ no exclusions — *skip to section 2b4*

A select group of patient care locations in hospitals are excluded from the measure scope largely on clinical grounds e.g.) psychiatric wards, obstetrics and gynecology units. Antimicrobial use in these locations typically are not the primary focus of antimicrobial stewardship programs.

2b2.1. Describe the method of testing exclusions and what it tests (*describe the steps—do not just name a method; what was tested, e.g., whether exclusions affect overall performance scores; what statistical analysis was used*)

Not tested

2b2.2. What were the statistical results from testing exclusions? (*include overall number and percentage of individuals excluded, frequency distribution of exclusions across measured entities, and impact on performance measure scores*)

Does not apply

2b2.3. What is your interpretation of the results in terms of demonstrating that exclusions are needed to prevent unfair distortion of performance results? (*i.e., the value outweighs the burden of increased data collection and analysis. Note: **If patient preference is an exclusion**, the measure must be specified so that the effect on the performance score is transparent, e.g., scores with and without exclusion*)

Does not apply

2b3. RISK ADJUSTMENT/STRATIFICATION FOR OUTCOME OR RESOURCE USE MEASURES

If not an intermediate or health outcome, or PRO-PM, or resource use measure, skip to section 2b5.

2b3.1. What method of controlling for differences in case mix is used?

☐ **No risk adjustment or stratification**

☒ **Statistical risk model with** Click here to enter number of factors **risk factors**

☐ **Stratification by** Click here to enter number of categories **risk categories**

☐ **Other,** Click here to enter description

2b3.1.1 If using a statistical risk model, provide detailed risk model specifications, including the risk model method, risk factors, coefficients, equations, codes with descriptors, and definitions.

Statistical risk models were used to identify factors associated with differences in rates of antimicrobial use at the patient care location-level. Predictive models were constructed using forward stage wise Negative Binomial regression. In order to maximize objectivity in decision making and confidence in model results, two analysts independently developed predictive models for each group of antimicrobial agents. Final models developed by each analyst were compared and any differences in results were discussed among a team of analysts, statisticians, and subject matter experts. Final models are used to calculate predicted days of therapy adjusting for factors found to be statistically significantly associated with rates of antimicrobial use.

Hospital- and location-level variables were considered in predictive models.

Hospital-level factors:

- Hospital type: general acute care, critical access, children's, military, oncology, surgical, VA, women's, women's & children's
- Hospital teaching status: non-teaching, undergraduate only (medical students), graduate only (residents and/or fellows), major (medical school and post-graduate training)
- Total number of hospital beds
- Total number of hospital ICU beds
- Percentage of beds designated ICU beds- calculated as (total no. ICU beds/total no. beds)x100
- Average length of stay- calculated as (annual patient days/annual admissions)

Location-level factors:

- Patient care location: parameterized in different ways, levels with similar parameter estimates were grouped further
    - Example parameterization:

- Unique patient care locations: medical ICU vs. medical-surgical ICU vs. surgical ICU vs. medical ward vs. medical-surgical ward vs. surgical ward vs. general hematology-oncology ward vs. step-down unit
- Grouped by location type: ICU vs. ward vs. Onc vs. step-down

Significance of variables was assessed using Wald and likelihood ratio Chi-square tests at a 0.05 significance level and model fit improvement was assessed using Akaike Information Criterion and Bayesian Information Criterion.

2b3.2. If an outcome or resource use component measure is <u>not risk adjusted or stratified</u>, provide <u>rationale and analyses</u> to demonstrate that controlling for differences in patient characteristics (case mix) is not needed to achieve fair comparisons across measured entities.
Does not apply

2b3.3a. Describe the conceptual/clinical <u>and</u> statistical methods and criteria used to select patient factors (clinical factors or social risk factors) used in the statistical risk model or for stratification by risk (*e.g., potential factors identified in the literature and/or expert panel; regression analysis; statistical significance of p<0.10; correlation of x or higher; patient factors should be present at the start of care*)

The variables considered for the predictive models of antimicrobial use rates are hospital- and patient care level-characteristics that hospitals report to the National Healthcare Safety Network (NHSN). No patient-level data are reported to the NHSN Antimicrobial Use and Resistance (AUR) Module. The potential factors considered for the model were: hospital teaching status (major [medical school and post-graduate training], graduate only [residents and/or fellows], undergraduate only [medical students], not a teaching hospital); hospital bedsize; hospital ICU bedsize; percentage of ICU beds (among all beds); average length of stay (number of annual patient days divided by annual admissions); patient care location. The predictive models were constructed using forward stagewise Negative Binomial regression assessing Wald and likelihood ratio Chi-square tests at a 0.05 significance level and model fit improvement using Akaike and Bayesian Information Criterion.

No patient level factors are included in this statistical risk models. However hospital and patient care location characteristics are used because of empirical evidence of AU variation associated with these characteristics.

Patient care location is indicative of patient mix and taken into account in predictive models. Adult patient care locations assessed in predictive models include: medical, medical-surgical, and surgical ICUs and wards, general hematology-oncology wards, and step-down units. Pediatric patient care locations assessed include: medical and medical-surgical ICUs and wards, and surgical wards. There were 0 pediatric surgical ICUs reporting to the AU Option in 2017. Patient care locations eligible for inclusion in predictive models were selected through expert panel discussions.

Also discuss any "ordering" of risk factor inclusion; for example, are social risk factors added after all clinical factors? No ordering used

2b3.3b. How was the conceptual model of how social risk impacts this outcome developed? Please check all that apply:

☐ **Published literature**

☐ **Internal data analysis**

☒ **Other (please describe)** No social risk factors

2b3.4a. What were the statistical results of the analyses used to select risk factors?

First, for each group of SAAR antimicrobial agents, univariate analyses were conducted to assess each hospital- and patient care-level factor using Negative Binomial regression. Univariate analysis results were assessed using Wald and likelihood ratio Chi-square tests at a 0.05 significance level. Forward stage wise regression allowed for models to be built by incrementally using Wald and likelihood Chi-square tests at a 0.05 significance level to assess variable significance and Akaike & Bayesian Information Criteria and likelihood ratio tests to assess improvement in model fit. Best fit models were chosen by assessing Akaike and Bayesian Information Criteria along with standard errors and sample size in each level of stratification.

Adult SAAR Model Results:

BROAD SPECTRUM ANTIBACTERIAL AGENTS PREDOMINANTLY USED FOR HOSPITAL-ONSET INFECTIONS						
AIC=32426.53, BIC=32505.99						
Parameter	**Estimate**	**Standard Error**	**Wald 95% Confidence Limits**		**Wald χ2**	**χ2 P-Value**
Intercept	-2.3357	0.049	-2.432	-2.239	2260.98	<.0001
Location type						
Medical ICU	1.0084	0.044	0.923	1.094	531.59	<.0001
Medical-Surgical ICU, Surgical ICU	0.8825	0.028	0.827	0.938	982.03	<.0001
General Hematology-Oncology Ward	0.3795	0.058	0.266	0.493	43.13	<.0001
Step-down Unit	0.2197	0.031	0.158	0.281	49.33	<.0001
Medical Ward	0.0781	0.027	0.025	0.132	8.25	0.0041
Medical-Surgical Ward, Surgical Ward	REF	.	.	.	.	.
Facility type						
Veteran's Affairs	-0.1821	0.030	-0.240	-0.124	37.96	<.0001
Critical access	-0.2465	0.088	-0.418	-0.075	7.92	0.0049
Military	-0.6278	0.063	-0.751	-0.505	99.86	<.0001
Women's	-1.1920	0.328	-1.834	-0.550	13.25	0.0003
General acute care, Oncology, Surgical, Women's & Children's	REF	.	.	.	.	.
Number of ICU beds, facility-wide						
≥8	0.1734	0.048	0.079	0.268	13.07	0.0003
<8	REF	.	.	.	.	.
Average length of stay, facility-wide (in days)						
≥3.6	0.1091	0.026	0.059	0.160	18.02	<.0001
<3.6	REF	.	.	.	.	.
Medical school affiliation type						
Undergraduate only	0.1394	0.030	0.081	0.198	22.09	<.0001
None, Graduate, Major	REF	.	.	.	.	.

BROAD SPECTRUM ANTIBACTERIAL AGENTS PREDOMINANTLY USED FOR COMMUNITY-ACQUIRED INFECTIONS

AIC=31167.69, BIC=31241.48

Parameter	Estimate	Standard Error	Wald 95% Confidence Limits		Wald χ2	χ2 P-Value
Intercept	-3.9491	0.182	-4.305	-3.593	472.89	<.0001
Location type						
Medical ICU, Medical Ward, General Hematology-Oncology Ward	0.3598	0.026	0.310	0.410	197.15	<.0001
Medical-Surgical ICU, Medical-Surgical Ward	0.2938	0.025	0.245	0.343	136.88	<.0001
Step-down Unit	0.2083	0.030	0.149	0.268	46.93	<.0001
Surgical ICU, Surgical Ward	REF	.	.	.	.	.
Facility type						
Critical access, General acute care, Oncology	1.5380	0.180	1.186	1.890	73.19	<.0001
Surgical, Veteran's Affairs	1.2805	0.182	0.924	1.638	49.42	<.0001
Military	1.0781	0.187	0.712	1.444	33.36	<.0001
Womens, Womens Childrens	REF	.	.	.	.	.
Average length of stay, facility-wide (in days)						
<4.5	0.1714	0.022	0.129	0.214	63.64	<.0001
4.5 - 5.1	0.1128	0.023	0.068	0.158	23.95	<.0001
≥5.2	REF	.	.	.	.	.
Number of hospital beds, facility-wide						
<135	0.2505	0.025	0.201	0.300	99.39	<.0001
135 - 330	0.1545	0.019	0.117	0.192	65.90	<.0001
≥331	REF	.	.	.	.	.
ICU beds (as a percentage of total beds)						
<7.6%	0.1198	0.025	0.071	0.169	23.20	<.0001
≥7.6%	REF	.	.	.	.	.

ANTIBACTERIAL AGENTS PREDOMINANTLY USED FOR RESISTANT GRAM-POSITIVE INFECTIONS (e.g., MRSA)

AIC=30225.82, BIC=30271.22

Parameter	Estimate	Standard Error	Wald 95% Confidence Limits		Wald χ2	χ2 P-Value
Intercept	-4.0018	0.200	-4.393	-3.611	402.38	<.0001
Location type						
Medical ICU, Medical-Surgical ICU, Surgical ICU	0.8382	0.032	0.775	0.902	667.29	<.0001
Med Ward, Med-Surg Ward, General Hematology-Oncology, Step-down	0.1443	0.029	0.088	0.201	24.78	<.0001
Surgical Ward	REF	.	.	.	.	.
Facility type						
Critical access, General acute care, Oncology, Surgical, Veteran's Affairs	1.1291	0.195	0.748	1.510	33.69	<.0001
Military	0.7007	0.202	0.305	1.097	12.02	0.0005
Women's, Women's & Children's	REF	.	.	.	.	.
Number of hospital beds, facility-wide						
≥66	0.1619	0.036	0.091	0.233	19.98	<.0001
<66	REF	.	.	.	.	.
Average length of stay, facility-wide (in days)						
≥3.3	0.1913	0.027	0.139	0.244	51.19	<.0001
<3.3	REF	.	.	.	.	.

NARROW SPECTRUM BETA-LACTAM AGENTS

AIC=30014.69, BIC=30071.45

Parameter	Estimate	Standard Error	Wald 95% Confidence Limits		Wald χ2	χ2 P-Value
Intercept	-3.2228	0.070	-3.361	-3.085	2101.80	<.0001
Location type						
Surgical ICU, Surgical Ward	1.1285	0.068	0.995	1.262	276.13	<.0001
Medical-Surgical ICU, Medical-Surgical Ward	0.5004	0.064	0.374	0.626	60.55	<.0001
Step-down Unit	0.2857	0.068	0.152	0.420	17.51	<.0001
Medical ICU, Medical Ward	0.2145	0.065	0.087	0.342	10.85	0.001
General Hematology-Oncology Ward	REF	.	.	.	.	.
Average length of stay, facility-wide (in days)						
<3.5	0.2612	0.030	0.202	0.320	75.22	<.0001
≥5.8	0.1726	0.034	0.107	0.238	26.40	<.0001
3.5 - 5.7	REF	.	.	.	.	.
ICU beds (as a percentage of total beds)						
≥8.6%	0.1633	0.030	0.105	0.221	30.48	<.0001
<8.6%	REF	.	.	.	.	.
Number of hospital beds, facility-wide						
≥222	0.1112	0.023	0.066	0.156	23.37	<.0001
<222	REF	.	.	.	.	.

ANTIFUNGAL AGENTS PREDOMINANTLY USED FOR INVASIVE CANDIDIASIS

AIC=25043.86, BIC=25128.99

Parameter	Estimate	Standard Error	Wald 95% Confidence Limits		Wald χ2	χ2 P-Value
Intercept	-6.7391	0.359	-7.443	-6.036	352.37	<.0001
Location type						
Surgical ICU, General Hematology-Oncology Ward	1.2644	0.060	1.146	1.382	440.72	<.0001
Medical ICU, Medical-Surgical ICU	0.8993	0.036	0.828	0.971	609.34	<.0001
Step-down Unit	0.1635	0.043	0.079	0.248	14.45	0.0001
Medical Ward, Medical-Surgical Ward, Surgical Ward	REF	.	.	.	.	.
Number of ICU beds, facility-wide						
≥78	0.7605	0.072	0.620	0.901	113.16	<.0001
15 - 77	0.4798	0.066	0.351	0.608	53.60	<.0001
8 - 14	0.3311	0.071	0.192	0.471	21.62	<.0001
<8	REF	.	.	.	.	.
Number of hospital beds, facility-wide						
176 - 306	0.1611	0.035	0.092	0.231	20.69	<.0001
<176 or ≥307	REF	.	.	.	.	.
Facility type						
Oncology	3.5368	0.592	2.377	4.696	35.74	<.0001
Critical access, General acute care, Surgical	1.9238	0.355	1.229	2.619	29.42	<.0001
Military, Veteran's Affairs	1.5391	0.358	0.838	2.240	18.53	<.0001
Women's, Women's & Children's	REF	.	.	.	.	.
Average length of stay, facility-wide (in days)						
≥5.2	0.3836	0.055	0.276	0.492	48.45	<.0001
4.5 - 5.1	0.2550	0.055	0.148	0.362	21.86	<.0001
3.0 - 4.4	0.1685	0.052	0.067	0.270	10.56	0.0012
<3.0	REF	.	.	.	.	.

COMPLEMENTARY AGENTS (aka Other antibacterials)

AIC=30919.59, BIC=30970.67

Parameter	Estimate	Standard Error	Wald 95% Confidence Limits		Wald χ2	χ2 P-Value
Intercept	-2.5547	0.053	-2.658	-2.451	2343.21	<.0001
Location type						
Medical ICU	0.4980	0.039	0.422	0.574	163.19	<.0001
Medical-Surgical ICU	0.3324	0.030	0.273	0.392	120.09	<.0001
Surgical ICU, Medical Ward, Medical-Surgical Ward	0.2152	0.025	0.166	0.265	73.30	<.0001
General Hematology-Oncology Ward, Step-down Unit	0.1477	0.030	0.090	0.205	25.14	<.0001
Surgical Ward	REF	.	.	.	.	.
Average length of stay, facility-wide (in days)						
<3	0.1850	0.025	0.136	0.234	54.60	<.0001
≥3	REF	.	.	.	.	.
Facility type						
Critical access, General, Oncology, Surgical, Women's, Women's & Children's	0.3758	0.049	0.281	0.471	59.71	<.0001
Veteran's Affairs	0.2708	0.052	0.168	0.374	26.75	<.0001
Military	REF	.	.	.	.	.

ANTIBACTERIAL AGENTS POSING THE HIGHEST RISK FOR CDI

AIC=32174.91, BIC=32237.35

Parameter	Estimate	Standard Error	Wald 95% Confidence Limits		Wald χ2	χ2 P-Value
Intercept	-2.6749	0.056	-2.784	-2.566	2323.70	<.0001
Location type						
Medical ICU, Medical-Surgical ICU, General Hematology-Oncology Ward	0.4922	0.028	0.437	0.547	309.10	<.0001
Medical Ward	0.3392	0.029	0.283	0.395	141.54	<.0001
Surgical ICU, Medical-Surgical Ward, Step-down Unit	0.2687	0.026	0.218	0.320	107.26	<.0001
Surgical Ward	REF	.	.	.	.	.
Facility type						
Critical access, General acute care, Oncology, Surgical	0.5005	0.048	0.407	0.594	109.63	<.0001
Veteran's Affairs	0.2427	0.052	0.141	0.345	21.85	<.0001
Military, Women's, Women's & Children's	REF	.	.	.	.	.
Medical school affiliation type						
None, undergraduate, graduate	0.0810	0.018	0.046	0.116	21.04	<.0001
Major	REF	.	.	.	.	.
Average length of stay, facility-wide (in days)						
<5.2	0.0773	0.019	0.039	0.115	15.87	<.0001
≥5.2	REF	.	.	.	.	.
Number of hospital beds, facility-wide						
<442	0.1469	0.022	0.105	0.189	46.73	<.0001
≥442	REF	.	.	.	.	.
Number of ICU beds, facility-wide						
<15	0.0628	0.020	0.023	0.103	9.60	0.0019
≥15	REF	.	.	.	.	.

Pediatric SAAR Model Results:

BROAD SPECTRUM ANTIBACTERIAL AGENTS PREDOMINANTLY USED FOR HOSPITAL-ONSET INFECTIONS

AIC=2240.68, BIC=2256.36

Parameter	Estimate	Standard Error	Wald 95% Confidence Limits		Wald χ2	χ2 P-Value
Intercept	-3.0042	0.160	-3.317	-2.691	354.40	<.0001
ICU beds (as a percentage of total beds)						
≥16.6%	0.2954	0.121	0.059	0.532	5.99	0.0144
<16.6%	REF	.	.	.	.	.
Location/Facility type combination						
Med-Surg ICUs in Children's, General, Military, Women's & Children's hospitals ; Med Wards in Children's hospitals	0.7558	0.182	0.400	1.112	17.30	<.0001
Med-Surg Wards in Children's, General, Military, Women's & Children's hospitals	0.4056	0.164	0.084	0.727	6.11	0.0134
Med ICUs in General acute care hospitals ; Med Wards in General, Military, Women's & Children's hospitals ; Surgical Wards in Children's, General acute care hospitals	REF	.	.	.	.	.

BROAD SPECTRUM ANTIBACTERIAL AGENTS PREDOMINANTLY USED FOR COMMUNITY-ACQUIRED INFECTIONS

AIC=2234.81, BIC=2250.49

Parameter	Estimate	Standard Error	Wald 95% Confidence Limits		Wald χ2	χ2 P-Value
Intercept	-2.8862	0.118	-3.118	-2.655	596.67	<.0001
Number of hospital beds, facility-wide						
<450	0.4758	0.081	0.318	0.634	34.87	<.0001
≥450	REF	.	.	.	.	.
Facility Type						
General acute care, Womens Childrens	0.5476	0.091	0.369	0.726	36.23	<.0001
Childrens, Military	REF	.	.	.	.	.
Location type						
Medical-Surgical ICU, Medical ICU	0.2629	0.075	0.116	0.410	12.28	0.0005
Medical-Surgical Ward, Medical Ward, Surgical Ward	REF	.	.	.	.	.

NARROW SPECTRUM BETA-LACTAM AGENTS

AIC=2116.13, BIC=2128.67

Parameter	Estimate	Standard Error	Wald 95% Confidence Limits		Wald χ2	χ2 P-Value
Intercept	-2.4703	0.043	-2.555	-2.386	3261.86	<.0001
Location/Facility type combination						
Med-Surg ICUs in Children's hospitals ; Surgical Wards in Children's, General acute care hospitals	0.6571	0.132	0.399	0.915	24.98	<.0001
Med and Med-Surg Wards in Children's, General, Military, Women's & Children's hospitals ; Med-Surg ICUs in General, Military, Women's & Children's hospitals ; Med ICUs in General acute care hospitals	REF	.	.	.	.	.
Number of hospital beds, facility-wide						
<204 or ≥450	0.3161	0.063	0.193	0.440	25.19	<.0001
204 - 449	REF	.	.	.	.	.

ANTIBACTERIAL AGENTS PREDOMINANTLY USED FOR RESISTANT GRAM-POSITIVE INFECTIONS (e.g., MRSA)

AIC=2165.44, BIC=2174.85

Parameter	Estimate	Standard Error	Wald 95% Confidence Limits		Wald χ2	χ2 P-Value
Intercept	-2.5518	0.052	-2.653	-2.451	2436.92	<.0001
ICU location	0.4608	0.095	0.274	0.648	23.35	<.0001

AZITHROMYCIN

AIC=1834.04, BIC=1855.99

Parameter	Estimate	Standard Error	Wald 95% Confidence Limits		Wald χ2	χ2 P-Value
Intercept	-6.2324	0.388	-6.993	-5.472	258.24	<.0001
Location type						
Medical-Surgical ICU, Medical ICU	2.3497	0.387	1.590	3.109	36.79	<.0001
Medical-Surgical Ward, Medical Ward	1.9694	0.380	1.224	2.715	26.82	<.0001
Surgical Ward	REF	.	.	.	.	.
Number of hospital beds, facility-wide						
<204	0.3832	0.185	0.020	0.746	4.28	0.0385
204 - 276	1.3460	0.187	0.980	1.712	51.91	<.0001
277 - 449	0.8030	0.149	0.512	1.094	29.23	<.0001
≥450	REF	.	.	.	.	.

ANTIFUNGAL AGENTS PREDOMINANTLY USED FOR INVASIVE CANDIDIASIS

AIC=1503.42, BIC=1515.94

Parameter	Estimate	Standard Error	Wald 95% Confidence Limits		Wald χ2	χ2 P-Value
Intercept	-5.5790	0.139	-5.852	-5.306	1608.54	<.0001
ICU beds (as a percentage of total beds)						
≥16.6%	0.9569	0.171	0.621	1.292	31.25	<.0001
<16.6%	REF	.	.	.	.	.
Location type						
Medical-Surgical ICU, Medical ICU	1.1361	0.174	0.795	1.477	42.61	<.0001
Medical-Surgical Ward, Medical Ward, Surgical Ward	REF	.	.	.	.	.

COMPLEMENTARY AGENTS (aka Other antibacterials)

AIC=2049.02, BIC=2064.69

Parameter	Estimate	Standard Error	Wald 95% Confidence Limits		Wald χ2	χ2 P-Value
Intercept	-3.3383	0.087	-3.509	-3.168	1467.30	<.0001
ICU beds (as a percentage of total beds)						
≥16.6%	0.6375	0.087	0.468	0.808	54.01	<.0001
<16.6%	REF	.	.	.	.	.
Average length of stay, facility-wide (in days)						
≥4.5	0.2378	0.087	0.068	0.408	7.49	0.0062
<4.5	REF	.	.	.	.	.
Number of hospital beds, facility-wide						
≥450	0.2367	0.097	0.047	0.426	5.99	0.0144
<450	REF	.	.	.	.	.

ANTIBACTERIAL AGENTS POSING THE HIGHEST RISK FOR CDI AIC=2379.54, BIC=2398.36						
Parameter	Estimate	Standard Error	Wald 95% Confidence Limits		Wald χ2	χ2 P-Value
Intercept	-2.2753	0.118	-2.507	-2.044	371.36	<.0001
Location type						
Medical-Surgical ICU, Medical ICU	0.2711	0.081	0.113	0.429	11.33	0.0008
Medical-Surgical Ward, Medical Ward, Surgical Ward	REF	.	.	.	.	.
Number of hospital beds, facility-wide						
<386	0.3145	0.079	0.159	0.470	15.72	<.0001
>=386	REF	.	.	.	.	.
Average length of stay, facility-wide (in days)						
4.1 - 4.7	0.2139	0.083	0.051	0.377	6.65	0.0099
<4.1 or ≥4.8	REF	.	.	.	.	.
Facility type						
General acute care, Women's & Children's	0.3656	0.100	0.170	0.562	13.36	0.0003
Children's, Military	REF	.	.	.	.	.

2b3.4b. Describe the analyses and interpretation resulting in the decision to select social risk factors *(e.g. prevalence of the factor across measured entities, empirical association with the outcome, contribution of unique variation in the outcome, assessment of between-unit effects and within-unit effects.)* **Also describe the impact of adjusting for social risk (or not) on providers at high or low extremes of risk.**

Did not incorporate social risk factors due to the absence of data demonstrating that these factors should be included in AU predictive models.

2b3.5. Describe the method of testing/analysis used to develop and validate the adequacy of the statistical model or stratification approach (*describe the steps—do not just name a method; what statistical analysis was used*)

See 2b3.4a

Risk adjustment is done with a statistical risk model. The variables considered for the predictive models of antimicrobial use rates are hospital- and patient care level-characteristics that hospitals report to the National Healthcare Safety Network (NHSN). No patient-level data are reported to the NHSN Antimicrobial Use and Resistance (AUR) Module. Adult and pediatric patient care locations and antimicrobial agent groups were modeled separately to further control for differences in use across different locations and patient populations. The predictive models were constructed using forward stagewise Negative Binomial regression assessing Wald and likelihood ratio Chi-square tests at a 0.05 significance level and model fit improvement using Akaike and Bayesian Information Criterion.

First, for each group of SAAR antimicrobial agents, univariate analyses were conducted to assess each hospital- and patient care-level factor using Negative Binomial regression. Univariate analysis results were assessed using Wald and likelihood ratio Chi-square tests at a 0.05 significance level. Forward stagewise regression allowed for models to be built by incrementally using Wald and likelihood Chi-square tests at a 0.05 significance level to assess variable significance and Akaike & Bayesian Information Criteria and likelihood ratio tests to assess improvement in model fit. Best fit models were chosen by assessing Akaike and Bayesian Information Criteria along with standard errors and sample size in each level of stratification.

Models with lowest AIC and/or BIC were selected, data was tested for influential observations, and sample sizes of stratification groups were assessed. NHSN required sufficient sample size across strata to evaluate any particular stratification that resulted in any parameterizations of the final models.

Bootstrap validation was conducted to validate performance of predictive models selected for each group of adult and pediatric antimicrobial agent categories.

Provide the statistical results from testing the approach to controlling for differences in patient characteristics (case mix) below.
If stratified, skip to 2b3.9

2b3.6. Statistical Risk Model Discrimination Statistics (*e.g., c-statistic, R-squared*)**:**

Likelihood Ratio Test, Akaike and Bayesian Information Criterion.

Adult Model		Akaike Information Criterion (AIC)	Bayesian Information Criterion (BIC)	Likelihood Ratio Test^ Statistic (p-value)
Broad spectrum antibacterial agents predominantly used for hospital-onset infections	Intercept only model	33735.30	33746.65	--
	Final model	32426.53	32505.99	1332.8 (<0.001)
Broad spectrum antibacterial agents predominantly used for community-acquired infections	Intercept only model	31756.13	31767.48	--
	Final model	31167.69	31241.48	610.4 (<0.001)
Antibacterial agents predominantly used for resistant Gram-positive infections (e.g., MRSA)	Intercept only model	31325.34	31336.69	--
	Final model	30225.82	30271.22	1111.5 (<0.001)
Narrow-spectrum beta-lactam agents	Intercept only model	30824.46	30835.81	--
	Final model	30014.69	30071.45	825.8 (<0.001)
Antifungal agents predominantly used for invasive candidiasis	Intercept only model	26260.25	26271.60	--
	Final model	25043.86	25128.99	1242.4 (<0.001)
Antibacterial agents posing the highest risk for CDI	Intercept only model	32801.06	32812.42	--
	Final model	32174.91	32237.35	644.1 (<0.001)
Complementary antibacterial agents not found in other mutually exclusive SAAR groups	Intercept only model	31223.64	31234.99	--
	Final model	30919.59	30970.67	318.1 (<0.001)
^Final model compared to intercept only model				

Pediatric Model		Akaike Information Criterion (AIC)	Bayesian Information Criterion (BIC)	Likelihood Ratio Test^ Statistic (p-value)
Broad spectrum antibacterial agents predominantly used for hospital-onset infections	Intercept only model	2263.49	2269.76	--
	Final model	2240.68	2256.36	28.8 (<0.001)
Broad spectrum antibacterial agents predominantly used for community-acquired infections	Intercept only model	2279.84	2286.11	--
	Final model	2234.81	2250.49	51.0 (<0.001)
Antibacterial agents predominantly used for resistant Gram-positive infections (e.g., MRSA)	Intercept only model	2186.58	2192.85	--
	Final model	2165.44	2174.85	23.1 (<0.001)
Narrow-spectrum beta-lactam agents	Intercept only model	2157.71	2163.98	--
	Final model	2116.13	2128.67	45.6 (<0.001)
Azithromycin	Intercept only model	1896.47	1902.74	--
	Final model	1834.04	1855.99	72.4 (<0.001)
Antifungal agents predominantly used for invasive candidiasis	Intercept only model	1578.00	1584.26	--
	Final model	1503.42	1515.94	78.6 (<0.001)
Antibacterial agents posing the highest risk for CDI	Intercept only model	2407.35	2413.62	--
	Final model	2379.54	2398.36	35.8 (<0.001)
Complementary antibacterial agents not found in other mutually exclusive SAAR groups	Intercept only model	2100.82	2107.09	--
	Final model	2049.02	2064.69	57.8 (<0.001)
^Final model compared to intercept only model				

2b3.7. Statistical Risk Model Calibration Statistics (*e.g., Hosmer-Lemeshow statistic*):

Likelihood Ratio Test, Akaike and Bayesian Information Criterion. See 2b3.6. for results.

2b3.8. Statistical Risk Model Calibration – Risk decile plots or calibration curves:

All Likelihood Ratio Tests for the best models indicated significant improvement as well as the lowest Akaike and/or Bayesian Information Criterion values.

2b3.9. Results of Risk Stratification Analysis:

Does not apply.

2b3.10. What is your interpretation of the results in terms of demonstrating adequacy of controlling for differences in patient characteristics (case mix)? (i.*e., what do the results mean and what are the norms for the test conducted*)

The models are the best models using all facility and location characteristics available for analysis based on Likelihood Ratio Tests, Akaike and/or Bayesian Information Criterion. Differences in patient characteristics were addressed by including patient care location characteristics in the model where appropriate. Each antimicrobial agent group was modeled separately to further control for differences in use across different locations.

2b3.11. Optional Additional Testing for Risk Adjustment (*not required, but would provide additional support of adequacy of risk model, e.g., testing of risk model in another data set; sensitivity analysis for missing data; other methods that were assessed*)

Not Applicable

2b4. IDENTIFICATION OF STATISTICALLY SIGNIFICANT & MEANINGFUL DIFFERENCES IN PERFORMANCE

2b4.1. Describe the method for determining if statistically significant and clinically/practically meaningful differences in performance measure scores among the measured entities can be identified (*describe the steps—do not just name a method; what statistical analysis was used? Do not just repeat the information provided related to performance gap in 1b)*

The models calculated the predicted number of antimicrobial days. These predicted numbers were summed to the appropriate level (either adult ICUs, adult wards, pediatric ICUs, pediatric wards, all adult ward and ICU locations, all pediatric ward and ICU locations). The Standardized Antimicrobial Administration Ratio (SAAR) and confidence interval were calculated as: reported number of antimicrobial days/predicted number of antimicrobial days. If observed and predicted antimicrobial days are <=100, the mid-p exact test is used to calculate 95% confidence intervals around SAAR values and calculate p-values by comparing SAAR values to a SAAR of 1. If observed and predicted antimicrobial days are >100, the Byar approximation method is used, assuming identical results with mid-P, based on the large sample theory, to calculate 95% confidence intervals and p-values for each SAAR value.

The SAAR is a summary metric that compares observed days of therapy to predicted days of therapy. A SAAR=1 indicates that antimicrobial use is equivalent to the referent group's antimicrobial use. A SAAR that is less than 1 may indicate antimicrobial underuse and a SAAR that is greater than 1 may indicate antimicrobial overuse. A SAAR alone, however, is not a definitive measure of the appropriateness or judiciousness of antimicrobial use and any SAAR value may warrant further investigation. Also, a SAAR that is statistically different from 1 does not mean that further investigation will be productive in identifying opportunities for improvement in antimicrobial use.

2b4.2. What were the statistical results from testing the ability to identify statistically significant and/or clinically/practically meaningful differences in performance measure scores across measured entities? (e.g., *number and percentage of entities with scores that were statistically significantly different from mean or some benchmark, different from expected; how was meaningful difference defined*)

A meaningful difference in the SAAR was defined as a SAAR and a confidence interval that was statistically different from 1. See attached Table 3. NHSN Antimicrobial Use Measure Proposal for the number of patient care locations that had a statistically significant higher SAAR and the number of patient care locations that a statistically significant lower SAAR for each antimicrobial agent group.

First, for each group of SAAR antimicrobial agents, univariate analyses were conducted to assess each hospital- and patient care-level factor using Negative Binomial regression. Univariate analysis results were assessed using Wald and likelihood ratio Chi-square tests at a 0.05 significance level. Forward stagewise regression allowed for models to be built by incrementally using Wald and likelihood Chi-square tests at a 0.05 significance level to assess variable significance and Akaike & Bayesian Information Criteria and likelihood ratio tests to assess improvement in model fit. Best fit models were chosen by assessing Akaike and Bayesian Information Criteria along with standard errors and sample size in each level of stratification. Models with lowest AIC and/or BIC were selected, data was tested for influential observations, and sample sizes of stratification groups were assessed. NHSN required sufficient sample size across strata to evaluate any particular stratification that resulted in any parameterizations of the final models.

2b4.3. What is your interpretation of the results in terms of demonstrating the ability to identify statistically significant and/or clinically/practically meaningful differences in performance across measured entities? (i.e., *what do the results mean in terms of statistical and meaningful differences?*)

SAAR calculations for each combination of patient care locations and antimicrobials provide summary results, which if statistically significant (i.e., statistically higher or lower than 1.0) are an indicator of possible overuse or underuse of antimicrobials. In practical terms, statistically significant SAAR results provide signals that

warrant further evaluation, which can include medication use evaluations by antimicrobial stewardship programs (ASPs) for specific antimicrobial agents included in a SAAR antimicrobial category. While SAAR results that achieve statistical significance are not a definitive measure of appropriateness or judiciousness of antimicrobial use, ASPs use SAAR values to identify priorities for further analysis and possible action, and ASPs use SAAR values as a way to gauge impact of ASP interventions.

2b5. COMPARABILITY OF PERFORMANCE SCORES WHEN MORE THAN ONE SET OF SPECIFICATIONS

If only one set of specifications, this section can be skipped. Does not apply

Note: *This item is directed to measures that are risk-adjusted (with or without social risk factors)* ***OR*** *to measures with more than one set of specifications/instructions (e.g., one set of specifications for how to identify and compute the measure from medical record abstraction and a different set of specifications for claims or eMeasures). It does not apply to measures that use more than one source of data in one set of specifications/instructions (e.g., claims data to identify the denominator and medical record abstraction for the numerator).* ***Comparability is not required when comparing performance scores with and without social risk factors in the risk adjustment model. However, if comparability is not demonstrated for measures with more than one set of specifications/instructions, the different specifications (e.g., for medical records vs. claims) should be submitted as separate measures.***

2b5.1. Describe the method of testing conducted to compare performance scores for the same entities across the different data sources/specifications (*describe the steps—do not just name a method; what statistical analysis was used*)

2b5.2. What were the statistical results from testing comparability of performance scores for the same entities when using different data sources/specifications? (*e.g., correlation, rank order*)

2b5.3. What is your interpretation of the results in terms of the differences in performance measure scores for the same entities across the different data sources/specifications? (i.*e., what do the results mean and what are the norms for the test conducted*)
Does not apply

2b6. MISSING DATA ANALYSIS AND MINIMIZING BIAS

2b6.1. Describe the method of testing conducted to identify the extent and distribution of missing data (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias (*describe the steps—do not just name a method; what statistical analysis was used*)
Data files submitted with missing data elements are rejected electronically. Corrected data files are then resubmitted and accepted when complete.

2b6.2. What is the overall frequency of missing data, the distribution of missing data across providers, and the results from testing related to missing data? (*e.g., results of sensitivity analysis of the effect of various rules for missing data/nonresponse; if no empirical sensitivity analysis, identify the approaches for handling missing data that were considered and pros and cons of each*)
AU numerator and denominator must be complete before the data files can be accepted into NHSN and used in the analyses.

2b6.3. What is your interpretation of the results in terms of demonstrating that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the

specified handling of missing data minimizes bias**?** (i.*e., what do the results mean in terms of supporting the selected approach for missing data and what are the norms for the test conducted; if no empirical analysis, provide rationale for the selected approach for missing data*)

Does not apply

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.

Generated or collected by and used by healthcare personnel during the provision of care (e.g., blood pressure, lab value, diagnosis, depression score)

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 a combination of electronic sources

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.

Use of electronic data sources for numerator and denominator data collection has proven feasible across multiple hospital settings; data are routinely available as a byproduct of electronic medication administration record keeping at the point of care and can be reported in a timely manner. Data for antimicrobial days and days present in specified patient care locations are reported as sums for all patients in those locations. i.e., 100% sample. Patient-identifiable data is not reported (aggregate data only), hence the risk of a breach of patient confidentiality is extremely low. Upfront implementation costs and technical challenges are the main operational issues for initial data collection and reporting; costs and level of effort vary across settings. Technical assistance provided by CDC facilitates implementation.

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)*.

Does not apply--no fees, license, or other requirements

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)
	Public Health/Disease Surveillance National Healthcare Safety Network http://www.cdc.gov/nhsn/acute-care-hospital/aur/index.html National Healthcare Safety Network http://www.cdc.gov/nhsn/acute-care-hospital/aur/index.html Quality Improvement (Internal to the specific organization) National Healthcare Safety Network http://www.cdc.gov/nhsn/acute-care-hospital/aur/index.html

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

National Healthcare Safety Network (NHSN), Centers for Disease Control and Prevention
NHSN is the system used by CDC and its partners in clinical care and public health for surveillance of healthcare-associated infections, healthcare worker safety, blood safety, antimicrobial use and resistance, and adherence to prevention practices. The system is designed to provide actionable data for healthcare facilities and systems, public health agencies at the state and federal levels, and prevention collaboratives. NHSN is the data source for multiple NQF-endorsed measures for which CDC reports measure results on behalf of

healthcare facilities to the Centers for Medicare and Medicaid Services (CMS) quality measurement reporting programs.

Percentage of hospitals in each state/territory submitting at least one month of AU data to NHSN and gaining access to national AU benchmarks provided via NHSN (as of March 7, 2019):

AL	4
AK	25
AZ	5
AR	13
CA	20
CO	33
CT	14
DC	6
DE	29
FL	22
GA	18
HI	6
ID	9
IL	20
IN	32
IA	7
KS	9
KY	25
LA	11
ME	7
MD	13
MA	9
MI	19
MN	8
MS	4
MO	36
MT	17
NE	5
NV	12
NH	6
NJ	7
NM	7
NY	21
NC	16
ND	7
OH	18
OK	9
OR	30
PA	6
PR	0
RI	25
SC	22
SD	21
TN	17
TX	11
UT	49

VT 11
VA 44
WA 15
WV 10
WI 18
WY 0

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?*)
The continued use of this measure is for public health/disease surveillance, quality improvement with benchmarking (external benchmarking to multiple organizations), and quality improvement (internal to the specific organization). Voluntary participation in reporting AU data to NHSN in accordance with the NHSN AU measure specification has increased to over a thousand hospitals nationwide. These hospitals are deriving benefits from the measure for patient care practices without use of the measure data in public reporting or payment programs. Lessons learned from this increasing field use of the measure, coupled with further development of the measure, specifically the predictive models, will enable use of the measure for accountability purposes. The measure steward, NHSN, seeks to add to predictive models data about infectious disease burden and use of antimicrobials for prophylaxis. These are strong predictors of AU, and NHSN seeks to include them in predictive models.

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.*)

NHSN serves as a national data aggregating system for AU and engages with multiple antimicrobial stewardship programs that use of AU data for stewardship purposes on a voluntary basis. The continuing growth in AU reporting to NHSN —a greater than five-fold increase in hospital participation since NQF initially endorsed the NHSN AU measure —is indicative of the measure´s value even without an external accountability application. As a result of this increased participation in AU reporting, much more AU data was available for NHSN to develop AU predictive models used in this measure proposal than were used in the initial proposal. Additional data, e.g., extent of infectious disease burden and indications for antimicrobial prophylaxis, are candidates for additions to NHSN´s AU predictive models. NHSN is working to identify or develop sources for these additional data, and will apply this work and work products in the next iteration of its AU predictive models. NHSN also continues to work with hospitals and healthcare systems that report AU data to NHSN to further evaluate the measure's usefulness for antimicrobial stewardship programs and to refine the measure as needed to improve its value for assessing variation in antimicrobial use intra- and inter-organizationally.

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.

All hospitals participating in NHSN's AU surveillance have immediate access to the AU data submitted to the system and are able to use NHSN's AU analytic features to analyze their data, including analyses based on the SAAR measure. In addition, NHSN provides direct technical support to hospitals participating in AU surveillance and NHSN publishes FAQ's and other resource documents that assist with interpretation of AU results.

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.

Results are available via the NHSN application at all times and on an ongoing basis.

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.

Feedback from multiple antimicrobial stewardship programs nationwide has been positive, as evidenced by the increase in hospital participation in NHSN's AU surveillance to over 1,000 hospitals, all of which participate voluntarily in AU reporting.

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

Hospitals participating in the NHSN AU Option surveillance have used their SAARs for stewardship program purposes, including priority setting and interventions designed to improve antimicrobial use. . Many hospitals have shared their feedback and changes in stewardship practices on quarterly NHSN AU Option users calls, in conference abstracts and presentations and formal journal publications. Further, feedback was received via conference call and email during the remodeling process.

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

Increasing numbering of state health departments (SHD) have gained access to AU data reported by hospitals in their jurisdictions to NHSN and these states are using the AU data to guide statewide antimicrobial stewardship methods.

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.

AMS programs participating in NHSN's AU surveillance have provided feedback that NHSN has used to update the measure specifications and facilitate use of measure data via the NHSN application.

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.

Hospitals participating in the NHSN AU Option surveillance have used their SAARs for stewardship program purposes, including priority setting and interventions designed to improve antimicrobial use. Many hospitals have shared their feedback and changes in stewardship practices on quarterly NHSN AU Option users calls, in conference abstracts and presentations and formal journal publications. Further, feedback was received via conference call and email during the remodeling process.

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.

No unintended negative consequences identified or reported.

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

•Better understanding of antimicrobial use data collected in the steward's hospital.

•Better understanding of how to use EHR or 3rd party vendor software responsible for collecting/reporting AU data.

•Improving hospital workflow based on data quality issues (ex: changes in order sets to fix free text entry; allowing off label administrations to be entered into EHR correctly).

•Improvement in the quality of data captured (ex: correcting errors in how routes are mapped).

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.

No

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

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

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?

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

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.)

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.

Attachment:

Contact Information

Co.1 Measure Steward (Intellectual Property Owner): Centers for Disease Control and Prevention

Co.2 Point of Contact: Daniel, Pollock, dap1@cdc.gov, 404-639-4237-

Co.3 Measure Developer if different from Measure Steward: Centers for Disease Control and Prevention

Co.4 Point of Contact: Daniel, Pollock, dap1@cdc.gov, 404-639-4237-

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.

Measure Developer/Steward Updates and Ongoing Maintenance

Ad.2 Year the measure was first released:

Ad.3 Month and Year of most recent revision:

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

Ad.5 When is the next scheduled review/update for this measure?

Ad.6 Copyright statement:

Ad.7 Disclaimers:

Ad.8 Additional Information/Comments: