NATIONAL QUALITY FORUM

Memo

TO: Pediatric Performance Measures Standing Committee

FR: NQF Staff

RE: Post-Comment Call to Discuss Public and Member Comments

DA: May 25, 2017

Purpose of the Call

The Pediatric Performance Measures Standing Committee will meet via conference call on **Wednesday, May 31, 2017, from 2-4 pm ET**. The purpose of this call is to:

- Review and discuss comments received during the post-evaluation public and member comment period.
- Provide input on proposed responses to the post-evaluation comments.
- Review and discuss the additional materials provided by the developer for #3154, which did not reach consensus, and revote on the measure.
- Determine whether reconsideration of any measure or other course of action is warranted. If so, review and discuss the two requests for reconsideration (measures #3189 and #2816).

Due to time constraints, we will review comments by exception—i.e., we will discuss any comments Committee members flag because they disagree with the proposed responses.

Standing Committee Actions

1. Review this briefing memo and Draft Report.
2. Review and consider the full text of all comments received and the proposed responses to the post-evaluation comments.
3. Be prepared to provide feedback and input on proposed post-evaluation comment responses, in particular identify specific comments you wish placed on the agenda for full Committee discussion.
4. Review the additional materials submitted by the two developers that submitted requests for reconsideration.

Conference Call Information

Please use the following information to access the conference call line and webinar:

Speaker dial-in #: 844-729-4895 *(NO CONFERENCE CODE REQUIRED)*
Web Link: http://nqf.commpartners.com/se/Rd/Mt.aspx?520949
Registration Link: http://nqf.commpartners.com/se/Rd/Rg.aspx?520949

Background

For this project, the 23-member Standing Committee evaluated 11 newly submitted measures against NQF's standard evaluation criteria. The Committee recommended three measures for endorsement, did not reach consensus on one measure, and did not recommend seven measures.

Comments Received

NQF solicits comments on measures undergoing review in various ways and at various times throughout the evaluation process. First, NQF solicits comments on endorsed measures on an ongoing basis through the Quality Positioning System (QPS). Second, NQF solicits member and public comments prior to the evaluation of the measures via an online tool located on the project webpage. Third, NQF opens a 30-day comment period to both members and the public after measures have been evaluated by the Committee and once a report of the proceedings has been drafted.

Pre-evaluation comments

The pre-evaluation comment period was open from January 23, 2017 to February 6, 2017, for the 12 measures[1] under review. No pre-evaluation comments were received during this comment period.

Post-evaluation comments

The Draft Report was posted for member and public comment from April 11, 2017 to May 11, 2017. NQF received 11 comments from four member organizations:

Consumers – 0	Professional – 0
Purchasers – 0	Health Plans – 0
Providers – 3	QMRI – 1
Supplier and Industry – 0	Public & Community Health - 0

Comments and Their Disposition

Two major themes were identified in the post-evaluation comments, as follows:

1. Support for Committee recommendations
2. Gaps for future measure development

In addition, one measure, *#3136: GAPPS: Rate of preventable adverse events per 1,000 patient-days among pediatric inpatients*, received specific comments requiring a developer response and Committee discussion.

Theme 1 - Support for Committee recommendations

Five of the comments received offered support for the Committee's endorsement recommendations, both for decisions to recommend endorsement and not to recommend endorsement. These comments provided support for the Committee's recommendations on measures #3153, #3166, #3220, and #3221. Commenters agreed with the Committee's decision not to recommend *#3220: Ask About Parental Concerns* and *#3221: Family Centered Care*, noting that despite the clear importance of these topics, there is "difficulty in attributing outcomes within these areas to specific providers and experiences." Two commenters supported the Committee's decision to recommend *#3166: Antibiotic Prophylaxis Among Children with Sickle Cell Anemia*, and one commenter supported the Committee's recommendation to endorse *measure #3153: Continuity of Primary Care for Children with Medical Complexity*.

[1] A developer withdrew one measure before the Committee review so the final number of measures in the project is 11

Proposed Committee Response: Thank you for providing this comment.

Theme 2 – Gaps for future measure development

Several gap areas were identified during the comment period for consideration by the Committee. Specifically, commenters suggested several NQF measure gaps that could be addressed by the measure concepts listed below at the clinic/systems levels:

- The identification of a team to work together to plan and test improvements in eliciting parental strengths and needs within a practice site.
- Defining parental strengths and needs within a practice site.
- Integrating tools such as process flows, prompts, and reminders into practice flow to support the engagement of parents.
- Clinic/systems-level measures that offer more specificity about appropriate "antibiotic prophylaxis."

Proposed Committee Response: Thank you for providing this comment. These gaps have been added to the gaps list.

Measure-Specific Comments

#3136: GAPPS: Rate of preventable adverse events per 1,000 patient-days among pediatric inpatients

Two of the comments received focused on measure *#3136: GAPPS: Rate of preventable adverse events per 1,000 patient-days among pediatric inpatients*. The developer has addressed each concern separately below.

The American Academy of Pediatrics submitted questions and suggested updates intended to clarify automated triggers to increase the specificity and clarity of the measure specifications.

Developer Response to the Academy of Pediatrics:

- Trigger: Consider rewording to "Hepatotoxic medications and RISING liver enzymes (AST, ALT)"

  Thank you for the suggestion. A consideration here is that if there were not previous hepatic enzyme measurements and the first measurement showed elevated enzymes, this would need to be investigated. If this were written to only include those that are rising (therefore requiring a previous lower value), the process might miss a possible hepatoxic injury. Therefore our preference is to retain the language as "elevated."

- Please explain how "Physician orders: Abrupt medication stop" is defined in the automated trigger tool? Most medication stops are abrupt (with rare exceptions like steroid weans or PCA infusions)

  The definition in our Manual of Operations reads as follows, "An abrupt medication stop is best described as an unexpected stop or deviation from typical ordering practice (e.g., discontinuation of a recently started medication)." Since this type of clinical decision making may be challenging to automate, it is not recommended for electronic trigger review.

- Please define “Transfer to higher level of care” more specifically. Many hospitals have observation units where most patients go home but some patients are admitted to the floor (higher level of care) after a specified time.

  The definition from our Manual of Operations reads as follows, “All transfers from an acute care area to an intensive care unit or intermediate care unit (“step-up unit”) should be considered a trigger.” Therefore the scenario presented in the comment would not meet the defined criteria.

- Consider changing Pressure ulcer documentation to “>= Stage 2” instead of just stage 2.

  Thank you for identifying this discrepancy. This was a typographical error and should read as you suggest. During our testing, the reviewers were instructed to investigate exactly as is suggested by the comment, meaning all pressure injuries Stage 2 and higher and unstageable. We will edit the relevant documents to reflect this change.

- Many places will start patients on laxatives simultaneously with opioids, but patients will still get constipated. Would this qualify as a trigger, or is it only a trigger if laxatives are started after (e.g. >=24 hours after) opioids are started? Latter would be more specific, less sensitive.

  Thank you for pointing out this ambiguity. We agree that excluding cases where laxatives are introduced concurrently (<24 hours after) with opioids is reasonable. The trigger is looking at cases where laxatives were given subsequent to the initial prescription of opioids (>=24 hours after). We will edit the relevant documents to reflect this change.

- Consider adding “positive coagulase-negative staphylococcus species blood culture” as a trigger for review; per algorithm, it should have a higher than 10% rate of being a true contaminant (i.e., an adverse event).

  Thank you for this comment. Since we currently look at a more broadly based trigger (positive blood culture 48 hours after admission), all of the occurrences of the suggested trigger would be included in the trigger as written. We hesitate to insert a new trigger into the recently reviewed tool at this stage.

- Please clarify the denominator of whether a partial day counts as a day. For example, is 1.5 days = 2 days or 1.5 days? What is the start and stop time for determining LOS duration (e.g. start of: time of arrival to floor, time of admission from ED; end of: time of discharge order, time of leaving hospital?)

  Length of stay is calculated as the number of days (discharge date minus admit date). For example, a patient who arrives at 4am on May 17th and is discharged at 4pm on May 18th has a length of stay of 1.0 day. However, a patient who arrives at 10pm on May 17th and is discharged at 10am on May 19th has a length of stay of 2.0 days. Start and stop times were not used to determine length of stay duration, only admit and discharge date.

- Step 2: Line 4. Please describe whether the unit of study (whether entire hospital, division, etc.) should remain stable over time.

  Thank you for the opportunity to clarify. We would suggest that the unit of interest remain stable over time.

Another commenter from the Armstrong Institute for Patient Safety and Quality at Johns Hopkins University, commented that implementing the trigger tool might be difficult and require significant resources; this commenter also was concerned that the tool lacks validity in identifying adverse events

Developer Response to Dr. Austin of Armstrong Institute for Patient Safety and Quality at Johns Hopkins University:

We would like to thank Dr. Austin for his comments. As the measure is implemented, the resource burden, while not trivial, should be manageable while providing a great deal of benefit in terms of increased safety events identified. The primary reviewer, typically an experienced nurse, is asked to perform chart review quarterly on 60 patient records per quarter with a limit of 30 minutes per chart. This would represent a total of 30 hours per quarter or 10 hours per month or 2.5 hours a week. The secondary reviewer, typically a physician, reviews the primary reviewer's findings. Assuming a high rate of harm or 33 events per 100 admissions, this would be 20 events to review each quarter. During validation testing, our physician team required on average 4 minutes per chart to review. Therefore, the typical time burden on the secondary reviewer would be approximately 80 minutes per quarter. Based on the frequency of events and the resources required, it is our view that the benefit of this modest resource requirement would far outweigh the burden.

In regards to validity, we developed the draft trigger tool used in the GAPPS measure through the RAND/UCLA Appropriateness Method, which is a modified Delphi process.(1–3) We first compiled a set of 78 candidate triggers from a literature review of existing pediatric and adult trigger tools and input from trigger tool experts.(4–6) We then recruited nine panelists from national pediatric and patient safety organizations and asked them to rate separately the validity and feasibility of the candidate triggers on a nine-point scale (where 1 is the least valid/feasible and 9 is the most valid/feasible). A trigger was considered valid if it was judged to be reasonably likely to identify an underlying AE, indicating that harm potentially occurred. A trigger was considered feasible if it was judged likely to be accurately and consistently documented in either paper or electronic medical records as part of patient care at a wide range of hospitals, from smaller community sites to larger tertiary care centers. Applying the RAND/UCLA Appropriateness Method, we accepted triggers that had both median validity and feasibility ratings greater than or equal to seven. This approach resulted in inclusion of 54 of the initial 78 candidate triggers in the draft GAPPS trigger list.

We focused our validity testing on evaluation of how accurately and completely "typical reviewers" (i.e., clinicians who are trained in GAPPS methodology but not necessarily trigger tool experts) were able to identify preventable AEs using the measure as compared to expert reviewers. The expert reviewers had extensive experience with using trigger tools for preventable AE identification and consequently were most likely to identify preventable AEs accurately and completely. To evaluate the validity of the GAPPS measure, we assessed the performance of the National Field Test hospitals' internal reviewers relative to the performance of external expert reviewers in applying the measure (as documented in our NQF submission materials).

REFERENCES

1. Fitch K, Bernstein S, Aguilar MD, Burnand B, LaCalle JR, Lázaro P, et al. The RAND/UCLA Appropriateness Method User's Manual. Santa Monica, CA: RAND; 2001.

2. Brown B. DELPHI PROCESS: A Methodology Used for the Elicitation of Opinions of Experts. Rand Corp. 1968 Sep;1–14.

3. Sweidan M, Williamson M, Reeve JF, Harvey K, O'Neill JA, Schattner P, et al. Identification of features of electronic prescribing systems to support quality and safety in primary care using a modified Delphi process. BMC Med Inform Decis Mak. 2010 Apr 15;10(1):21.

4. Stockwell D, Bisarya H, Classen D, Kirkendall E, Landrigan C, Lemon V, et al. A trigger tool to detect harm in pediatric inpatient settings. Pediatrics. 2015;

5. Griffin FA, Resar RK. IHI Global Trigger Tool for Measuring Adverse Events (Second Edition). Institute for Healthcare Improvement; 2009. (IHI Innovation Series white paper).

6. Kirkendall ES, Kloppenborg E, Papp J, White D, Frese C, Hacker D, et al. Measuring adverse events and levels of harm in pediatric inpatients with the Global Trigger Tool. Pediatrics. 2012 Nov;130(5):e1206-1214.

Committee Action Items:

The Committee should review the comments and the developer's responses, evaluate whether the concerns have been adequately addressed and, if so, approve the proposed responses provided. Additional information may be added to the response, depending on the discussion.

Proposed Committee Response:

Thank you for providing this comment on measure #3136. The Committee discussed the measure specifications and validity during the in-person meeting. The Committee did note that that the highest possible score for reliability was a moderate, since the measure is tested at the data-element level only; the highest possible score for validity is also moderate, since validity testing is patient-level data element. Overall, the Committee determined that the measure as specified and tested offered sufficient validity for endorsement.

The Committee also will evaluate the developer's responses on the post-comment call.

Consensus Not Reached

3154: Informed Coverage

During the in-person meeting, Committee members agreed this was an important outcome to assess, but they were concerned about the measure's ability to discern differences among states due to the overlap of the 95% confidence intervals of the performance scores provided for score-level reliability testing. The Committee believed that the measure would be a useful self-assessment tool for states to improve their coverage rates, but questioned whether this measure could be used for accountability purposes. The Committee did not reach consensus on Reliability (1-H; 11-M; 9-L; 3-I).

During the post-comment call, the Committee should review and discuss the additional material provided by the developer, and then vote on both Reliability and an overall recommendation for endorsement. A memo from the developer responding to the issue of overlap, as well as other questions brought up by the Committee during the in-person meeting, is provided in Appendix A. With respect to the issue of the overlapping performance scores, the developer summarized the graph (previously provided) as follows:

- 24 of 43 states (55.8%) can be distinguished from more than 1/2 of the other states;
- 11 (25.6%) states can be distinguished from more than 2/3 of the other states;
- At each end of the spectrum (high and low performers), 3 of 43 states (7.0%) and 3 of 43 states (7.0%), respectively, can be distinguished from 3/4 of the other states.

One NQF member, the American Academy of Pediatrics (AAP), commented on #3154. AAP agrees with the intent of the measure to more accurately capture the continuity of coverage in the Medicaid program, but recommends that this measure be further validated and re-evaluated for endorsement in the future.

Developer Response:

We appreciate that the AAP agrees with the intent of our measure to more accurately capture the continuity of coverage in the Medicaid program so that states can improve coverage. The AAP suggested that our measure "requires assumptions that may not be universally accepted," without telling us which assumptions are objectionable. We would point out that with our assumptions, our results were carefully validated against the gold standard ACS (American Community Survey). Our results, in both development and validation, were superior to the current metrics of Continuity Ratio (Ku et al.) and Duration (currently used by CMS). Informed Coverage had better correlation with the ACS and less error deviation than the other metrics. See Validity Testing, Section 2b2.3, Table 2: Pearson Correlations. Also, see Validity Testing, Section 2b2.3, Table 3: Median Absolute Errors.

Action Item: After review of additional materials provided by the developer and the comment above from the developer, the initial primary discussants—Amy Houtrow, Kerri Fei, Kraig Knudsen, David Keller, and Jeff Schiff—will lead the discussion on the new reliability information during the post-comment call. The Committee will vote on reliability and the overall recommendation.

Requests for Re-consideration

3189: Rate of Emergency Department Visit Use for Children Managed for Identifiable Asthma: Visits per 100 Child-years

During the in-person meeting, the Committee concluded that the submitted testing information was insufficient to meet NQF's minimum standards and the measure did not pass Reliability (N/A-H; 1-M; 4-L; 18-I). No comments were received specific to this measure during the post-meeting commenting period. The developer submitted a request for reconsideration. A brief summary of the request for reconsideration follows and a memo from the developer outlining in detail why the measure should be reconsidered and what changes have been made since the last review are included in Appendix B. The original submission, which includes the staff preliminary analysis and Committee's original comments, is provided in Appendix C.

Developer Rationale for Reconsideration:

"At the in-person meeting, measure #3189 passed on Evidence and Gap, and was voted insufficient for Reliability. In general, the sense of the group [the Committee] at the in-person meeting was that measure #3189 is a very viable measure, but having to conform to the NQF procedure, the group required a little bit more data, which is provided herein:

1) Reliability
2) Inclusion/Exclusion
3) Pharmacy Data
4) Race Disparities
5) Data Element Validity"

Action Item: After reviewing the information provided by the developer, does the Committee wish to reconsider this measure? If so, the initial primary discussants, Karen Dorsey, Jonathan Finkelstein, Carol Stanley, and Ricardo Quinonez, will lead the discussion of each criterion, starting with Reliability, and the Committee will vote on each criterion to reach a recommendation.

2816: Appropriateness of Emergency Department Visits for Children and Adolescents with Identifiable Asthma

During the in-person, the Committee raised a number of significant concerns with the construction and testing of the measure and the appropriateness criteria; specifically the measure was tested in only one hospital making it difficult to discern meaningful differences among institutions; and not all critical data elements were tested. The Committee also noted that the measure specifications permit variable use of pharmacy data, as available. Measure #2816 did not achieve consensus on Evidence (12-Pass; 10-No Pass) and did not pass Validity (N/A-H; 1-M; 17-L; 5-I); therefore it was not recommended for endorsement. The developer used data element level validity testing, which is accepted under NQF guidance to assess both Reliability and Validity; therefore, the Committee did not vote separately on Reliability.

One comment was received for this measure from the American Academy of Allergy, Asthma and Immunology (AAAI), which supported concerns about the lack of risk adjustment brought up by the NQF Pulmonary and Critical Care Standing Committee during a previous review; the Pediatric Committee did not discuss this issue since the discussion did not progress to that aspect of validity given the other concerns. A summary of the request for reconsideration is below, and a memo from the developer outlining in detail why the measure should be reconsidered and what changes have been made since the last review are included in Appendix B. The original submission, which includes the staff preliminary analysis and Committee's original comments is provided in Appendix D.

Developer Rationale for Reconsideration:

At the in person meeting, for measure #2816, consensus was not reached for Evidence, the measure passed on Gap, and did not pass on Reliability. While the developer has requested reconsideration for both measures, they did not provide a separate, specific rationale for this measure. The developer noted that the data

provided for #3819 also informs this measure; no additional information was provided specifically related to appropriateness. The updated data are included with the information on #3189 begins on page 17 of Appendix B.

Developer Response to Comment:

"We have submitted this to the Pediatric Committee in part because of its greater sensitivities to the issues specific to children and in this case asthma in children. Nearly half of US children are covered by public health insurance programs. Equity of outcomes across race and social class is a preeminent concern and value in child health, especially for asthma. As the internationally accepted NHLBI guidelines states, "As a general rule, patients with well-controlled asthma should have:

- Few, if any, asthma symptoms.
- Few, if any, awakenings during the night caused by asthma symptoms.
- No need to take time off from school or work due to asthma.
- Few or no limits on full participation in physical activities.
- No emergency department visits.
- No hospital stays.
- Few or no side effects from asthma medicines."

Further it is not clear whether those stressors that increase asthma burden are likely to increase or to decrease the level of appropriateness of ED use for asthma. Cogent arguments can be made in either direction, or for not at all.

Measure 2816, Appropriateness of Emergency Department Visits for Children and Adolescents with Identifiable Asthma, is stratified by age, specifically the measure is reported for children ages 2-5,6-11, 12-18, and optionally, 19-21. This is because asthma and its management are related both to the child's age and stage of development. Hence comparing performance in young children is very different from performance in adolescents.

Further, this measure of appropriateness is best interpreted in light of other measures, such as the rate of ED use for asthma. High levels of appropriateness may suggest a highly functional primary care and outpatient approach to asthma. However high levels of appropriateness and high levels of utilization may together suggest that asthma outcomes form outpatient management are not as desired. Hence, this measure informs but is not dispositive without other data.

Our formal RAND style panel of national experts did not recommend risk adjustment by race or social class. They recommend stratification by the age groups indicated below. Environmental differences may produce unequal burdens on various health plans, but the field's capacity to discriminate and risk adjust in that manner is of uncertain value and such data for adjustment are neither readily available, nor is there a consensus on what and when and how to adjust for such exposure.

Establishment of asthma control should occur from an early age. Because of challenges in identifying asthma before the age of 2, we have not included this age group in our specification.

For purchasers who are interested in stratification beyond race and age we provide OPTIONAL specifications that allow them to ask health plans to incorporate additional stratification in the measure (e.g. insurance status, county rates of poverty, and rurality/urbanicity). Contracting health plans can negotiate with purchasers and other accountability agencies to demonstrate stratified performance if they so desire.

This measure requires stratification by the following age groups:

- Age 2-5 years (second birthday to the day before the 6th birthday);
- Age 6-11 years (sixth birthday to the day before the 12th birthday);
- Age 12-18 years (twelfth birthday to the day before the 18th birthday); and
- Age 19-21 years (nineteenth birthday to the day before the 21st birthday).

These age strata are to be reported distinctly and not combined for reasons noted above.

This measure has optional stratifications for the following that can be determined by the reporting agency to use all or none, as appropriate:

- Race/Ethnicity: Hispanic, Non-Hispanic Black, Non-Hispanic White; Non- Hispanic Asian/Pacific Islander, other Non-Hispanic
- Insurance type (Public, Commercial, Uninsured)
- Benefit type (if insured): HMO, PPO, Medicaid Primary Care Case Management (PCCM) Plan, Fee for Service (FFS), other relevant enrollment categories (e.g., TANF, SSI)
- Urban influence codes: Identify the Urban Influence Code or UIC. (2013 urban influence codes available at: http://www.ers.usda.gov/data-products/urban- influence-codes.aspx#.UZUvG2cVoj8). Use parent or primary caregiver's place of residence to determine UIC. State and county names can be linked or looked up directly or zip codes can be linked to county indirectly, using the Missouri Census Data Center (http://mcdc.missouri.edu/). These data will link to county or county equivalents as used in various states.
- Urban Influence Codes (UIC) have been developed by the USDA to describe levels of urbanicity and rurality. While each UIC has its own meaningful definition, some researchers choose to aggregate various codes. Well regarded schemas for aggregation of codes include Bennett and colleagues at the South Carolina Rural Research Center. Their aggregation scheme brings together Codes 1 & 2 as Urban; 3,5, & 8 as micropolitan rural; 4,6, & 7 as rural adjacent to a metro area; and 9, 10, 11, & 12 as remote rural. We acknowledge that UIC 5 (adjacent rural area) may appropriately be aggregated with 4,6,&7 as rural. Frontier health care may be approximated by analysis of the remote rural categories (UIC 9, 11 and 12). Alternatively, Gary Hart, Director of the Center for Rural Health at the University of North Dakota School of Medicine & Health Science suggests that UIC 9-12 is the best overall approach to using county level

data to study frontier health. Inclusion of UIC 8 would make the analysis more sensitive to including frontier areas but at a meaningful cost in specificity.
- Those interested in care specific to large cities may wish to aggregate the rural area and analyze UIC 1 and 2 separately.
- When stratifying by urbanicity or UIC, the reporting and accountability entities should specify clearly what if any aggregating schema was used.
- Identify the Level of Poverty in the parent or primary caregiver's county of residence. The percent of all residents in poverty by county or county equivalent are available from the US Department of Agriculture at http://www.ers.usda.gov/data-products/county-level-data-sets/download- data.aspx. Our stratification standards are based on 2011 US population data that we have analyzed with SAS 9.3. Using parent or primary caregiver's state and county of residence (or equivalent) or FIPS code, use the variable PCTPOVALL_2011 to categorize into one of 5 Strata:
 o Lowest Quartile of Poverty if percent in poverty is <=12.5%
 o Second Quartile of Poverty if percent in poverty is >12.5% and <=16.5%
 o Third Quartile of poverty if percent in poverty is >16.5% and <=20.7%
 o First Upper Quartile (75th-90th) if percent in poverty is >20.7% and <=25.7%
 o Second Upper Quartile (>90th percentile)

These classification standards may be updated by the accountability entity using more recent data if desired.

To summarize:

Appropriateness of ED visits is a new construct for pediatric asthma. As such, there are no pre-existing data to suggest a disparate burden of either appropriate or inappropriate ED visits by socioeconomic class or by health plans caring for them. The NHLBI guideline is clear in articulating the expectation that outcomes should be equally good across the general population of individuals regardless of who they are and even how severe their asthma is (obviously there are true exceptions here, but they would not be well accounted for in any risk adjustment or stratification schema that we have ever seen). The Pediatric Committee is in a better position to understand and appreciate the implications of all of this for children and to incorporate such insights into their evaluation of this measure (and the similar rate measure).

The lack of required stratifications by risk does not lead to misinterpretation of results as a potential unintended consequence if the measure is implemented. In fact, this measure is specified to give flexibility to plans and to purchasers to respond to local conditions and needs by using stratification as needed and desired to compare performance within specified strata. These are desirable attributes for child health quality measures."

Action Item:

After review of the comment received and the information provided by the developer, does the Committee wish to reconsider this measure? If so, the initial primary discussants, Ricardo Quinonez, Marlene Miller, Jeffrey Susman, and James Bost, will lead the discussion of each criterion, starting with Evidence and the Committee will vote on each criterion to reach a recommendation.

Appendix A: Additional Information on Measure #3154

Response to Comments regarding Reliability at the NQF March 2017 Meeting (Rev 4/17/17)

Dear NQF Committee:

What follows is a point-by-point response to the comments I received at the NQF session where the Informed Coverage Metric 3154 was discussed on March 2, 2017. As I recall, the measure passed all elements of the voting except for reliability. What follows is my best effort to answer each and every concern raised in the discussion.

I continue to believe that the Informed Coverage (IC) metric would be a very beneficial tool for measuring participation in eligible Medicaid and CHIP children.

1. A problem with "clustering" noted in the Reliability Figure 2. I believe the worry raised by the reviewers was that not enough states look different from one another.

In response to the concern, we provide three new sets of analyses to view the issue in different ways.

1.A. We first present the original Figure 2, (now described as Figure 1.A. in this document and associated with it is Table 1.A.1, a 24 x 24 table to describe whether a specific state's Informed Coverage estimate was statistically above or below that of each other state. Table 1.A.1 provides the states in the same order as in the original Figure 2, going from lowest IC to highest IC. Each row describes a state. Each cell in the designated state's row is either shaded gray or white. If gray, it signifies that the row state has a higher informed coverage than the column state using a two tailed $P < 0.05$ level of significance.

At the far right of the spreadsheet we provide the number of states that the column state has a higher Informed Coverage rate in column labeled "GT" for "Greater Than". To its right, in the column labeled "LT" are the number of states with Informed Coverage "Less Than" the column state. For example, the state of MD, which has IC that is greater than 14 other states, and less than 5 other states, is different from a total of 19 states (we provide a column describing how many states were different from the column, the sum of both the GT and LT columns).

Finally we provide a test or whether the column state's mean is different from the mean of all remaining states (with a P-value).

It can be seen that many states differ from many other states on their IC values. They are not all clustered together.

i. 24 of 43 states (55.8%) can be distinguished from more than 1/2 of the other states.

ii. 11 (25.6%) states can be distinguished from more than 2/3 of the other states.
iii. At each end of the spectrum (high and low performers), 3 of 43 states (7.0%) and 3 of 43 states (7.0%), respectively, can be distinguished from 3/4 of the other states.

Figure 1. Original Graph of IC in States Ranked by IC

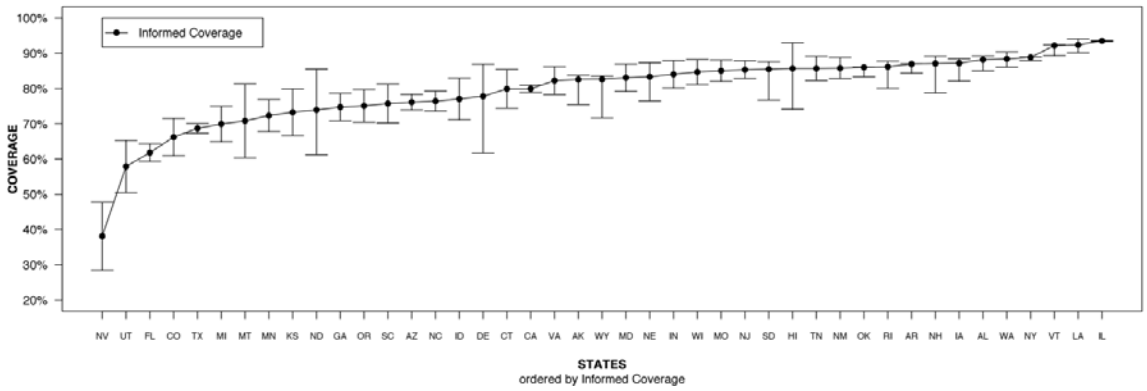
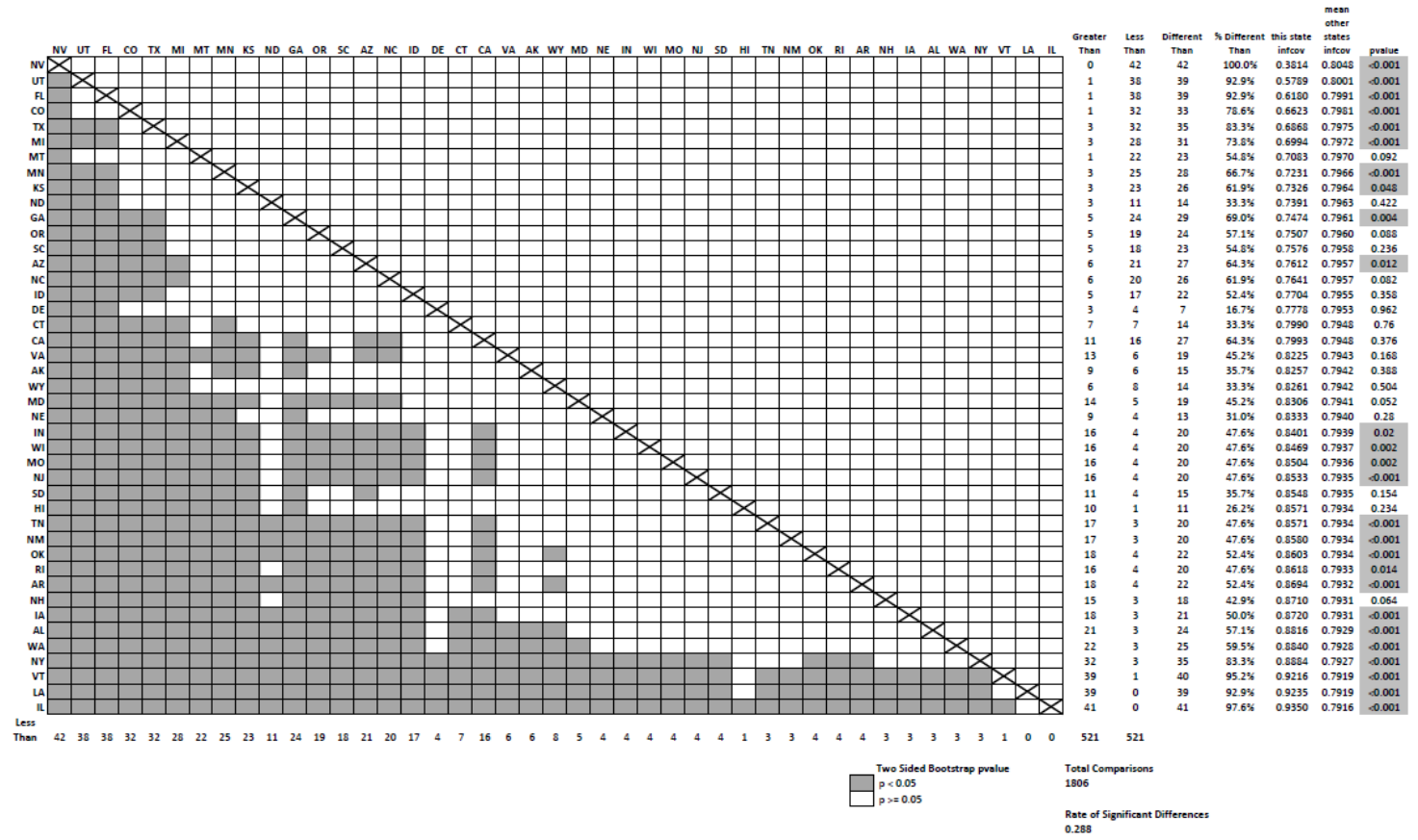


Table 1. State by State Comparison of IC in States Ranked by IC: When reading across a row, shaded squares suggest row state has higher IC than the defined column state. When reading down a column, a shaded square suggests a column state has lower IC than the row state.



	Greater Than	Less Than	Different Than	% Different Than	this state infcov	mean other states infcov	pvalue
NV	0	42	42	100.0%	0.3814	0.8048	<0.001
UT	1	38	39	92.9%	0.5789	0.8001	<0.001
FL	1	38	39	92.9%	0.6180	0.7991	<0.001
CO	1	32	33	78.6%	0.6623	0.7981	<0.001
TX	3	32	35	83.3%	0.6868	0.7975	<0.001
MI	3	28	31	73.8%	0.6994	0.7972	<0.001
MT	1	22	23	54.8%	0.7083	0.7970	0.092
MN	3	25	28	66.7%	0.7231	0.7966	<0.001
KS	3	23	26	61.9%	0.7326	0.7964	0.048
ND	3	11	14	33.3%	0.7391	0.7963	0.422
GA	5	24	29	69.0%	0.7474	0.7961	0.004
OR	5	19	24	57.1%	0.7507	0.7960	0.088
SC	5	18	23	54.8%	0.7576	0.7958	0.236
AZ	6	21	27	64.3%	0.7612	0.7957	0.012
NC	6	20	26	61.9%	0.7641	0.7957	0.082
ID	5	17	22	52.4%	0.7704	0.7955	0.358
DE	3	4	7	16.7%	0.7778	0.7953	0.962
CT	7	7	14	33.3%	0.7990	0.7948	0.76
CA	11	16	27	64.3%	0.7993	0.7948	0.376
VA	13	6	19	45.2%	0.8225	0.7943	0.168
AK	9	6	15	35.7%	0.8257	0.7942	0.388
WY	6	8	14	33.3%	0.8261	0.7942	0.504
MD	14	5	19	45.2%	0.8306	0.7941	0.052
NE	9	4	13	31.0%	0.8333	0.7940	0.28
IN	16	4	20	47.6%	0.8401	0.7939	0.02
WI	16	4	20	47.6%	0.8469	0.7937	0.002
MO	16	4	20	47.6%	0.8504	0.7936	0.002
NJ	16	4	20	47.6%	0.8533	0.7935	<0.001
SD	11	4	15	35.7%	0.8548	0.7935	0.154
HI	10	1	11	26.2%	0.8571	0.7934	0.234
TN	17	3	20	47.6%	0.8571	0.7934	<0.001
NM	17	3	20	47.6%	0.8580	0.7934	<0.001
OK	18	4	22	52.4%	0.8603	0.7934	<0.001
RI	16	4	20	47.6%	0.8618	0.7933	0.014
AR	18	4	22	52.4%	0.8694	0.7932	<0.001
NH	15	3	18	42.9%	0.8710	0.7931	0.064
IA	18	3	21	50.0%	0.8720	0.7931	<0.001
AL	21	3	24	57.1%	0.8816	0.7929	<0.001
WA	22	3	25	59.5%	0.8840	0.7928	<0.001
NY	32	3	35	83.3%	0.8884	0.7927	<0.001
VT	39	1	40	95.2%	0.9216	0.7919	<0.001
LA	39	0	39	92.9%	0.9235	0.7919	<0.001
IL	41	0	41	97.6%	0.9350	0.7916	<0.001
Less Than	521	521					

2. A concern that state income is driving the differences in IC between states. This was a concern if the states with patients on the margin of poverty are falling in and out of eligibility, it will bias the Informed Coverage metric. We were therefore asked to correlate the poverty burden of the state with Informed Coverage.

We performed 4 analyses to provide the committee with information concerning level of poverty of a state and IC levels. We believe that Figure 2.C. is the most important analysis. It shows no correlation between state income level and informed coverage.

2.A. Figure 2.A presents the same states as in the original Figure 2, with IC on the y-axis and now state ordering based on percentage of the population below the poverty line (most poverty to the left, to least poverty to the right, where state poverty was defined by the American Community Survey (ACS), a Census Bureau statistic, updated annually for inflation. We chose to use the poverty rates reported by the ACS because the Census Bureau recommends using the ACS for one year estimates of poverty at the state level.[1-3]

2.B. Figure 2.B plots the states also in the order of their poverty burden on the x-axis (most poverty to the left, to least poverty to the right) versus actual percent poverty on the vertical axis.

2.C. We next plot the correlation between state poverty rate and Informed Coverage. As can be seen, there are 44 states on the correlation plot. The correlation was only 0.03 and was not significant.

2.D. Table 2 finally provides the same information as we described in Table 1, but now ordered by state poverty percentage.

From these analyses, we conclude that we see no evidence that state poverty level explains our state informed coverage estimates.

Figure 2.A. Graph of IC in States Ranked by State Poverty Level (highest poverty (left) to least poverty(right)).

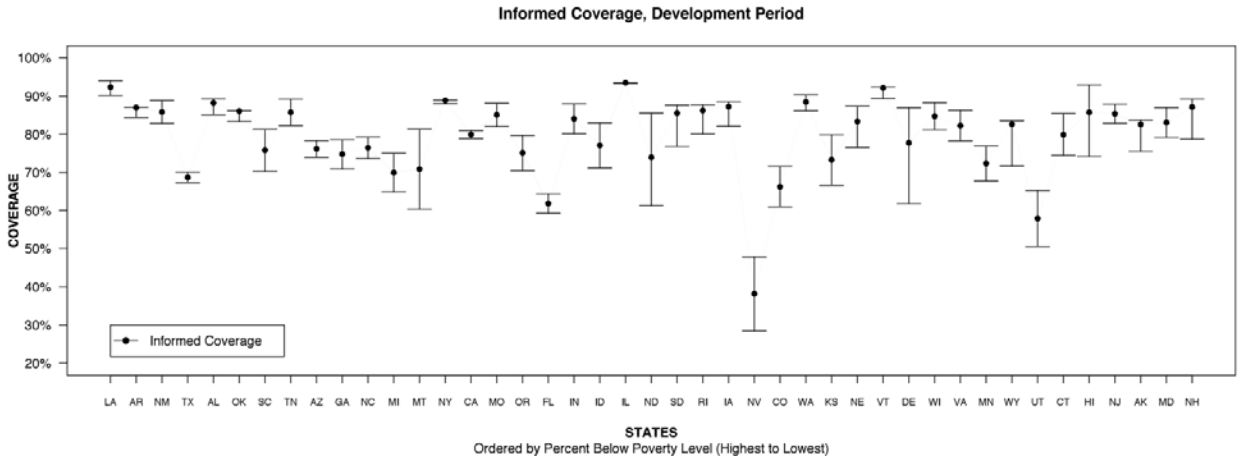


Figure 2.B. Percent Below Poverty Line by ordered by state poverty ((highest poverty (left) to least poverty(right)).

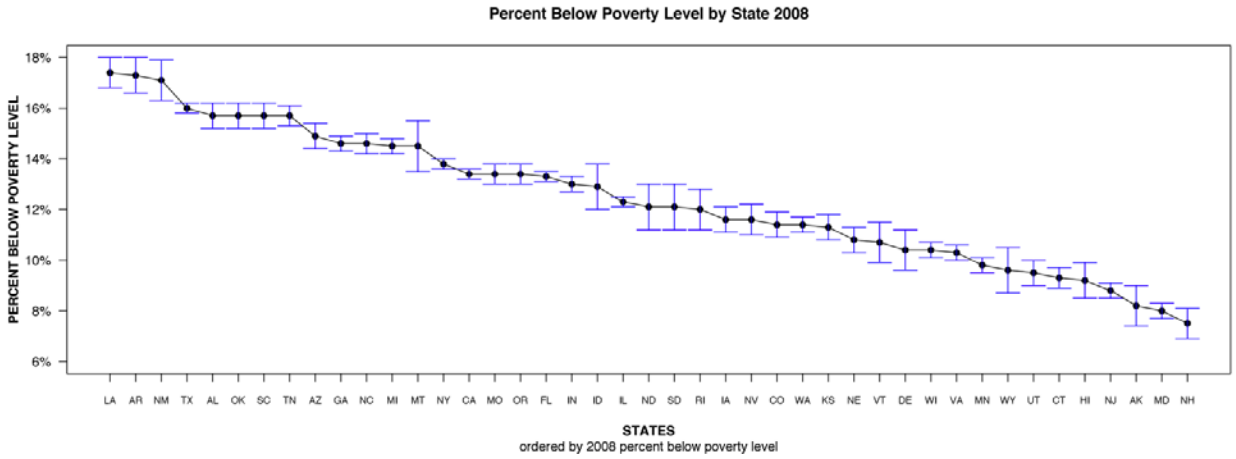


Figure 2.C. Correlation Between IC and Poverty (highest poverty (left) to least poverty(right)).

Scatter plot of State Informed Coverage by Percent Below Poverty Level,
with linear regression line and correlation

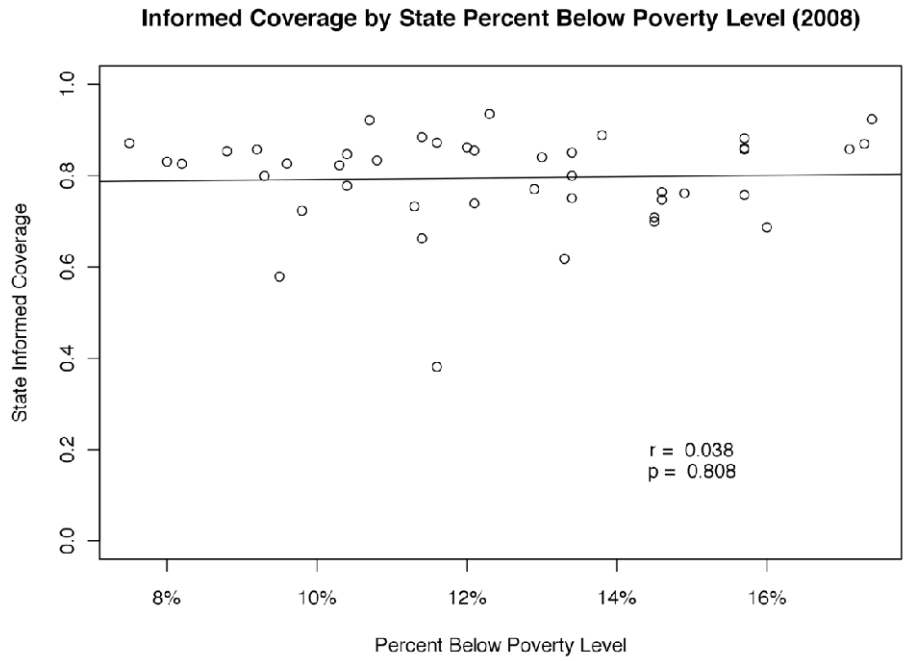
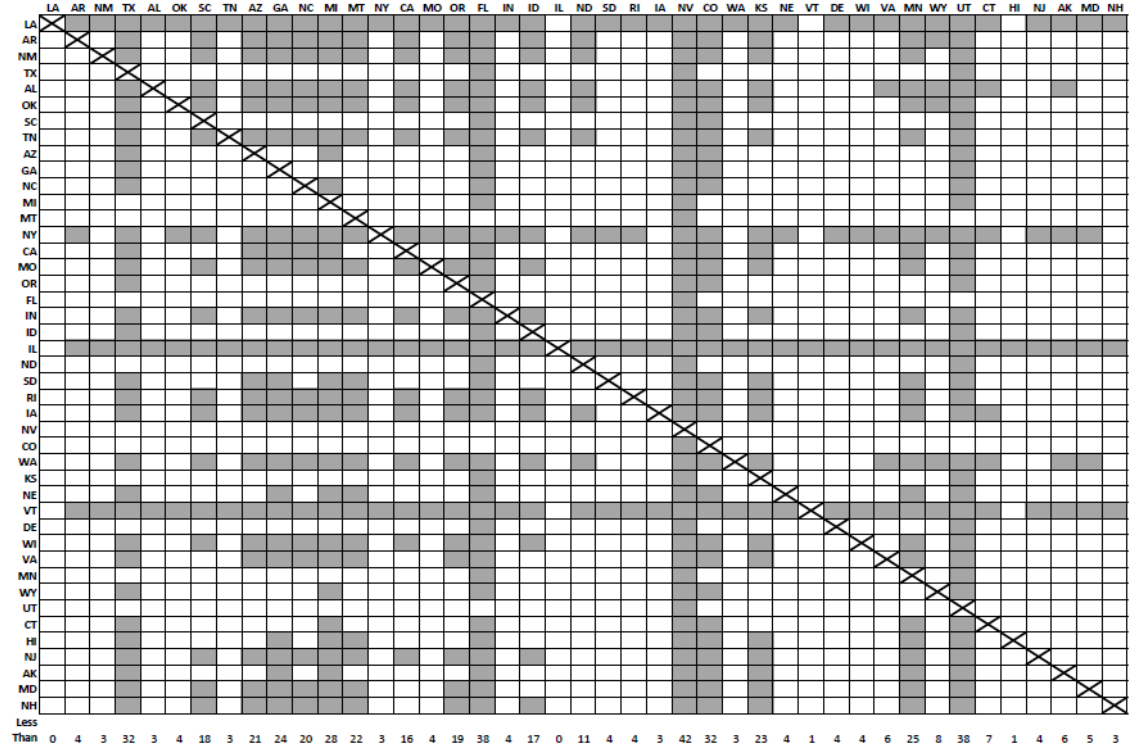


Table 2. State by State Comparison of IC in States Ranked by Statewide Percent Below the Poverty Line (highest poverty (left) to least poverty(right)).

Is the row state significantly larger than the column state?
Ordered By State's Percent Below Poverty Level (Highest to Lowest)



	Greater Than	Less Than	Different Than	% Different Than	this state infcov	mean other states infcov	pvalue
LA	39	0	39	92.9%	0.9235	0.7919	<0.001
AR	18	4	22	52.4%	0.8694	0.7932	<0.001
NM	17	3	20	47.6%	0.8580	0.7934	<0.001
TX	3	32	35	83.3%	0.6868	0.7975	<0.001
AL	21	3	24	57.1%	0.8816	0.7929	<0.001
OK	18	4	22	52.4%	0.8603	0.7934	<0.001
SC	5	18	23	54.8%	0.7576	0.7958	0.236
TN	17	3	20	47.6%	0.8571	0.7934	<0.001
AZ	6	21	27	64.3%	0.7612	0.7957	0.012
GA	5	24	29	69.0%	0.7474	0.7961	0.004
NC	6	20	26	61.9%	0.7641	0.7957	0.082
MI	3	28	31	73.8%	0.6994	0.7972	<0.001
MT	1	22	23	54.8%	0.7083	0.7970	0.092
NY	32	3	35	83.3%	0.8884	0.7927	<0.001
CA	11	16	27	64.3%	0.7993	0.7948	0.376
MO	16	4	20	47.6%	0.8504	0.7936	0.002
OR	5	19	24	57.1%	0.7507	0.7960	0.088
FL	1	38	39	92.9%	0.6180	0.7991	<0.001
IN	16	4	20	47.6%	0.8401	0.7939	0.02
ID	5	17	22	52.4%	0.7704	0.7955	0.358
IL	41	0	41	97.6%	0.9350	0.7916	<0.001
ND	3	11	14	33.3%	0.7391	0.7963	0.422
SD	11	4	15	35.7%	0.8548	0.7935	0.154
RI	16	4	20	47.6%	0.8618	0.7933	0.014
IA	18	3	21	50.0%	0.8720	0.7931	<0.001
NV	0	42	42	100.0%	0.3814	0.8048	<0.001
CO	1	32	33	78.6%	0.6623	0.7981	<0.001
WA	22	3	25	59.5%	0.8840	0.7928	<0.001
KS	3	23	26	61.9%	0.7326	0.7964	0.048
NE	9	4	13	31.0%	0.8333	0.7940	0.28
VT	39	1	40	95.2%	0.9216	0.7919	<0.001
DE	3	4	7	16.7%	0.7778	0.7953	0.962
WI	16	4	20	47.6%	0.8469	0.7937	0.002
VA	13	6	19	45.2%	0.8225	0.7943	0.168
MN	3	25	28	66.7%	0.7231	0.7966	<0.001
WY	6	8	14	33.3%	0.8261	0.7942	0.504
UT	1	38	39	92.9%	0.5789	0.8001	<0.001
CT	7	7	14	33.3%	0.7990	0.7948	0.76
HI	10	1	11	26.2%	0.8571	0.7934	0.234
NJ	16	4	20	47.6%	0.8533	0.7935	<0.001
AK	9	6	15	35.7%	0.8257	0.7942	0.388
MD	14	5	19	45.2%	0.8306	0.7941	0.052
NH	15	3	18	42.9%	0.8710	0.7931	0.064
Less Than	521	521					

Two Sided Bootstrap pvalue
p < 0.05
p >= 0.05

Total Comparisons
1806

Rate of Significant Differences
0.288

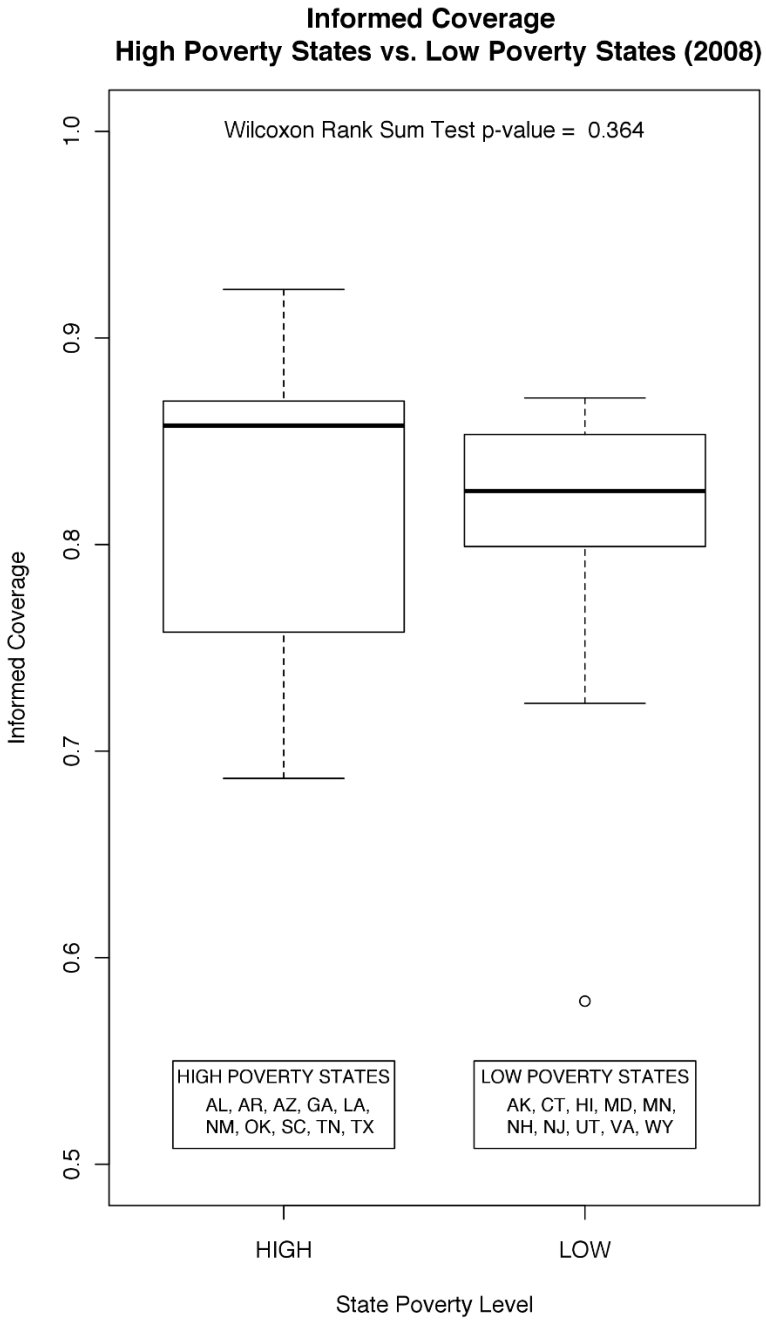
3. A concern that the algorithm for determining if a patient was enrolled four months prior to an appendectomy would lead to biased estimates of IC based on the poverty level of the state. The concern was that the Informed Coverage metric may be biased if there is much movement in and out of the Medicaid system because parents are on the cusp of eligibility. The discussion included a prediction by some committee members that richer states would have lower Informed Coverage because people would be moving in and out of poverty (a mention was made of "churning"). (Note, I responded that because we are using a point-in-time analysis to get the rate of enrollment (4 months before the appendectomy) we should not see the bias that was suggested in the comments that day).

To address this worry, we present 2 analyses:

3.A. We first display the 10 richest and 10 poorest states and ask if they differ in IC. We find no significant difference between these states in terms of IC. The median of the 10 richest states showed an IC of 0.858 versus 0.826 ($P < 0.364$) in the poorest states, and no statistically significant difference.

3.B. We next provide the overall correlation between the state poverty rate and IC (that was already described in Figure 2.C.). Using all the state data, we saw no association between poverty and IC rate.

In other words, previously we saw no correlation between IC and poverty as shown in Figure 2.C., and now we show no association between grouping the ten highest and ten lowest states with respect to poverty. We see no evidence to support the concern that poverty is biasing the state IC rate estimate.


Informed Coverage
High Poverty States vs. Low Poverty States (2008)
Wilcoxon Rank Sum Test p-value = 0.364
Informed Coverage
1.0
0.9
0.8
0.7
0.6
0.5
HIGH POVERTY STATES
AL, AR, AZ, GA, LA,
NM, OK, SC, TN, TX
LOW POVERTY STATES
AK, CT, HI, MD, MN,
NH, NJ, UT, VA, WY
HIGH
LOW
State Poverty Level

4. In a related concern, the committee wanted to see how different our results would be had we used 5 month rather than 4 month look back, and wanted to see if the poverty of the state would influence this difference.

To examine this question we first looked at the relationship between 4-month and 5-month look-backs by examining the correlation in IC for the 44 state estimates comparing the 4-month versus the 5-month look-back. We found the following correlations: Pearson = 0.992(95% CI = 0.986, 0.996), Spearman = 0.986, and a Kendall's Tau yielding a probability of concordance of 0.956. The probability of concordance says that with a probability of 0.956, a state that was ranked higher (or lower) than another state using the 4 month look back would also be ranked higher (or lower) using the IC estimate based on a 5 month look-back. In short, we find no evidence that we would get systematically different relationships between the states based on using a 5 versus 4 month look-back to estimate IC.

We next examined whether poverty level of the state was associated with the difference in IC estimates between the 4-month and 5-month look backs. That is, did the high or low poverty of a state bias the 4 versus 5-month relationship. To do this we examined the correlation between the difference between the 4 versus 5-month look back based estimates of IC versus state poverty rate. We found the Pearson correlation between state poverty level and the difference between 4 and 5-month IC was -0.0596 (95% CI = -0.354, 0.245), P = 0.704 (not significant), and the Spearman correlation coefficient was -.0240 (P = 0.8785), and the probability of concordance was 0.490, suggesting almost a random chance of ordering the state differences in 4 versus 5-month differences by poverty level.

In summary, there was a concern of the committee that state poverty level may bias the IC estimate because of "churning" in higher income states. We find no evidence to support this concern.

References

1. United States Census Bureau. Income and poverty: Which data source to use. 2016. Available at: https://www.census.gov/topics/income-poverty/guidance/data-sources.html. Accessed March 27, 2017.
2. Bishaw A, Macartney S. Poverty: 2008 and 2009. American Community Survey Briefs. Issued September 2010. Available at: https://www.census.gov/prod/2010pubs/acsbr09-1.pdf. Accessed March 27, 2017.
3. Bishaw A. Poverty: 2010 and 2011. American Community Survey Briefs. Issued September 2012. Available at: https://www.census.gov/prod/2012pubs/acsbr11-01.pdf. Accessed March 27, 2017.

Appendix B: Additional Analysis and Information #3189 and #2816

TO: National Quality Forum's Pediatric Committee

FROM: Dr. Lawrence Kleinman

Date: May 12, 2017

Re: Additional Analysis and Information

Dear Pediatric Committee,

It was a pleasure to present two measures to you on March 2, 2017. To summarize, we submitted the following two measures to the Pediatric Project: (1) Measure 3189: Rate of Emergency Department Visit Use for Children Managed for Identifiable Asthma - Visits per 100 Child years and (2) Measure 2816: Appropriateness of Emergency Department Visits for Children and Adolescents with Identifiable Asthma.

At the in person meeting, measure 3189 passed on evidence and gap, and was voted insufficient for reliability, and for measure 2816 consensus was not reached for evidence, the measure passed on gap, and did not pass on reliability. In general, the sense of the group at the in-person meeting was that measure 3189 is a very viable measure, but having to conform to the NQF procedure, the group required a little bit more data, which is provided herein:

1) Reliability
2) Inclusion/Exclusion
3) Pharmacy Data
4) Race Disparities
5) Data Element Validity

Please note that the analyses and additional information are applicable for both measures - #3189 and #2816.

Thank you for your time and consideration.

Kind regards,
Larry

Lawrence C. Kleinman, MD, MPH, FAAP
Director, Center for Child Health and Policy
Frederick C. Robbins MD Professor in Child and Adolescent Health
Vice Chair for Child Health and Policy

Department of Pediatrics
UH Rainbow Babies' and Children's Hospital

Case Western Reserve University School of Medicine

Center: 216 844-0067
Cell: 617 699 3357

1) RELIABILITY

It is worth noting that measure #3189 is a claims-based measure and measure #2816 includes claims, EHRs and paper records. The whole notion of reliability, inter-rater reliability, test/retest reliability, does not really hold in the same way for a claims-based measure as with other measures put forth, as it requires a demonstration of difference of performance. With guidance from NQF and the committee, we conducted several statistical summaries of reliability for this distribution – Poisson, zero-inflated Poisson, and the hurdle. Using the single data set we had available, we looked at various plans and we looked at counties and found that it worked in both. All of which are theoretically coherent, all came up with very similar results, and all of which were all coherent. Confidence intervals were sufficiently narrow to demonstrate differences between various reference entities (county, plan) and many other of the entities. We provide both tables and charts for the following three analysis that further support reliability:

a. By county – using NYC as the comparator county
b. By county – using Westchester as the comparator county
c. By Plans

a. By County - Using NYC as the comparator county

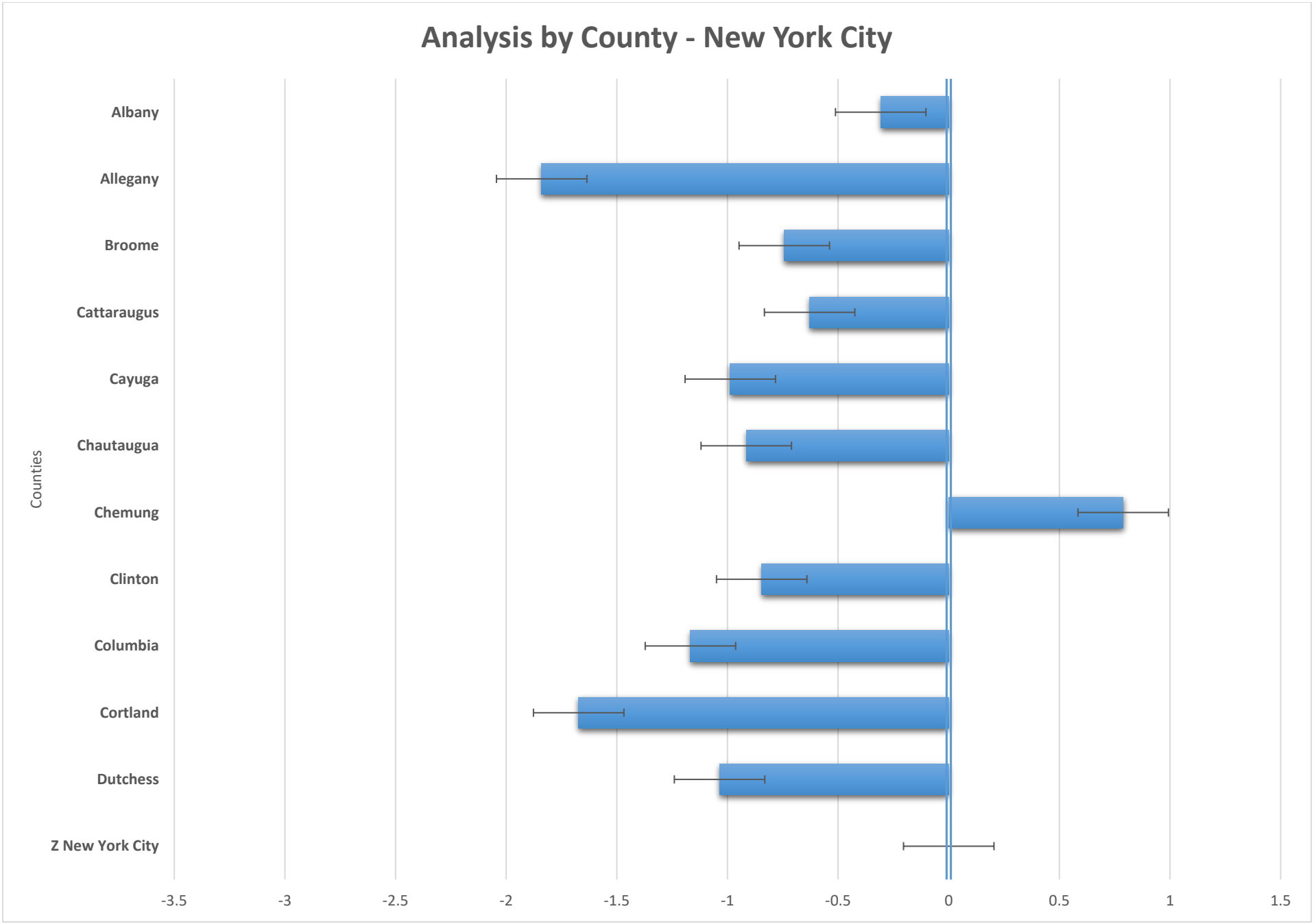
To demonstrate reliability, we conducted the following analysis of Maximum Likelihood that provides the rates, 95% CIs, and significance levels that indicates difference of performance by county. As you can see, all but 1 county (Chemung) has significant differences at least $p<0.05$ level, with most having a significance at $p<0.0001$ level. This demonstrates the difference of performance by county (Table 1) using an atypical region (the five counties of NYC) as a comparator. Figure 1, shows this analysis in a chart with the respective standard errors. As you can see, the standard error bars do not overlap and again shows significant differences between counties. These analyses control for age group.

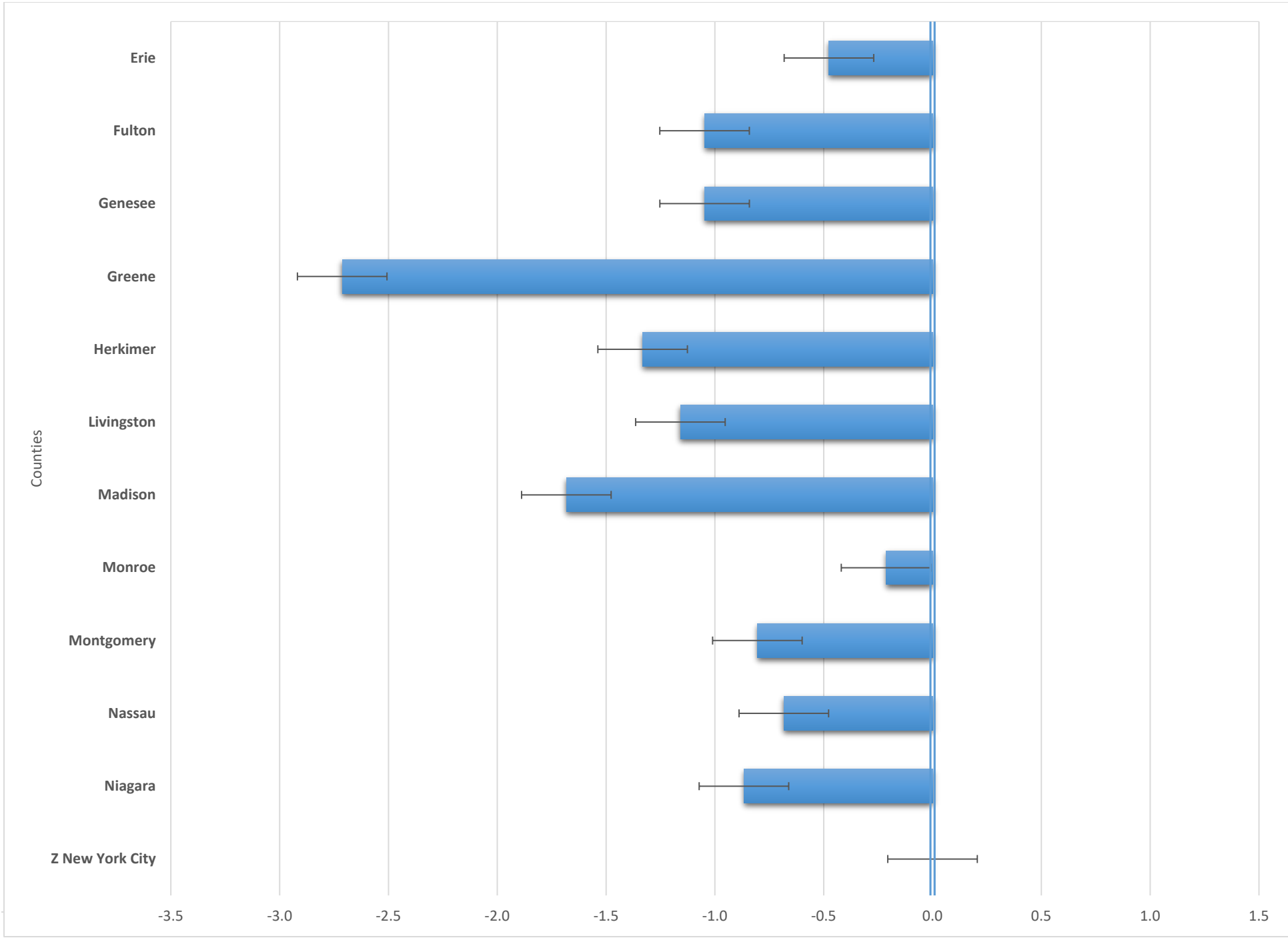
Table 1: Analysis by county, using NYC as the comparator

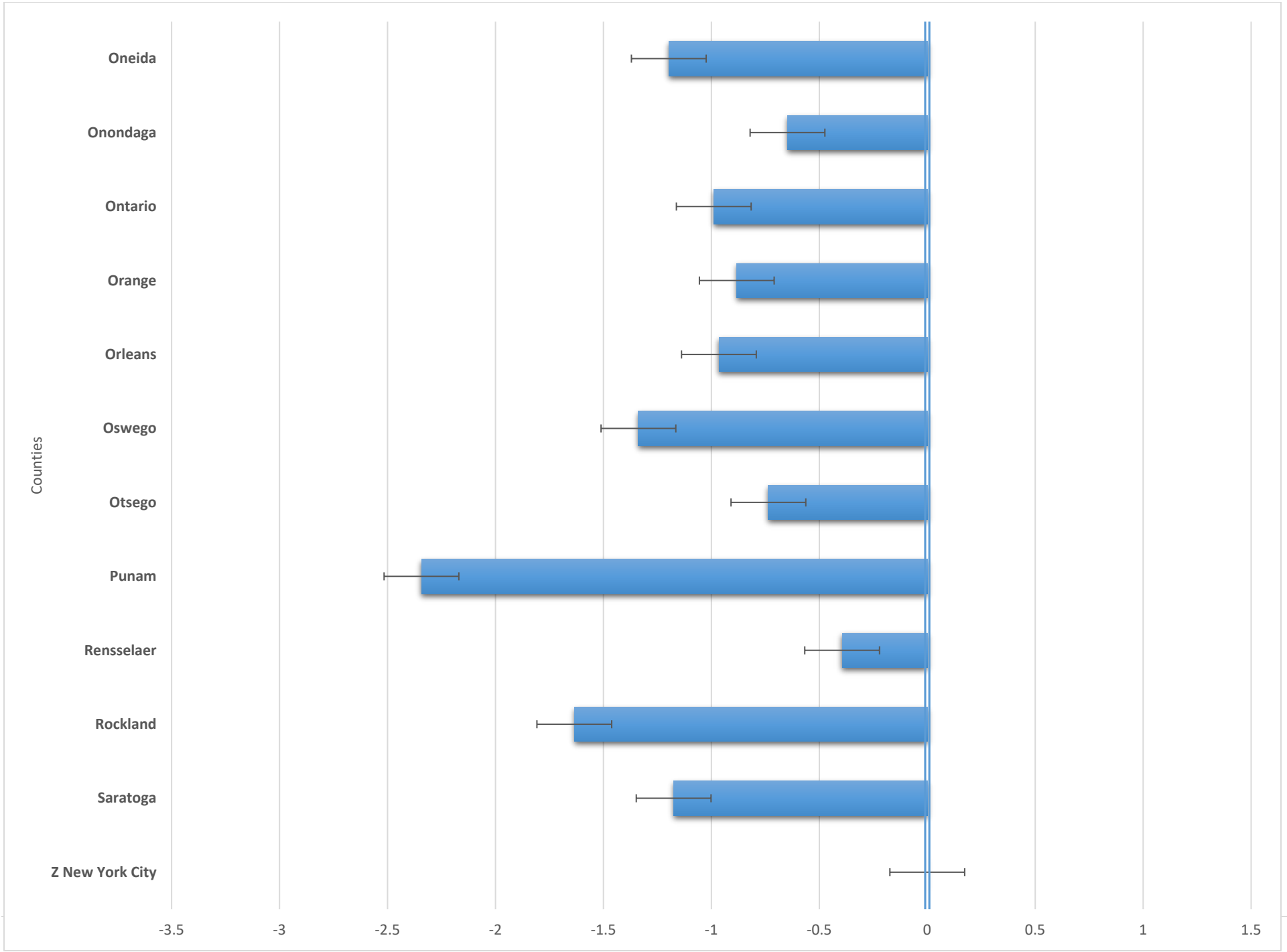
Analysis of Maximum Likelihood Parameter Estimates								
Parameter		DF	Estimate	Standard Error	Wald 95% Confidence Limits		Walk Chi-Square	Pr>ChiSq
Intercept			-3.9019	0.0237	-3.9485	-3.8554	27020.2	<.0001
mbr_county_desc	Albany	1	-0.3075	0.0632	-0.4314	-0.1837	23.68	<.0001
mbr_county_desc	Allegany	1	-1.8396	0.3536	-2.5327	-1.1465	27.06	<.0001
mbr_county_desc	Broome	1	-0.7433	0.0956	-0.9306	-0.5560	60.50	<.0001
mbr_county_desc	Cattaraugus	1	-0.6288	0.1338	-0.8910	-0.3665	22.09	<.0001
mbr_county_desc	Cayuga	1	-0.9872	0.2358	-1.4494	-0.5251	17.53	<.0001
mbr_county_desc	Chautaugua	1	-0.9151	0.0956	-1.1024	-0.7278	91.71	<.0001
mbr_county_desc	Chemung	1	0.7884	0.7075	-0.5983	2.1751	1.24	0.2651
mbr_county_desc	Clinton	1	-0.8452	0.2673	-1.3692	-0.3212	9.99	0.0016
mbr_county_desc	Columbia	1	-1.1674	0.1858	-1.5316	-0.8032	39.47	<.0001
mbr_county_desc	Cortland	1	-1.6724	0.2426	-2.1479	-1.1968	47.51	<.0001
mbr_county_desc	Dutchess	1	-1.0357	0.1326	-1.2956	-0.7758	61.00	<.0001
mbr_county_desc	Erie	1	-0.4760	0.0394	-0.5533	-0.3987	145.60	<.0001
mbr_county_desc	Fulton	1	-1.0475	0.1492	-1.3399	-0.7550	49.28	<.0001
mbr_county_desc	Genesee	1	-1.0473	0.2237	-1.4858	-0.6089	21.92	<.0001
mbr_county_desc	Greene	1	-2.7126	0.4473	-3.5892	-1.8360	36.78	<.0001

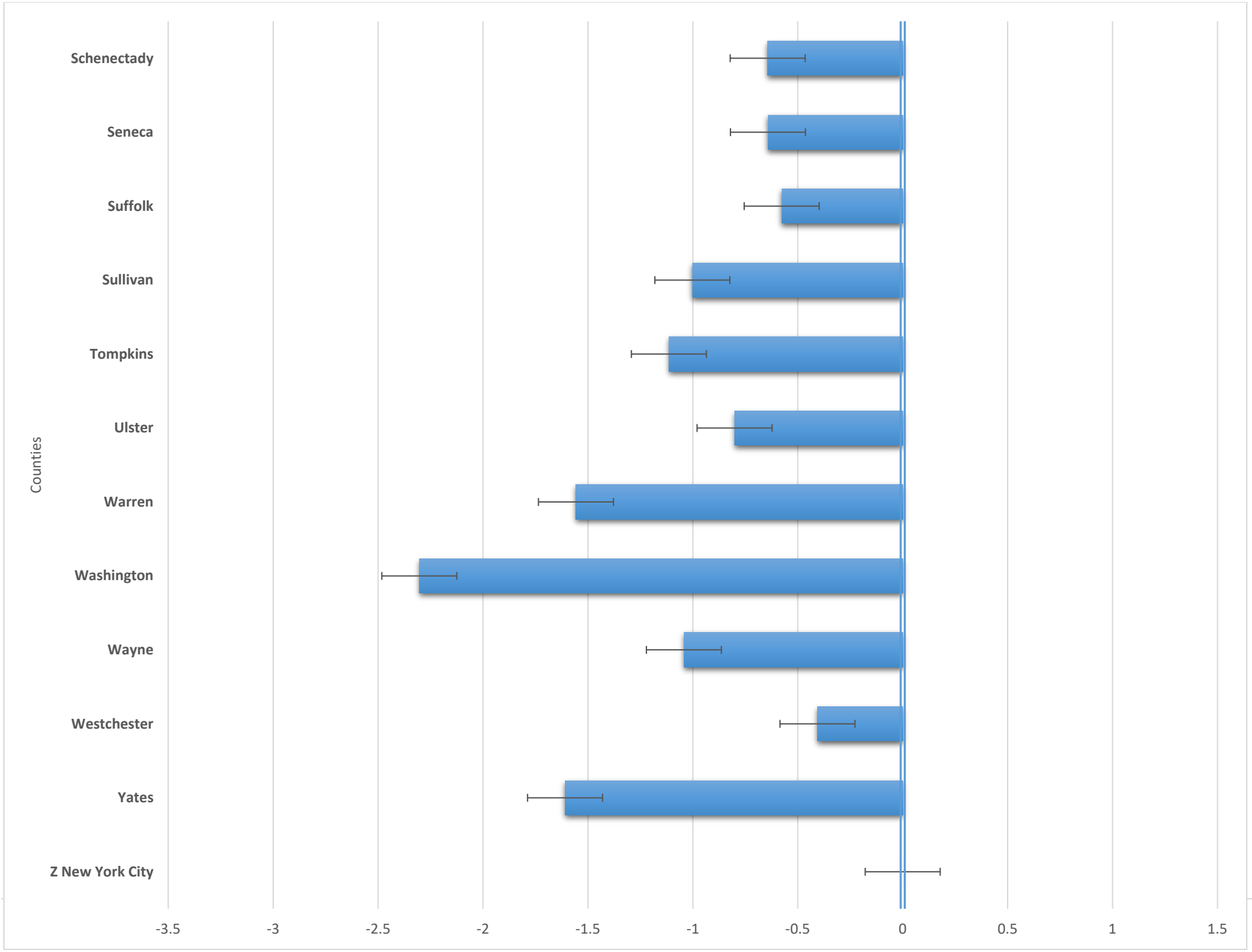
mbr_county_desc	Herkimer	1	-1.3321	0.2133	-1.7501	-0.9140	39.00	<.0001
mbr_county_desc	Livingston	1	-1.1585	0.2426	-1.6340	-0.6830	22.80	<.0001
mbr_county_desc	Madison	1	-1.6826	0.3163	-2.3025	-1.0627	28.30	<.0001
mbr_county_desc	Monroe	1	-0.2136	0.0361	-0.2845	-0.1428	34.93	<.0001
mbr_county_desc	Montgomery	1	-0.8048	0.1214	-1.0429	-0.5668	43.92	<.0001
mbr_county_desc	Nassau	1	-0.6833	0.0456	-0.7727	-0.5939	224.50	<.0001
mbr_county_desc	Niagara	1	-0.8666	0.0904	-1.0438	-0.6895	91.92	<.0001
mbr_county_desc	Oneida	1	-1.1972	0.0927	-1.3789	-1.0156	166.90	<.0001
mbr_county_desc	Onondaga	1	-0.6474	0.0581	-0.7612	-0.5335	124.23	<.0001
mbr_county_desc	Ontario	1	-0.9889	0.1827	-1.3469	-0.6308	29.30	<.0001
mbr_county_desc	Orange	1	-0.8823	0.0698	-1.0191	-0.7455	159.79	<.0001
mbr_county_desc	Orleans	1	-0.9651	0.2774	-1.5088	-0.4213	12.10	0.0005
mbr_county_desc	Oswego	1	-1.3378	0.1445	-1.6210	-1.0546	85.74	<.0001
mbr_county_desc	Otsego	1	-0.7358	0.1827	-1.0938	-0.3777	16.22	<.0001
mbr_county_desc	Punam	1	-2.3427	0.4473	-3.2193	-1.4660	27.43	<.0001
mbr_county_desc	Rensselaer	1	-0.3943	0.0830	-0.5570	-0.2316	22.57	<.0001
mbr_county_desc	Rockland	1	-1.6348	0.1050	-1.8407	-1.4289	242.20	<.0001
mbr_county_desc	Saratoga	1	-1.1747	0.1544	-1.4774	-0.8720	57.86	<.0001
mbr_county_desc	Schenectady	1	-0.6435	0.0935	-0.8266	-0.4603	47.39	<.0001
mbr_county_desc	Seneca	1	-0.6422	0.2673	-1.1662	-0.1182	5.77	0.0163
mbr_county_desc	Suffolk	1	-0.5768	0.0369	-0.6491	-0.5045	244.43	<.0001
mbr_county_desc	Sullivan	1	-1.0025	0.1716	-1.3389	-0.6662	34.12	<.0001
mbr_county_desc	Tompkins	1	-1.1143	0.2237	-1.5527	-0.6758	24.81	<.0001
mbr_county_desc	Ulster	1	-0.8012	0.1223	-1.0409	-0.5614	42.89	<.0001
mbr_county_desc	Warren	1	-1.5572	0.4473	-2.4338	-0.6806	12.12	0.0005
mbr_county_desc	Washington	1	-2.3035	0.3780	-3.0444	-1.5626	37.13	<.0001
mbr_county_desc	Wayne	1	-1.0427	0.1858	-1.4068	-0.6785	31.49	<.0001
mbr_county_desc	Westchester	1	-0.4062	0.0415	-0.4876	-0.3248	95.67	<.0001
mbr_county_desc	Yates	1	-1.6091	0.4473	-2.4857	-0.7325	12.94	0.0003
mbr_county_desc	Z New York City	0	0.0000	0.0000	0.0000	0.0000		
agegrp	1	1	0.3905	0.0254	0.3407	0.4402	236.53	<.0001
agegrp	2	1	-0.1096	0.0253	-0.1592	-0.0599	18.71	<.0001
agegrp	3	1	-0.3269	0.0269	-0.3797	-0.2742	147.74	<.0001
agegrp	4	0	0.0000	0.0000	0.0000	0.0000		

Figure 1: Analysis by county, using NYC as the comparator – including standard errors


Analysis by County - New York City
Albany
Allegany
Broome
Cattaraugus
Cayuga
Chautaugua
Chemung
Clinton
Columbia
Cortland
Dutchess
Z New York City
Counties
-3.5
-3
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5


Counties
Erie
Fulton
Genesee
Greene
Herkimer
Livingston
Madison
Monroe
Montgomery
Nassau
Niagara
Z New York City
-3.5
-3.0
-2.5
-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5


Counties
Oneida
Onondaga
Ontario
Orange
Orleans
Oswego
Otsego
Punam
Rensselaer
Rockland
Saratoga
Z New York City
-3.5
-3
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5


Schenectady
Seneca
Suffolk
Sullivan
Tompkins
Ulster
Warren
Washington
Wayne
Westchester
Yates
Z New York City
Counties
-3.5
-3
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5

b. Using Westchester as the comparator county

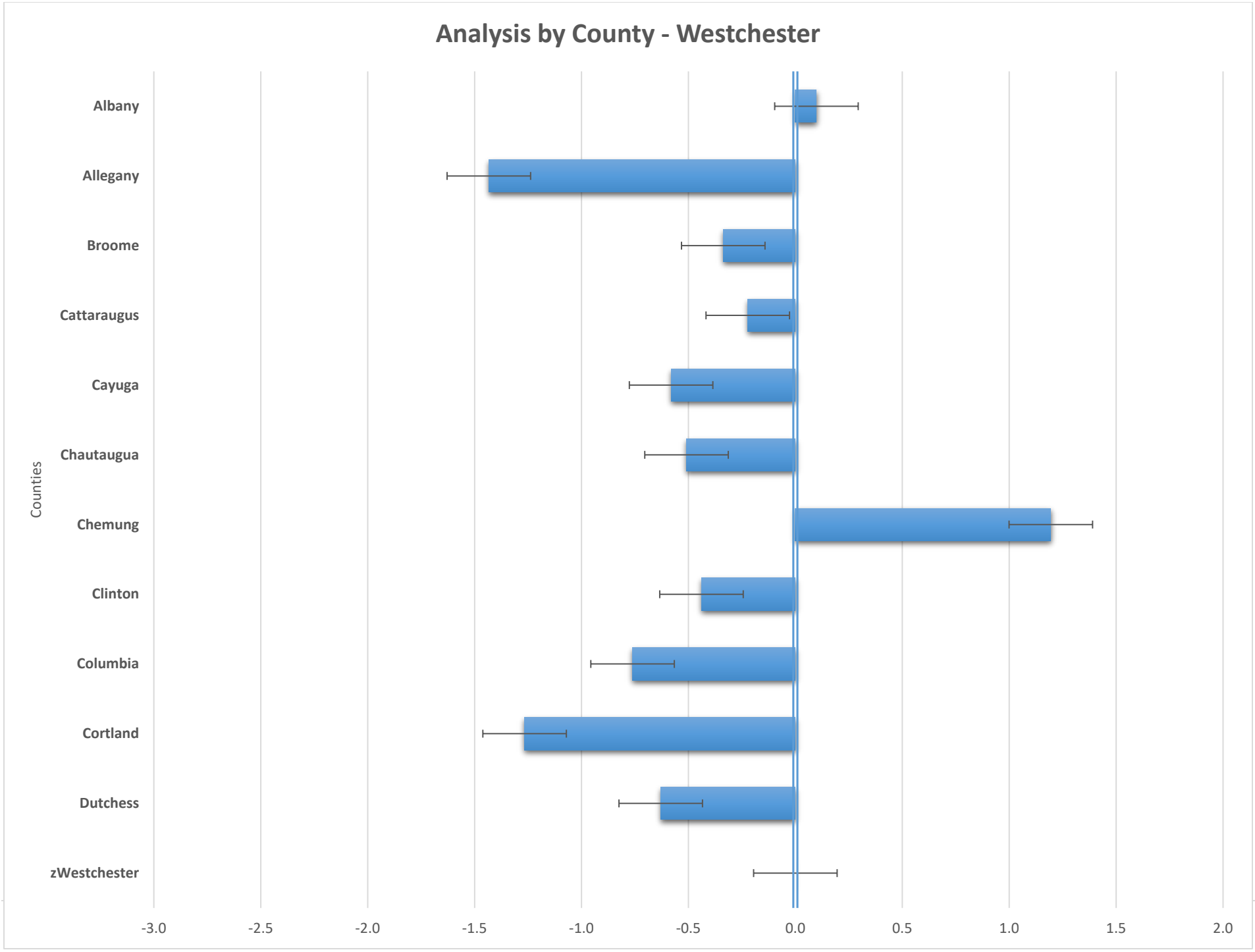
Recognizing the exceptionalism of New York City (NYC), we reanalyzed the data using a normal suburban county (Westchester) as the comparator county to show that meaningful differences remain. Following the findings from above, Table 2 and Figure 2 show the rates, 95% CIs, and significance levels that indicates difference of performance by county. Not as drastic as above, but there are still all but 8 counties that have significant differences at least at the p<0.05 level, with most having a significance at p<0.0001 level. This again demonstrates the difference of performance by county (Table 2). Figure 2, like Figure 1, shows this analysis in a chart with the respective standard errors. As you can see, the standard error bars again do not overlap and again shows significant differences between counties.

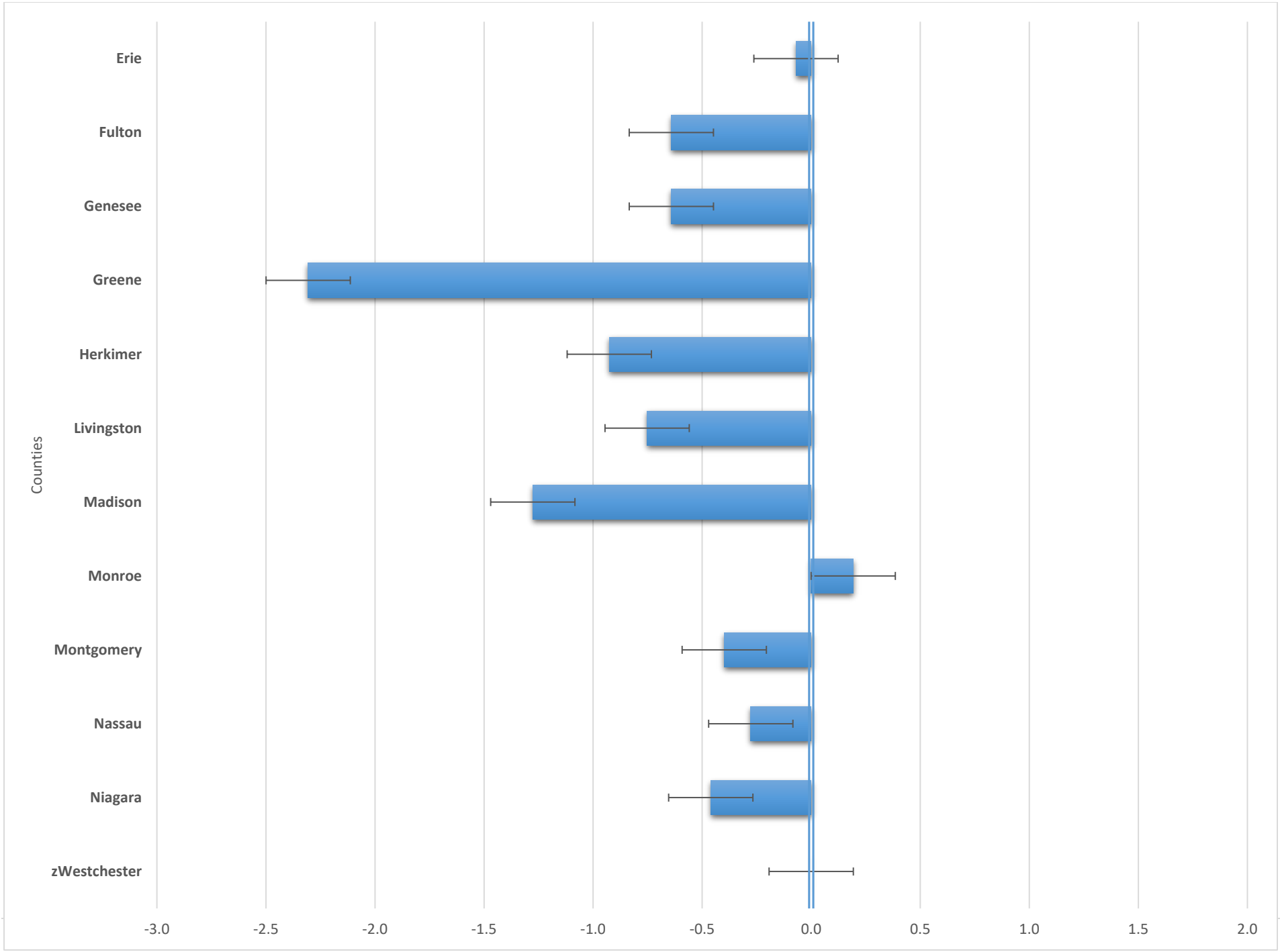
Table 2: Analysis by county, using a normal suburban county (Westchester) as the comparator

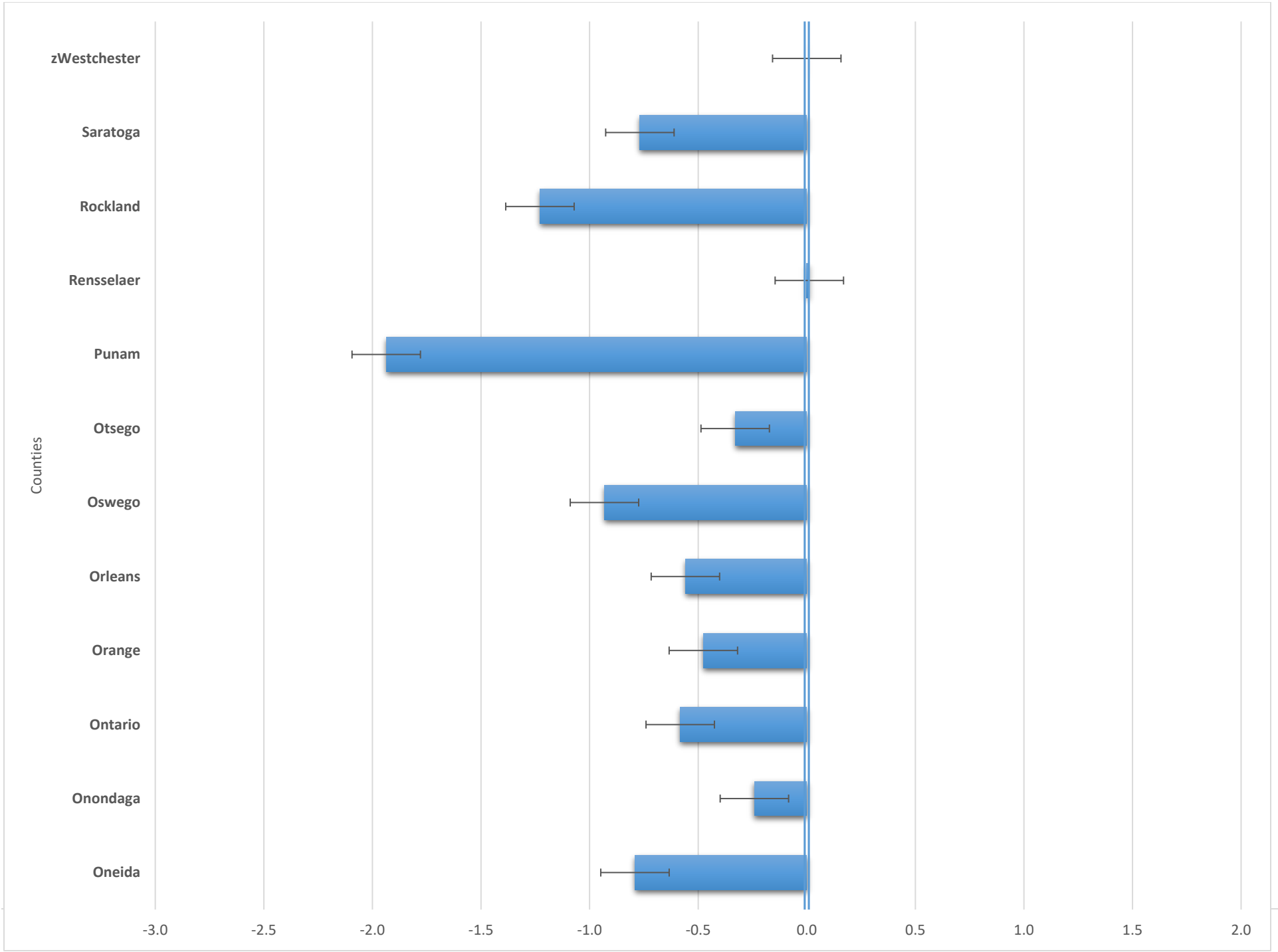
Analysis of Maximum Likelihood Parameter Estimates								
Parameter		DF	Estimate	Standard Error	Wald 95% Confidence Limits		Walk Chi-Square	Pr>ChiSq
Intercept			-4.3081	0.0470	-4.4003	-4.2159	8387.34	<.0001
mbr_county_desc	Albany	1	0.0986	0.0751	-0.0485	0.2458	1.73	0.1889
mbr_county_desc	Allegany	1	-1.4334	0.3559	-2.1311	-0.7358	16.22	<.0001
mbr_county_desc	Broome	1	-0.3372	0.1038	-0.5406	-0.1337	10.55	0.0012
mbr_county_desc	Cattaraugus	1	-0.2226	0.1398	-0.4966	0.0514	2.54	0.1113
mbr_county_desc	Cayuga	1	-0.5811	0.2392	-1.0500	-0.1121	5.90	0.0152
mbr_county_desc	Chautaugua	1	-0.5090	0.1038	-0.7124	-0.3055	24.04	<.0001
mbr_county_desc	Chemung	1	1.1946	0.7087	-0.1944	2.5835	2.84	0.0919
mbr_county_desc	Clinton	1	-0.4390	0.2704	-0.9690	0.0910	2.64	0.1045
mbr_county_desc	Columbia	1	-0.7613	0.1902	-1.1340	-0.3885	16.02	<.0001
mbr_county_desc	Cortland	1	-1.2662	0.2460	-1.7483	-0.7841	26.50	<.0001
mbr_county_desc	Dutchess	1	-0.6296	0.1387	-0.9013	-0.3578	20.61	<.0001
mbr_county_desc	Erie	1	-0.0698	0.0566	-0.1807	0.0411	1.52	0.2172
mbr_county_desc	Fulton	1	-0.6413	0.1546	-0.9444	-0.3383	17.20	<.0001
mbr_county_desc	Genesee	1	-0.6412	0.2273	-1.0868	-0.1956	7.95	0.0048
mbr_county_desc	Greene	1	-2.3064	0.4491	-3.1866	-1.4262	26.38	<.0001
mbr_county_desc	Herkimer	1	-0.9259	0.2171	-1.3515	-0.5004	18.19	<.0001
mbr_county_desc	Livingston	1	-0.7523	0.2460	-1.2345	-0.2702	9.35	0.0022
mbr_county_desc	Madison	1	-1.2765	0.3189	-1.9014	-0.6515	16.02	<.0001
mbr_county_desc	Monroe	1	0.1925	0.0543	0.0861	0.2990	12.56	0.0004
mbr_county_desc	Montgomery	1	-0.3987	0.1280	-0.6496	-0.1477	9.70	0.0018
mbr_county_desc	Nassau	1	-0.2771	0.0610	-0.3967	-0.1576	20.64	<.0001
mbr_county_desc	Niagara	1	-0.4605	0.0991	-0.6546	-0.2663	21.60	<.0001
mbr_county_desc	Oneida	1	-0.7911	0.1011	-0.9893	-0.5928	61.17	<.0001
mbr_county_desc	Onondaga	1	-0.2412	0.0708	-0.3801	-0.1024	11.60	0.0007
mbr_county_desc	Ontario	1	-0.5827	0.1871	-0.9495	-0.2159	9.70	0.0018
mbr_county_desc	Orange	1	-0.4762	0.0807	-0.6344	-0.3180	34.80	<.0001
mbr_county_desc	Orleans	1	-0.5589	0.2804	-1.1084	-0.0094	3.97	0.0462
mbr_county_desc	Oswego	1	-0.9316	0.1501	-1.2257	-0.6375	38.54	<.0001
mbr_county_desc	Otsego	1	-0.3296	0.1871	-0.6964	0.0372	3.10	0.0782

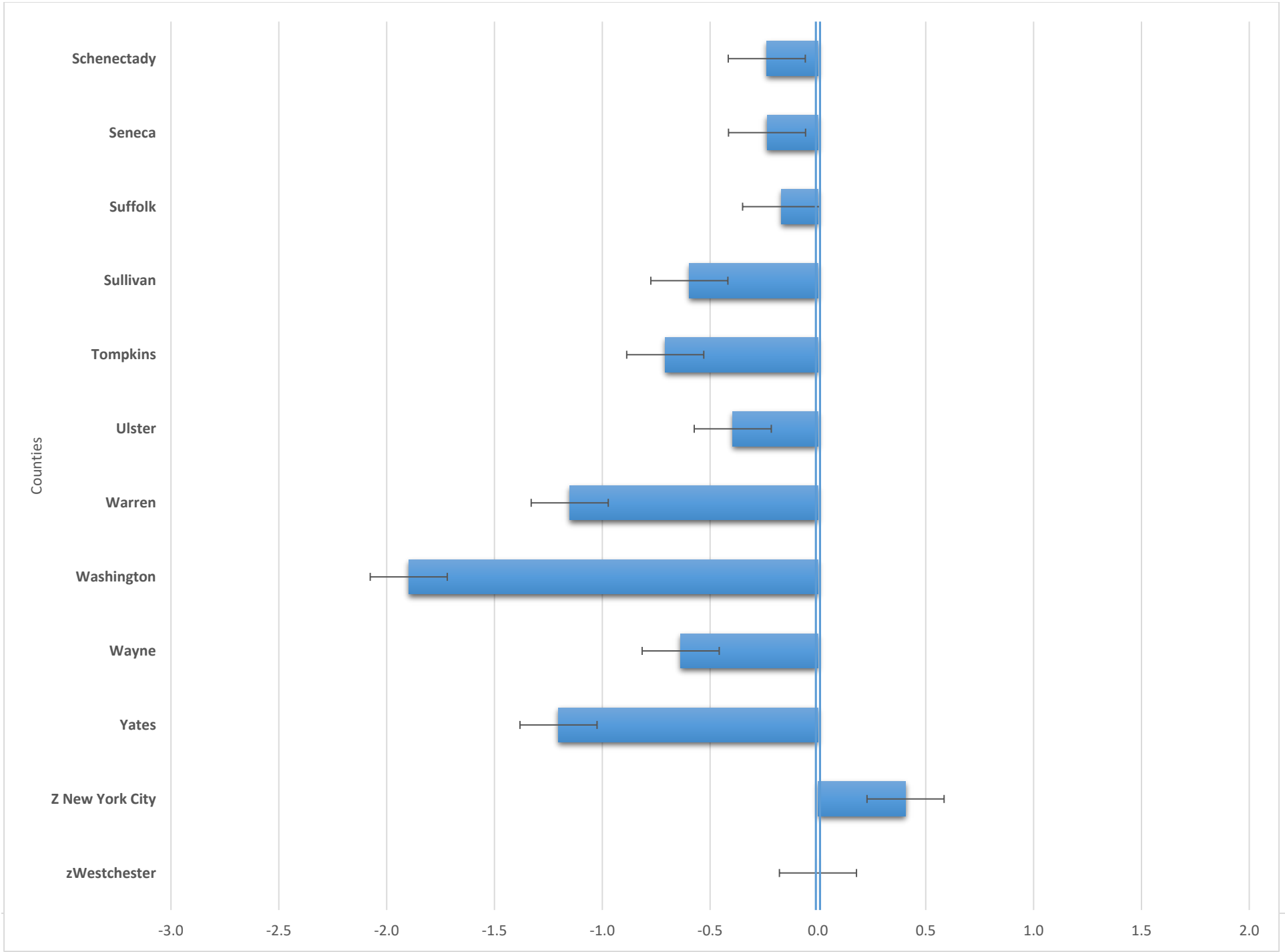
mbr_county_desc	Punam	1	-1.9365	0.4491	-2.8167	-1.0563	18.59	<.0001
mbr_county_desc	Rensselaer	1	0.0118	0.0924	-0.1692	0.1929	0.02	0.8980
mbr_county_desc	Rockland	1	-1.2287	0.1126	-1.4493	-1.0080	119.10	<.0001
mbr_county_desc	Saratoga	1	-0.7685	0.1597	-1.0815	-0.4556	23.17	<.0001
mbr_county_desc	Schenectady	1	-0.2373	0.1019	-0.4370	-0.0376	5.42	0.0199
mbr_county_desc	Seneca	1	-0.2361	0.2704	-0.7660	0.2939	0.76	0.3827
mbr_county_desc	Suffolk	1	-0.1707	0.0548	-0.2781	-0.0633	9.70	0.0018
mbr_county_desc	Sullivan	1	-0.5963	0.1763	-0.9420	-0.2507	11.44	0.0007
mbr_county_desc	Tompkins	1	-0.7081	0.2273	-1.1537	-0.2625	9.70	0.0018
mbr_county_desc	Ulster	1	-0.3950	0.1289	-0.6476	-0.1424	9.39	0.0022
mbr_county_desc	Warren	1	-1.1510	0.4491	-2.0312	-0.2708	6.57	0.0104
mbr_county_desc	Washington	1	-1.8974	0.3802	-2.6425	-1.1522	24.91	<.0001
mbr_county_desc	Wayne	1	-0.6365	0.1902	-1.0092	-0.2638	11.20	0.0008
mbr_county_desc	Yates	1	-1.2030	0.4491	-2.0832	-0.3228	7.18	0.0074
mbr_county_desc	Z New York City	1	0.4062	0.0415	0.3248	0.4876	95.67	<.0001
mbr_county_desc	zWestchester	0	0.0000	0.0000	0.0000	0.0000		
agegrp	1	1	0.3905	0.0254	0.3407	0.4402	236.53	<.0001
agegrp	2	1	-0.1096	0.0253	-0.1592	-0.0599	18.71	<.0001
agegrp	3	1	-0.3269	0.0269	-0.3797	-0.2742	147.74	<.0001
agegrp	4	0	0.0000	0.0000	0.0000	0.0000		
Scale		0	1.0000	0.0000	1.0000	1.0000		

Figure 2: Analysis by county, using a normal suburban county (Westchester) as the comparator – including standard errors


Analysis by County - Westchester
Albany
Allegany
Broome
Cattaraugus
Cayuga
Chautaugua
Chemung
Clinton
Columbia
Cortland
Dutchess
zWestchester
Counties
-3.0
-2.5
-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0


Counties
Erie
Fulton
Genesee
Greene
Herkimer
Livingston
Madison
Monroe
Montgomery
Nassau
Niagara
zWestchester
-3.0
-2.5
-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0


zWestchester
Saratoga
Rockland
Rensselaer
Punam
Otsego
Oswego
Orleans
Orange
Ontario
Onondaga
Oneida
Counties
-3.0
-2.5
-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0


Schenectady
Seneca
Suffolk
Sullivan
Tompkins
Ulster
Warren
Washington
Wayne
Yates
Z New York City
zWestchester
Counties
-3.0
-2.5
-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0

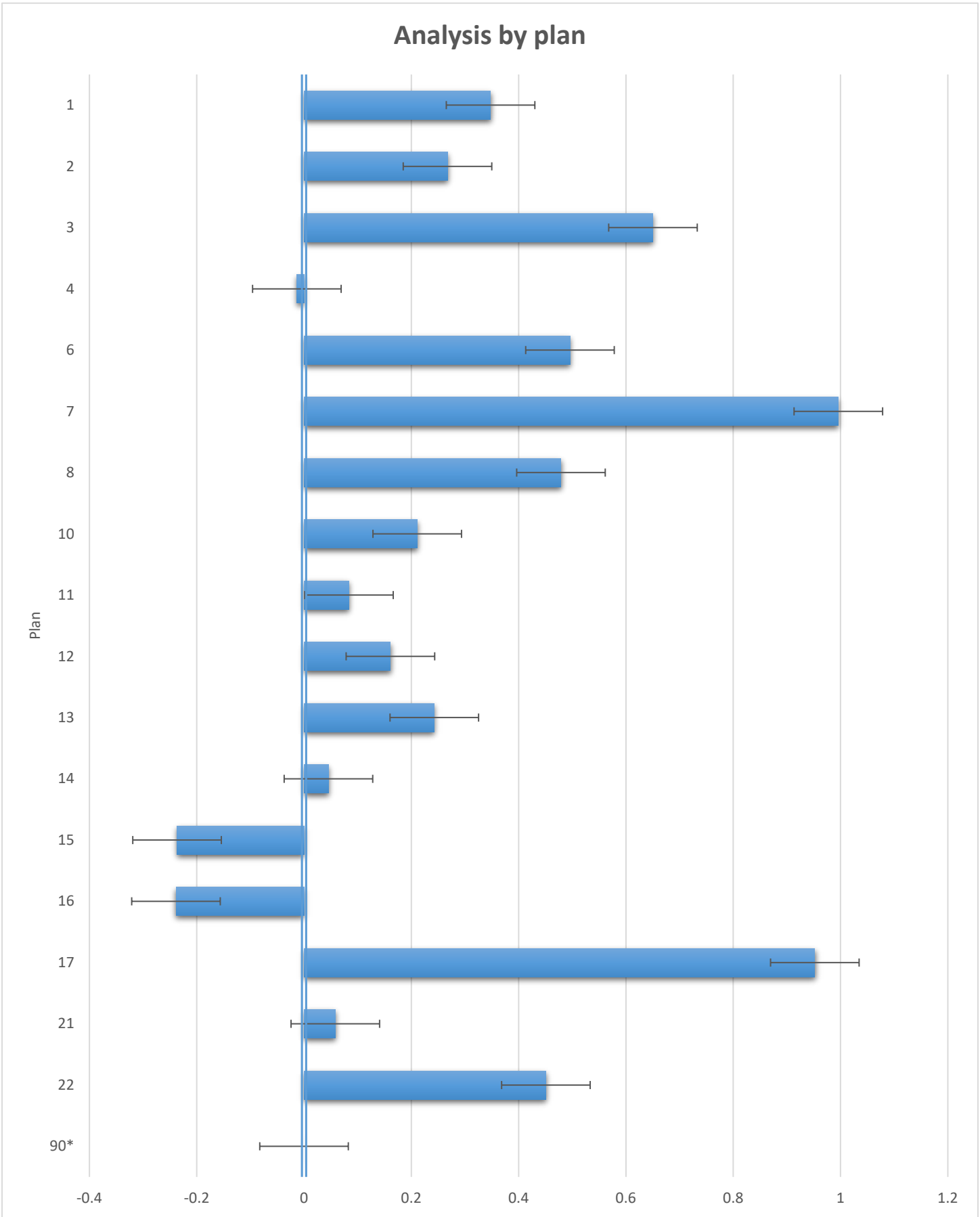
c. By Plan: Poisson Plan

When we analyze the data by plan in the Poisson model, we ran it with a number of different plans as the index plan, and had typical results, which show that some plans are similar and some plans are different, which is what you would want to see. In the analysis that we present, we found that 50% of the plans (9 of the 18 plans) have a p-value of less than 0.001 (Table 3). This precision analysis shows that the standard errors are typically small relative to the estimates, which indicates that there are tight confidence intervals, but that not every plan is different from every other plan (Figure 3). For example, if a coefficient of 0.34, with a standard error of 0.05, that tells us the 95 confidence interval is between 0.25 and 0.45. We have this information for each plan and for every county, which shows that the confidence intervals are narrow enough that plans differ and are outside of one another's confidence interval. Although this model does not produce an s-statistic, which means we are unable to calculate reliability in the standard way, but this model is adjusted for age-group as we are recommending it be done.

Table 3: Analysis by plan

Analysis of Maximum Likelihood Parameter Estimates								
Parameter		DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square	Pr > Chi Sq
Intercept		1	-4.6254	0.0509	-4.7250	-4.5257	8273.70	<.0001
plan	1	1	0.3477	0.0506	0.2485	0.4468	47.21	<.0001
plan	2	1	0.2674	0.0565	0.1567	0.3780	22.43	<.0001
plan	3	1	0.6504	0.0495	0.5534	0.7474	172.71	<.0001
plan	4	1	-0.0134	0.0838	-0.1776	0.1508	0.03	0.8733
plan	6	1	0.4957	0.0490	0.3997	0.5917	102.48	<.0001
plan	7	1	0.9959	0.0466	0.9046	1.0872	456.75	<.0001
plan	8	1	0.4789	0.0841	0.3141	0.6437	32.44	<.0001
plan	10	1	0.2110	0.0689	0.0760	0.3460	9.39	0.0022
plan	11	1	0.0837	0.0653	-0.0443	0.2117	1.64	0.2001
plan	12	1	0.1610	0.1173	-0.0689	0.3909	1.88	0.1698
plan	13	1	0.2428	0.0752	0.0953	0.3902	10.41	0.0013
plan	14	1	0.0455	0.0831	-0.1173	0.2082	0.30	0.5842
plan	15	1	-0.2367	0.0557	-0.3458	-0.1276	18.09	<.0001
plan	16	1	-0.2387	0.0715	-0.3790	-0.0985	11.14	0.0008
plan	17	1	0.9521	0.0464	0.8612	1.0429	421.66	<.0001
plan	21	1	0.0583	0.0906	-0.1192	0.2358	0.41	0.5197
plan	22	1	0.4508	0.0478	0.3571	0.5446	88.90	<.0001
plan	90	0	0.0000	0.0000	0.0000	0.0000		
agegrp	1	1	0.3378	0.0268	0.2852	0.3903	158.65	<.0001
agegrp	2	1	-0.1688	0.0268	-0.2213	-0.1162	39.65	<.0001
agegrp	3	1	-0.3785	0.0285	-0.4343	-0.3226	176.39	<.0001
agegrp	4	0	0.0000	0.0000	0.0000	0.0000		
Scale		0	1.0000	0.0000	1.0000	1.0000		

Figure 3: Analysis by plan – with standard errors



We provide additional supporting information to address questions raised by the committee in the in-person meeting. The analyses and additional information are applicable for both measures - #3189 and #2816.

2) INCLUSION/EXCLUSION

From the committee: The inclusion of visits with "bronchitis" is unusual as a criterion for identifiable asthma. I do not see data on how including or excluding these diagnoses would change the denominator in a given health plan or population. Rationale for why short-acting beta agonists are excluded in the denominator of "identifiable asthma" - Provide evidence for how changes in specifications (not using short-acting beta agonists) might affect validity.

Exclusions: Of 1.4 million children in this portion of the analysis (of whom approximately 125,000 had identifiable asthma), the following number of children had at least one exclusionary diagnosis (some had more than one) based on the two year look back period: COPD 2121, acute respiratory failure 2139, cystic fibrosis 650 and emphysema 482.

Leukotriene inhibitors: Leukotriene inhibitors may be used as a controller medication for asthma and also for children with allergy without asthma. In one analysis of the NYS Medicaid children, 795 of 70,898 children (1.1%) with an asthma diagnosis used leukotriene inhibitors exclusively among the potentially asthma related drugs. We do not know how many of these 795 otherwise qualified for the denominator because of other criteria. Overall ~125,000 children met denominator criteria. Since we don't know how many of the children who use leukotrienes alone were included based upon other criteria (or for how many of the children the medication was provided because of allergy and not asthma), we are left to calculate a maximum potential error: Assuming that none qualified for other criteria (highly unlikely) and that none were prescribed the leukotriene inhibitors because of their allergy and not their asthma (also unlikely), the maximum underestimation of the denominator = 795/(795+125,252)=or 0.6%. These numbers reassure us that the decision to exclude leukotriene inhibitors as a qualifying medication is both clinically reasonable and will not unduly influence the measure.

Our panel chose to include controller medications only and not rescue medications when setting their criteria for inclusion. Short acting beta agonists are rescue medications that are used more broadly than for asthma, for example in the context of acute wheezing. This exclusion is particularly important in light of the panel's decision to allow some visits for bronchitis to be considered for inclusion. Although, we do not have specific data regarding the impact of these decisions, we can state that we achieved the face validity desired, being more sensitive than the HEDIS persistent asthma measure but not indiscriminant allowing in the denominator all children who have ever had an asthma diagnosis the decision both to include. The measure reflects the expert panel judgment. As stated in our submission, our construct for the CAPQuaM measure was defined by the multidisciplinary national expert panel using a RAND type modified Delphi process, which produced a set of explicit criteria that were both substantive and addressed specification details, such as what combination of administrative codes could be used to identify a child with asthma, other inclusion and exclusion criteria for the measure, and preferences regarding how to report and stratify the measure. The panel initially used the term persistent asthma to describe asthma that was pre-existing and should have been recognized as asthma by the health care system prior to the timing of the ED visit. This construct was renamed by our stakeholder group to be identifiable asthma to avoid confusion with other uses of the term persistent asthma. The construct was intended to be more inclusive than HEDIS' persistent asthma diagnosis, while still removing from consideration those whose asthma was unlikely to have been actively managed at the time.

Holding steady the continuous enrollment criterion at the HEDIS-specified 12 months, HEDIS criteria identified a rate of persistent asthma of 3.1% with the CAPQuaM criteria identifying identifiable asthma at a rate of 8.6%. As expected, identifiable asthma was between 2 and 3 times more permissive than the intentionally restrictive

persistent asthma. We analyzed NSCH data to estimate a population rate of asthma in NY State Medicaid child population to be between 15 - 16%, indicating that our criteria did provide a meaningful filter as we had intended.

3) PHARMACY DATA

From the committee: Pharmacy data availability and its effect on the measure

In one analysis of New York State data, of 125,036 children identified with asthma, 70,898 of them met the medication criteria and of those, 1,037 were identified specifically because of the medication use criteria that was identified through pharmacy data. Thus the denominator would have been reduced by (1037./125,036) 0.008%.

4) RACE DISPARITIES

From the committee: It demonstrates disparities, however, it is not clear how complete New York Medicaid data are with regards to race. What percent of Medicaid enrollee race is incomplete?

We cannot locate the specific analysis from this validation study that confirmed that race was typically coded in New York State Medicaid. In another analysis of 321,586 children 0-20 in NY State Medicaid that our team conducted, race was not a missing variable.

Race			
White	115570	36%	321586
Hispanic	80077	25%	321586
Black	55453	17%	321586
Unknown	43963	14%	321586
Other	13487	4%	321586
Asian	12252	4%	321586
Native American	784	0%	321586

5) DATA ELEMENT VALIDITY

From the committee: The analysis conducted needs to show that claims can accurately identify children with asthma, children with the various exclusions, and ED/hospitalizations. The latter may not even require a complex statistical analysis. Given what has provided for the other measures, this one should be close to meeting all criteria terms of data element validity. Please note that NQF doesn't accept the score-level testing of other measures as demonstration of validity for this measure, so we suggest removing narrative about that part. The citations provided for 2816 may be applicable here.

See below for data element validity.

Data element	**Reference** (e.g., Quam, et al., 1993)	**Data source** (e.g., Medicare FFS outpatient data)	**Statistical results** (e.g., kappa, sensitivity, specificity, etc.)
Numerator: Validation of Numerator Data Elements was performed by the CAPQuaM development team and the results are summarized in this section following the table.			
Denominator			
Age	NYSDOH CAPQuaM Analysis – internal testing	NY State Medicaid Data	Meaningful variation by age groups as predicted, with peaks in younger children and older adolescents.
	CMS MMIS data requirements Exemplar specifications at https://www.cms.gov/Research-Statistics-Data-and-Systems/Computer-Data-and-Systems/MSIS/downloads/msisdd2010.pdf	State Medicaid MMIS systems	States are required to submit validated claims data including age or date of birth with a tolerance of 0.1%
Asthma diagnosis in inpatient/ED setting	Wilchesky, M., Tamblyna, R. M., & Huang, A. (2004). Validation of diagnostic codes within medical services claims. Journal of Clinical Epidemiology, 57, 131-141.	Drug utilization review, the Charlson comorbidity index and the Johns Hopkins Adjusted Care Group Case-Mix profile (ADGs).	Asthma claims were highly specific, Sp= 96.76 (95%CI 96.5, 97.0).
Asthma diagnosis in ambulatory setting	Fowles, J. B., Fowler, E. J., & Craft, C. (1998). Validation of claims diagnoses and self-reported conditions compared with medical records for selected chronic diseases. Journal of Ambulatory Care Management, 21(1), 24-34.	Multispecialty group practice in Minneapolis, Minnesota	Sensitivity and specificity was 0.82 and 0.99, respectively. Sensitivity of .82 using claims was higher than sensitivity using self-report at .64
Asthma diagnosis in clinic/outpatient setting	Wilchesky, M., Tamblyna, R. M., & Huang, A. (2004). Validation of diagnostic codes within medical services claims. Journal	Drug utilization review, the Charlson comorbidity index and the Johns Hopkins Adjusted Care Group Case-Mix profile (ADGs).	Asthma claims were highly specific, Sp= 96.76 (95%CI 96.5, 97.0).

	of Clinical Epidemiology, 57, 131-141.		
Bronchitis diagnosis in ambulatory setting	Improving Healthcare for the Common Good (IPRO). Avoidance of Antibiotic Treatment in Adults with Acute Bronchitis. May 2011. http://www.health.ny.gov/health_care/managed_care/reports/docs/adults_antibiotic.pdf	New York Medicaid managed care members	An IPRO analysis of ambulatory claims data in NY State Medicaid found that of 651 individuals with an administrative claim for bronchitis, 629 (96.6%) were confirmed by chart review.
Fill of short acting beta agonist Fill of asthma controller medication • anti- asthmatic combination • antibody inhibitor • inhaled steroid combinations • inhaled corticosteroids (alone or in combination) • leukotriene modifiers • methylxanthines (alone or in combination) • mast cell stabilizers	Samnaliev, M., Baxter, J. D., & Clark, R. E. (2009). Comparative evaluation of two asthma care quality measures among Medicaid beneficiaries. Chest, 135(5), 1193-1196. Mudd KE, Bollinger ME, Hus VD, et al. Concordance of Mediaciad and pharmacy record data in inner-city children with asthma. Contemporary Clinical Trials 29(2008) 13-20 Grymonpre R, Xheang M, Fraser M, et al. cvalidity of Precritpion Claims Database to Estimate Medication Adherence in Older Persons	Using complete claims and pharmaceutical data for 19,076 patients with persistent asthma (based on Health Effectiveness and Data Information Set criteria) in five Medicaid populations (Colorado, Georgia, Indiana, New Jersey, Washington) using ICD-9-CM code 493.x to measure filling prescriptions of asthma control medication. Comparison of pharmacy records and Medicaid clams Manitoba prescription claims and pill count for medication adherence A number of studies found that asthma drug data using the similar	Sensitivity and specificity were combined into one statistic, the area under the ROC curve. For controller medications, the area under ROC curve is 0.705, which represents good agreement. For inner city children on Medicaid, Medicaid claims was sensitive compared to pharmacy records, identifying 91.3% of pharmacy claims for ICS, 94.7% for SABA and 90.4% for leukotriene modifiers (Table 2) Using a much stronger standard of actual compliance, this study found for multiple condition for two conditions in adults that there was strong concordance (79% and 88% respectively) between pill counts and administrative claims data. Not specific for asthma meds Controller medication use was associated with fewer ED visits across 5 states, with OR ranging from 0.30 to 0.47, all significant, overall 0.34 (0.32-0.36). Used actual HEDIS pharmacy code set as do we.

	e.g. Samnaliev M, Baxter JD, and Clark RE. Comparative Evaluation of Two Asthma Care Quality Measure Among Medicaid Beneficiaries. Berger WE, Legorreta AP, Blaiss MS, et al. The Utility of the HEDIS Asthma Measure to predict asthma related outcomes. Annals of Allergy, Asthma, and Immunology. 93:538-545. 2004.	HEDIS data elements that we propose were valid for predicting things like emergency department use in asthma patients. As indicated in this article: "HEDIS has become an important industry standard…adopted by regulators, consumers, and public purchasers of health care" Commercial claims	Low Controller use had an adjusted odds ratio of 1.72 (1.42-2.08) of ED visit or hospitalization. Those with moderate and higher adherence had graded reductions in undesirable outcomes in the predicted fashion (OR, .84 and 0.72 respectively)
Exclusions			
Diagnosis of COPD	Rawson NS, Malcolm E., validity of the recording of ischaemic heart disease and chronic obstructive pulmonary disease in the Saskatchewan health care datafiles. State Med. 1995. Dec 30: 14 (24):2627-43.	Administrative health care datafiles of the Canadian province of Saskatchewan	Comparisons between hospital data and medical charts for chronic airways obstruction patients showed excellent diagnostic agreement at 94%. In other words, the charted discharge diagnosis from the patient's medical record showed exact agreement for 94.2% of these patients.
	Ginde AA, Tsai CL, Blanc PG, Camargo CA Jr. Positive predictive value of ICD-9-CM codes to detect acute exacerbation of COPD in the emergency department. Jt Comm J Qual Patient Saf.2008;34(11):678–680.	Two academic emergency departments.	The overall positive predictive value for the presence of any of the specified codes, including COPD, was 97%. The positive predictive value for a code of 496 alone was 60% (95% CI 32-84%).
	Gershon AS, Wang C, Guan J, Vasilevska-Ristovska J, Cicutto L, To T. Identifying individuals with physician diagnosed COPD in health administrative	Claims in Ontario, Canada	The combination of one or more outpatient ICD-9 codes (491.xx, 492.xx, 496.xx) and ICD-10 inpatient ICD-10 codes (J41, J43, J44) had a sensitivity of 85% and specificity of 78.4% among

	databases. Copd. 2009;6(5):388 –394. doi: 10.1080/15412550903140865.		113 patients with COPD and 329 patients without COPD.
Diagnosis of COPD Diagnosis of cystic fibrosis Diagnosis of emphysema (Exclusions identified anywhere are excluded. The measure is written to over exclude if need be, but our data suggest that exclusions are uncommon.)	Quan, H., Li, B., Saunders, L. D., Parsons, G. A., Nilsson, C. I., Alibhai, A., et al. (2008). Assessing validity of icd-9-cm and icd-10 administrative data in recording clinical conditions in a unique dually coded database. HSR: Health Services Research, 43(4), 1424.	Four teaching hospitals in Alberta, Canada	Claims had a PPV of 91.9, and a negative predictive value of 92.6, with *k* of 0.65 (substantial agreement[i]) compared to chart review for chronic pulmonary disease. ICD 10 performed similarly in this study
	NCQA: http://www.qualityforum.org/QPS/QPSTool.aspx?m=367&e=1	The presence of diagnostic exclusions was extensively tested on the entire field test population (>82,000 members) to determine the effect on eligible population and the measure results experienced as a result of the application of clinical exclusions.	This measure was deemed valid by the expert panel and approved by NCQA's Committee on Performance Measurement (CPM) for continued inclusion in HEDIS[ii]

Data element	**Reference** (e.g., Quam, et al., 1993)	**Data source** (e.g., Medicare FFS outpatient data)	**Statistical results** (e.g., kappa, sensitivity, specificity, etc.)
Race/ Ethnicity	Kressin, NR, Chang, BH, Hendricks, A, Kazis, LE. Agreement Between Administrative Data and Patients' Self-Reports of Race/Ethnicity. American Journal of Public Health. Oct. 2003. 93 (10): 1734-1739.	Federal administrative data	Among patients with known race/ethnicity, there was a 97.9%, 92.0%, and 83.4% agreement between self-report race/ethnicity and administrative data for white, African American, and Hispanic, respectively. (Table 2, p. 1736)
	Blustein, J. The Reliability of Racial Classifications in Hospital Discharge Abstract Data. American Journal Public Health. 1994; 84:1018-1021.	Statewide Planning and Research Cooperative System, a hospital discharge abstract database maintained by the New York State Department of Health.	Percentage of concordance and kappa of reported racial classifications: - Black: 99%; 089 (95% CI: 0.82, 0.96) - White: 95%; 0.72 (95% CI: 0.64, 0.80) (Table 3, page 1020)

	Klinger, EV, Carlini, SV, Gonzalez, I, et al., Accuracy of Race, Ethnicity, and Language Preference in an Electronic Health Record. 2014. J Gen Intern Med. 30(6):719-23.	Thirteen primary care clinics' electronic health records.	When comparing electronic health record to self-report the sensitivity, specificity and ppv for Black, Hispanic and white are as follows (Table 2, page 721): - Black: Se: 70.9, Sp: 98.8, PPV: 95.5 - Hispanic: Se: 83.8, Sp: 99.8; PPV: 98.9 - White: Se: 93.8; Sp: 97.0; PPV: 98.3
	Escarce, JJ and McGuire, TG., Methods for Using Medicare Data to Compare Procedure Rates among Asians, Blacks, Hispanics, Native Americans, and Whites. Health Services Research. Oct. 2003. 38(5): 1303-1318.	Physician claims data	When comparing enrollment database and survey, probability for White, Black, and Hispanic are 0.954, 0.943, 0.977, respectively. (Table 2, page 1309)

[i] The *k* value indicates a near perfect agreement (*k*: 0.81-1.0 between coded data and chart review data), and substantial agreement (*k*: 0.61-0.80).

[ii] We note that 1799 and 1800 are not directly applicable because they were tested at the score level. However, the scores were dependent upon definitions which use the same data element level as our measure and thus provide indirect evidence of the capacity of a measure using such data elements to produce valid scores.

Thus we cite them not as specific evidence of our score level performance of the submitted measure, but as evidence that the HEDIS measures that rely on the same administrative data elements for their denominator have the capacity to distinguish signal to noise at a very high level. While the evidence is indirect it is dispositive. That is, we assert that had the data elements been inadequate it would result in non-differential misclassification error which is a major bias towards the null thus introducing noise and reducing signal. That this does not happen to an appreciable degree specifically implies that the data elements function well – indeed this could be one rationale for why NQF allows the use of performance score level analysis in the first place. These findings provide strong indirect evidence of the validity of our approach to capturing the measure's denominator.

There is nearly complete overlap of the denominator codes and there is overlap of the denominator elements. Where codes differ it is specific to decisions made by the CAPQuaM expert panel which was aware of the NCQA measures. Review of the medication lists for 0036 reveal that all medication used by the submitted CAPQuaM measure are also in the HEDIS measure. The CAPQuaM measure excludes specifically short acting beta agonists and leukotriene inhibitors at the specific direction of the CAPQuaM expert panel. We also specify exclude indacaterol from the list of "asthma specific medications" since it is a long acting beta agonist which is only indicated in the USA for treatment of COPD, which is a specific exclusion criterion for this measure.

Further, we identify asthma visits and medications using the same data that an insurance company or Medicaid would use for payment, including ICD codes, CPT codes, and revenue codes. We have had conversations with expert coders and New York State Department of Health Office of Health Insurance Programs to confirm our choices. Our literature review found that while there is moderate agreement (kappa = 0.45 - 0.50) when comparing administrative data regarding the presence of constructs such as recent asthma attacks, use of asthma medications, attack or medication, attack and medication, using 1 year of administrative claims data to parent report, the agreement improves from 0.55 to 0.60 when using two years of data.(1) We expect that these kappas would be significantly higher were the analyses restricted to children with disease that met our construct criteria for identifiable asthma.

The literature further supports our work as highlighted above in the table and in more detail in our testing form 2b2.3 (validity testing).

NATIONAL QUALITY FORUM

Appendix C

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 #: 3189
Measure Title: Rate of Emergency Department Visit Use for Children Managed for Identifiable Asthma: Visits per 100 Child-years
Measure Steward: University Hospitals Cleveland Medical Center
Brief Description of Measure: This measure estimates the rate of emergency department visits for children ages 2 – 21 who are being managed for identifiable asthma, using specified definitions. The measure is reported in visits per 100 child-years.
Developer Rationale: Asthma is a critical problem with racial and ethnic disparities and varies by urbanicity. Adherence to the National Asthma Education and Prevention Programs (NAEPP) Guidelines improves outcomes. [1-32]. We have elsewhere provided other articles, studies, and summaries of evidence to document that ED visits and hospitalizations are typically outcome measures of choice when assessing asthma control.

ED visits for asthma in children are common and expensive. They may result from poor quality of care delivered (failure to adhere to guidelines) as well as from insufficient access to primary care. Asthma is the leading diagnosis leading to urgent care/emergent care provided in emergency departments for children. It is among the most common chronic diseases in children and expenses for asthma care are in the billions of dollars annually. Further, CMS and AHRQ assigned us this measure. In addition to data and citations provided, the team has analyzed 2007 and 2011 waves of the National Survey of Children's Health and confirmed that this parent reported measure both identified a high prevalence of asthma nationwide and significant consequences in terms of parent reported child health for children who have asthma.

Our analysis of National Survey of Children's Health [33] data (NSCH, 2011/12), estimates that 10.3 million children in the U.S. have been told that they have asthma. Of these children 7.6 million live in more urban areas that are characterized as metropolitan statistical areas (MSAs), an asthma prevalence rate of 15.4%. Table 1 shows that asthma is very consequential for health.

Table 1. Impacts of Asthma for Children Age 2-17, NSCH 2011/12
Parent/caregiver reports child's health status is excellent or very good

	2 - 5 years	6 – 11 years	12 – 17 years	Total
All Children living in Metropolitan Statistical Areas				
Asthma	59.8 %	69.6 %	74.3 %	70.1 %
No asthma	87.8 %	85.3 %	85.1 %	85.9 %
Overall	84.9 %	82.8 %	83.1 %	83.4 %
Difference	-28.0 %	-15.7 %	-10.8 %	-15.8 %
Children living in MSAs with Asthma				
All Children	59.8 %	69.6 %	74.3 %	70.1 %
Black or Latino	52.1 %	64.1 %	66.4 %	62.9 %
Not Black/Latino	66.5 %	74.6 %	80.4 %	76.1 %
Difference	-14.4 %	-10.5 %	-14.0 %	-13.2 %

We find overall a 15.8% drop in the proportion of parents who report their child's health as very good or excellent among those who have asthma, and almost twice that in younger children. Because 2 of our networks are in the greater NYC area, these data highlight children who live in more urban areas. Outside of urban areas both prevalence and gap between those with and without asthma are slightly higher (each ~17%). Effective delivery of guideline-based care can reduce the gap and decrease consequences of uncontrolled asthma, such as emergency room use and hospitalizations; better asthma care is beneficial and needed across the spectrum of children and primary care settings.[34-40] We find compelling evidence that the failure to effectively deliver guideline-based care contributes significantly to the lower health ratings for children with asthma, including for the 3.4 million urban Black and Hispanic children (age 2-17 years) with asthma. About 60% of these children are low income and have public insurance. We further are persuaded by evidence that quality of life and the quality of asthma management are associated specifically with such factors as family satisfaction with the nature of shared decision making.[41]

Citations for data demonstrating high priority

1. PCORI. PCORI Funding Annoouncement: Treatment Options for Afircan Americans and Hispanics/Latinos with Uncontrolled Asthma. 2013 [cited 2013 September 18]; Available from: http://pcori.org/assets/2013/06/PCORI-Asthma-PFA-061813.pdf.
2. Marcano-Belisario, J., Greenfield G, Huckvale K, Gunn LH, Car J, Apps for asthma self-management: a systematic assessment of content and tools. Cochrane Database Syst. Rev., 2012(8).
3. Health, O.o.M. Asthma and African Americans. [Fact Sheet]. 2012 [cited 2013 August 28]; Available from: http://minorityhealth.hhs.gov/templates/content.aspx?ID=6170.
4. Health, O.o.M. Asthma and Hispanic Americans. [Fact Sheet]. 2012 [cited 2013 August 28]; Available from: http://minorityhealth.hhs.gov/templates/content.aspx?ID=6173.
5. Wennergren, G., Strannegard I, Asthma hospitalizations continue to decrease in schoolchildren but hospitalization rates for wheezing illnesses remain high in young children. Acta Paediatr, 2002. 91(11): p. 1239-1245.
6. Wisnivesky, J., Lorenzo J, Lyn-Cook R, et al., Barriers to adherence to asthma management guidelines among inner-city primary care providers. Annals of Allergy, Asthma & Immunology, 2008. 101(3): p. 264-270.
7. DiSantostefano, R., Davis K, Yancey S, Crim C, Ecologic analysis of asthma-related events and dispensing of inhaled corticosteroid- and salmeterol-containing products. Ann Allergy Asthma Immunol, 2008. 100(6): p. 558-565.
8. Crocker, D., Kinyota S, et al., Effectiveness of home-based, multi-trigger, multicomponent interventions with an environmental focus for reducing asthma morbidity: a community guide systematic review. Am J Prev Med, 2011. 41(2): p. S5-S32.
9. Gustafson, D., Wise M, et al., The effects of combining Web-based eHealth with telephone nurse case management for pediatric asthma control: a randomized controlled trial. J Med Internet Res., 2012. 14(4): p. e101.
10. Program, T.N.A.E.a.P., Expert Panel Report 3: Guidelines for the Diagnosis and Management of Asthma. J Allergy Clin Immunol, 2007. 120(5 Suppl): p. S94-138.
11. Celano, M., Holsey CN, Kobrynski LJ, Home-based family intervention for low-income children with asthma: a randomized controlled pilot study. J Fam Psychol., 2012. 26(2): p. 171-178.
12. Zuniga, G., Kirk S, et al., The impact of asthma health education for parents of children attending head start centers. J Community Health, 2012. 37(6): p. 1296-1300.
13. Blaakman, S., Tremblay PJ, Halterman JS, Fagnano M, Borrelli B, Implementation of a community-based secondhand smoke reduction intervention for caregivers of urban children with asthma: process evaluation, successes and challenges. Health Educ Res, 2013. 28(1): p. 141-152.
14. Martin, M., Catrambone CD, et al., Improving asthma self-efficacy: developing and testing a pilot community-based asthma intervention for African American adults. J Allergy Clin Immunol, 2009. 123(1): p. 153-159.
15. Seid, M., D´Amico E, Varni JW, et al., The In Vivo Adherence Intervention For at Risk Adolescents With Asthma: Report of a Randomized Pilot Trial. Journal of Pediatric Psychology, 2012. 37(4): p. 390-403.
16. Edwards, J., INSPIRE curriculum delivered in a faith-based setting. Fam Community Health., 2010. 33(2): p. 117-122.PRINCIPAL INVESTIGATOR
17. Press, V., Pappalardo AA, et al., Interventions to improve outcomes for minority adults with asthma: a systematic review. J Gen Intern Med, 2012. 27(8): p. 1001-1015.
18. DeJongh, T., Gurol-Urganci I, Vodopivec-Jamsek V, Car J, Atun R, Mobile phone messaging for facilitating self-management of long-term illnesses. Cochrane Database Syst. Rev., 2012.
19. James, T., Fine M, Monitoring Asthma Control Using Claims Data And Patient-Reported Outcomes Measures. P.T., 2008. 33(8): p. 454-466.
20. Bender, B., Overcoming barriers to nonadherence in asthma treatment. J Allergy Clin Immunol, 2002. 109(6): p. 554-559.
21. Okelo, S., Eakin M, et al., The Pediatric Asthma Control and Communication Instrument asthma questionnaire: for use in diverse children of all ages. J Allergy Clin Immunol, 2013. 132(1): p. 55-62.

22. Benavides, S., Rodriquez JC, Maniscalco-Feichtl M, Pharmacist involvement in improving asthma outcomes in various healthcare settings: 1997 to present. Ann Pharmacother. , 2009. 43(1): p. 85-97.
23. Drotar, D., Physician behavior in the care of pediatric chronic illness: association with health outcomes and treatment adherence. J Dev Behav Pediatr., 2009. 30(33): p. 246-254.
24. Weinstein, A., The potential of asthma adherence management to enhance asthma guidelines. Ann Allergy Asthma Immunol, 2011. 106(4): p. 283-291.
25. Agency, E.P. President's Task Force on Environmental Health Risks and Safety Risks to Children:Coordinated Federal Action Plan to Reduce Racial and Ethnic Asthma Disparities. 2012 [cited 2013 August 28]; Available from: http://www.epa.gov/

Numerator Statement: The numerator estimates the number of emergency department (ED) visits for asthma among children being managed for asthma. Since most hospitalizations for asthma are from the ED and many ED visits that result in hospitalization are not captured in encounter data, a numerator event may be identified either as an ED visit or as a hospitalization.

Denominator Statement: The denominator represents the person time experience among eligible children with identifiable asthma. Assessment of eligibility is determined for each child monthly. The total number of child months experienced is summed and divided by 1200 to achieve the units of 100 child years for the denominator.

Assessing eligibility for the denominator requires 2 years of data, the reporting year and the 12 month period before the reporting year. (See Appendix 1, Figure 1)

Denominator Exclusions: Children with specified concurrent or pre-existing diagnosis and children who have not been consecutively enrolled in the reporting plan for at least three months, as specified in the details section.

Measure Type: Outcome
Data Source: Claims (Only), Claims (Other)
Level of Analysis: Health Plan, Population : Community, County or City, Population : Regional and State

IF Endorsement Maintenance – Original Endorsement Date: Most Recent Endorsement Date:

New Measure -- Preliminary Analysis

Criteria 1: Importance to Measure and Report

1a. Evidence

1a. Evidence. The evidence requirements for a health outcomes measure include providing rationale that supports the relationship of the health outcome to processes or structures of care. The guidance for evaluating the clinical evidence asks if the relationship between the measured health outcome and at least one clinical action is identified and supported by the stated rationale.

Evidence Summary

This measure was previously reviewed by the Pulmonary Standing Committee (March 2016) as NQF 2794 and has been revised and resubmitted. In the measure's evidence and testing forms, content submitted previously is in black; new information is in blue in the measure testing and evidence forms.

The developer provides the following rationale for this outcome (plan or population) measure:

- Accessible, high-quality primary care reduces the need for emergency department (ED) visits by decreasing the number of children who have acute breakthrough episodes requiring the ED.
- Accessible, high-quality primary care reduces the need for ED visits by decreasing the number of children who come to the ED for asthma care better performed in the office setting.
- ED visits/hospitalizations are undesirable outcomes that can be reduced by better primary care management.
- A systematic review of the body of evidence is not required for outcome measures.
- The evidence for this measure is based on clinical practice guidelines for asthma control from the National Heart and Lung and Blood Institutes (NHLBI) (2007): "As a general rule, patients with well-controlled asthma should have: ... no emergency department visits; no hospital stays ...". Grade C = Nonrandomized trials and observational studies. Evidence is from outcomes of uncontrolled or nonrandomized trials or from observational studies.

- Although not required per NQF guidance for outcome measures, the developer also provided three systematic reviews: Interventions to Modify Health Care Provider Adherence to Asthma Guidelines; Cochran Database of Systematic Reviews: Intermittent versus daily inhaled corticosteroids for persistent asthma in children and adults (Review); Quality of Care for Childhood Asthma: Estimating Impact and Implications.

Question for the Committee:

- *Is there at least one thing that the provider can do to achieve a change in the measure results?*

Guidance from the Evidence Algorithm
Assesses health outcome (Box 1) → Relationship between outcome and healthcare action (Box 2) → Pass

Preliminary rating for evidence: ☒ **Pass** ☐ **No Pass**
(Previous review by Pulmonary Committee: Pass - Evidence: Y-21; N-0)

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

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

The developer reports:

- ED visits for asthma are common, may be reduced through improved primary care or community-based interventions, and demonstrate disparities.
- NHLBI NAELPP guideline provides a description of clinical evidence of gaps.
- The developer reports 10.3 million children have asthma and that asthma has a significant impact on health, with an overall a 15.8% drop in the proportion of parents who report their child's health as very good or excellent among those who have asthma, and almost twice that in younger children.
- The developer reports overall rate of ED visits for asthma in NY State Medicaid Managed Care in 2012 is 20.65 per 100 child-years.
  - By age stratum the rates are 47.4 visits per 100 child-years for children 2 to 4 years, 26.0 for children 5 to 11 years, 22.7 for adolescents 12 to 18 years, and 34.1 for adolescents 19 to 21 years.
  - There are differences in performance by race, urbancity, and quartile of poverty.
  - The developer provides additional data demonstrating expected seasonal variations in performance rates.

Disparities
The developer reports:

- Asthma is a critical problem with racial and ethnic disparities and varies by urbanicity. The developer's analysis of National Survey of Children's Health data (NSCH, 2011/12), estimates that 10.3 million children in the United States have been told that they have asthma. Of these children, 7.6 million live in more urban areas that are characterized as metropolitan statistical areas (MSAs), with an asthma prevalence rate of 15.4%.
- The developer reports that, on a yearly and a monthly basis, differences exist in performance by age, urbanicity, race/ethnicity, and level of poverty. Additionally, it identifies disparities in cross tabulations—e.g., the performance rate for children 2 to 4 years in large metropolitan areas is 52.6 visits per 100 child-years compared to those in small metropolitan areas with 26.2 visits per child year, in micropolitan areas with 18.3 visits/100 child-years, and in rural areas with 12.3 visits per 100 child-years.
- The developer reports racial and ethnic differences were notable:
  - For children ages 2 to 4 years, the rate in non-Hispanic Whites was 18.4 visits per 100 child-years, in Asians 19.3 visits per 100 child-years, in Hispanics 53.9 child-years, and in non-Hispanic Blacks 74 visits per 100 child-years.
  - The disparities regardless of age were Black, 41.99 visits per 100 child-years; White, 14.79 visits per 100 child-years; Hispanic, 31.91 visit per 100 child-years.

Question for the Committee:

Does the Committee believe there is a gap in care that warrants a national performance measure?

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

Previous review by Pulmonary Committee: Pass - Performance Gap: H-6; M-14; L-1; I-0

Committee pre-evaluation comments

Criteria 1: Importance to Measure and Report (including 1a, 1b, 1c)

- Asthma ED visits and hospitalizations are widely viewed as health outcomes for children. The developers provide sufficient links data to conclude that a healthcare process (controller regimens) can affect the outcome.
- I have concerns about the lower age limit of 2 years old. Not seeing substantive guidelines to substantiate inclusion of the 2 year old age group. Where in NHLBI guidelines does it support this for inclusion.
- This is an outcome measures and the developers cite evidence that ED visits are a signal of quality and that interventions by providers can reduce rates of ED visits in this patient population
- The developers provide ample evidence of high rates of ED use and hospitalization, a significant fraction of which is preventable. The marked differences by race/ethnicity, urban residence, and SES are further strong evidence of performance gaps. I would rate this as "high".
- It demonstrates disparities, however, it is not clear how complete New York Medicaid data are with regards to race. What percent of Medicaid enrollee race is incomplete?
- "Distribution of ED visit rates by measured entity (health plan and county level) are shown in item 2b5.2 rather than here. The results shown do demonstrate variation in quality across measured entities. However they do show substantial racial and ethnic disparities in outcome rates (no in performance at the plan or provider level). However this demonstrates that the measure could be used to illuminate and potentially reduce disparities at the population level.
- The developers appear to provide differences in the performance on the measures across plans in the measure reliability section. Although they do not show the distribution of scores they show that performance is different."

Criteria 2: Scientific Acceptability of Measure Properties

2a. Reliability

2a1. Reliability specifications

2a1. Specifications requires the measure, as specified, to produce consistent (reliable) and credible (valid) results about the quality of care when implemented.

Data source(s):

- Administrative claims

Specifications:

- The level of analysis is plan or population
- The developer defines the numerator as: *The numerator estimates the number of emergency department (ED) visits for asthma among children being managed for asthma. To enhance validity, a numerator event may be identified either as an ED visit or as a hospitalization.*
- The denominator for this measure is: *The person time experience among eligible children with identifiable asthma. Assessment of eligibility is determined for each child monthly. The total number of child months experienced is summed and divided by 1200 to achieve the units of 100 child years.*
- The exclusions for the measure are: *Children with specified concurrent or pre-existing diagnosis and children who have not been consecutively enrolled in the reporting plan for at least three months, as specified in the details section.*
- The developer states “If pharmacy data are not available, the measure should be reported with notation that pharmacy data were not used for the assessment of eligibility. This avoids eliminating from the measure those facilities with no link to pharmacies. Our testing reveals that only a small proportion of patients are excluded by not including pharmacy data to establish eligibility.” However, specifics on the amount of missing data or the proportion of patients with missing data were not submitted.

- The numerator and denominator details include the CPT and ICD-9 codes; ICD-10 codes are included in an attachment.
- This outcome measure is not risk adjusted.
- The calculation algorithm is stated in S.14.
- There is no sampling for this measure.

Questions for the Committee:

- *Are the definitions and codes for "managed for identifiable asthma" and "asthma related medication" appropriate? Are they specific enough so they can be reliably collected by different parties?*
- *The developer notes the goal of the measure is to assess how many children are visiting the ED for asthma treatment. Is the numerator specifications 'either an ED visit **or hospitalization**' appropriate?*
- *Are all the data elements clearly defined? Are all appropriate codes included?*
- *Is it likely this measure can be consistently implemented?*
- *Is the potential variability in access to/inclusion of pharmacy data a concern?*

2a2. Reliability Testing, Testing attachment

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.

SUMMARY OF TESTING

Reliability testing level ☐ **Measure score** ☒ **Data element** ☐ **Both**

Reliability testing performed with the data source and level of analysis indicated for this measure ☒ **Yes** ☐ **No**

Method(s) of reliability testing

- Per NQF guidance separate reliability testing is not required if data element-level validity testing is performed, however, the developer also stated it performed zero inflated poisson (ZIP) approach. This approach is generally not considered a demonstration of reliability, but rather a demonstration of differences in performance.

Results of reliability testing

- The current information provided is not sufficient to demonstrate reliability. NQF staff have requested more information from the developer, but it is not yet available.

Questions for the Committee:

- *Is the test sample adequate to generalize widespread implementation?*
- *Do the results demonstrate sufficient reliability so that differences in performance can be identified?*

Guidance from the Reliability Algorithm: Precise specifications (Box 1)→ Empirical reliability testing (Box 2) → Empirical validity testing of patient level data (Box 3) → Insufficient

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

RATIONALE: The current information is not sufficient to demonstrate reliability.

Previous review by Pulmonary Committee: Pass - Reliability: H-2; M-17; L-2; I-0

2b. Validity

2b1. Validity: Specifications

2b1. Validity Specifications. This section should determine if the measure specifications are consistent with the evidence.

Specifications consistent with evidence in 1a. ☒ **Yes** ☐ **Somewhat** ☐ **No**

Specification not completely consistent with evidence

- The goal of the measure is to assess how many children are visiting the ED for asthma treatment. According to the developer, ED visits for asthma are a function of a sick child who needs to be seen; poor access to high-quality primary care; or poor quality management of a chronic condition. The rate should be low, but not zero. The numerator of children with undesirable visits and a denominator of children with identified asthma are consistent with this evidence.

Question for the Committee:

o *Are the specifications consistent with the evidence?*

2b2. Validity testing

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.

SUMMARY OF TESTING

Validity testing level ☐ **Measure score** ☒ **Data element testing against a gold standard** ☐ **Both**

Method of validity testing of the measure score:

- ☐ **Face validity only**
- ☐ **Empirical validity testing of the measure score**

Validity testing method:

- The developer states: "Assessment of the capacity to identify the eligible population and qualifying events was performed in NY State Medicaid data in both 2011 and 2012 reporting years. For MCO analysis we analyzed both with the 18 plans that had 900 or more children contributing to the denominator and with the 20 plans that contributed at least 1000 months of person time to the denominator."
- For both the numerator and denominator, the developer relies on literature to support its conclusion of the validity of administrative data elements to identify children who are being managed with identifiable asthma. Per NQF policy:
  - Prior evidence of validity of data elements can be used, including published data, provided it includes the same data elements; uses the same data type; and is conducted on an appropriate sample (i.e., representative, adequate numbers, etc.)
  - The developer attests that the data elements match those assessed in the literature.
- However, the developer did not cite the full range of literature (as cited in the other submission, #2816) to support this measure.
- The developer used NY State Medicaid Managed Care claims data for its analyses.
- The developer states it used face validity, but its face validity assessments involve construct of the measure. NQF specifically requires that face validity be at the measure score level and that the assessment be that the measure ***score*** can distinguish good from poor quality.
- The developer also cites score-level validity testing of two previously-endorsed asthma measures as evidence of data-element level validity. However, this does not meet NQF's requirements for demonstration of data element validity.
  - The developer states there is "nearly complete overlap of the denominator codes for this measure and there is overlap of the denominator elements." The developer states that where codes differ, they were specific decisions by its expert panel.
  - The information provided is insufficient for denominator validity as the specific differences and how they do or do not affect the validity are not described.

Validity testing results:

The developer reports:

- "The literature also supports the use of claims data to identify the presence of asthma. We use administrative data to identify the age of the child, various stratification variables and the presence of asthma, as well as the presence of an asthma ED visit or hospitalization. These are routinely used to support billing by CMS,

Medicaid, and private insurers and are routinely used in quality measurement. There is moderate agreement (kappa = 0.45 – 0.50) when comparing administrative data regarding the presence of constructs such as recent asthma attacks, use of asthma medications, attack or medication, attack and medication, using 1 year of administrative claims data. The agreement improves from 0.55 to 0.60 when using two years of data as this measure does."

Questions for the Committee:

- *Is the test sample adequate to generalize for widespread implementation?*
- *Do the results demonstrate sufficient validity so that conclusions about quality can be made?*
- *Do you agree that the score from this measure as specified is an indicator of quality?*

2b3-2b7. Threats to Validity

2b3. Exclusions:

The developer provides the following information:

- Denominator exclusions include: Children with concurrent or pre-existing: Chronic Obstructive Pulmonary Disease (COPD) diagnosis, Cystic Fibrosis diagnosis, or Emphysema diagnosis. Children who were not consecutively enrolled in the reporting plan for three consecutive months ending in the reporting month are excluded.
- There are no numerator exclusions.
- <=2.5% potentially eligible children were excluded by clinical diagnoses.
- Exclusions are clinical and represent construct validity rather than statistical considerations. Longer continuous enrollment requirements would harm the validity of the measure since more children with multiple diagnosis would have been excluded. The 3-month continuous enrollment requirement is provided so that the child is under the management of the health plan, which is the accountable unit.

Question for the Committee:

- *Are any patients or patient groups inappropriately excluded from the measure?*

2b4. Risk adjustment: **Risk-adjustment method** ☐ **None** ☐ **Statistical model** ☒ **Stratification**

Conceptual rationale for SDS factors included? ☒ **Yes** ☐ **No**

SDS factors included in risk model? ☐ **Yes** ☒ **No**

Risk adjustment summary

The developer provides the following information:

- "Specifications for this measure require stratification and reporting by age group only and also within age group by race/ethnicity." The submission indicates that the stratification is informational, not to control for patient characteristics.
- The developer states biological data and national guidelines do not support risk adjustment to control for patient characteristics.

Conceptual analysis of the need for SDS adjustment:

- The developer notes additional stratification variables are optional (e.g., rurality/urbanicity and county level of poverty), but may be required by the accountability entity or reported by the reporting entity. According to the developer, "risk adjustment is not critical for interpreting the results or for validity, but ... stratification is informative to help to promote like to like comparisons and allow for plans to demonstrate how they do on specified subgroups. Such voluntary stratification specified in the measure helps to mitigate against the potential for misinterpretation and unintended consequences."
- The developer acknowledges the association of the risk factors with performance on the measure, but states risk adjustment is not justified by such differences as "either acceptable or unmodifiable by health care," and

posits that evidence exists that primary care, adherence to guidelines, and other interventions can reduce or eliminate the impact of the risk factors.

Empirical analysis of the SDS factors:

- The developer also found that urban counties perform differently than rural counties and the size of the county impacts performance.
- The developer also found differences in performance by race: ED utilization of Blacks is significantly different from Whites ($p<0.01$); ED utilization of Hispanics is significantly different from Whites ($p<0.01$); ED utilization of Blacks and Hispanics are significantly different from one another ($p<0.01$).

Question for the Committee:

- *Do you agree with the developer's rationale that there is no conceptual basis for adjusting this measure for SDS factors?*

2b5. Meaningful difference *(can statistically significant and clinically/practically meaningful differences in performance measure scores can be identified)*:

- The developer used NY State Medicaid Managed Care claims data, comparing 18 plans that included at least 900 children in the denominator, which "yielded statistically significant differences among plans and among counties, whether or not we controlled for age group and/or race and ethnicity and/or urbanicity. This is true also when we analyzed stratified by age group."
  - The developer notes that partially managed plans had a statistically higher rate than fully managed HMO plans. Among all 18, the mean rate is 15.7, with a standard deviation of 6.0.
- The developer states that "Poisson Regression analyses indicate significant differences by health plans and by counties, whether or not controlling for age group and race/ethnicity."
  - The ZIP models showed that even after controlling for age, urban and rural counties performed differently, as did suburban and rural counties vs small urban counties. There were also statistically significant differences in ED usage between whites and blacks, blacks and Hispanics, and whites and Hispanics.
- The developer analyzed meaningful differences related to the stratification subpopulations (e.g., within a state), reporting it performed Chi-square analysis and t-statistics. The developer states "differences between major groups were statistically significant", $p<0.05$.
- The developer states the measure is sensitive enough to detect meaningful differences as observed within a population across counties and between counties and NYC.
- The developer also notes that "Comparing to a randomly selected index plan, 14 of 17 plans had statistically significant differences in performance with the median and modal p-value being <0.001. Non-significant plans' p-values=0.08, 0.16 and 0.88."

Questions for the Committee:

- *Does this measure identify meaningful differences about quality at the health plan level?*
- *Does this measure identify meaningful differences about quality at the population level?*

2b6. Comparability of data sources/methods:

N/A

2b7. Missing Data

The developer notes the following:

- Since administrative claims are used, the extent of missing data is expected to be low. There were a total of three children in our analysis of children with identifiable asthma who dropped out of the analyses because of any missing data element.

- The developer states, “Our analyses found that the absence of pharmacy data would reduce only slightly (as we recall, less than 1%) the number of children identified as having identifiable asthma; no specific data are provided in the submission. This finding became apparent during alpha testing of our specifications and was incorporated into our specifications as a permissive allowance when pharmacy data were not available. We have not located the original analysis and hope for the NY State team to replicate the analysis by the time of the Committee meeting.”
    - The developer also states that “Systems unable to integrate pharmacy data into the eligibility analysis would have a minimally higher risk population than those with pharmacy claims. The specifics of the definitions and the limited impact of pharmacy claims on eligibility combine to make the expected impact of this on the rate of ED visits to almost zero. They are included in the identification of denominator because our expert panel directed us to do so…. for this measure, pharmacy data is used only to complement other utilization data when determining eligibility.”

Guidance from the Validity Algorithm
Measure specifications consistent with evidence (Box 1)→ All potential threats to validity are empirically assessed (Box 2)→ Insufficient

The highest possible rating is INSUFFICIENT.

Preliminary rating for validity: ☐ **High** ☐ **Moderate** ☐ **Low** ☒ **Insufficient**
RATIONALE: The current submission is insufficient, until further information is submitted regarding risk stratification.
Previous review by Pulmonary Committee: Consensus Not Reached - Validity: H-0; M-10; L-11; I-0

Committee pre-evaluation comments

Criteria 2: Scientific Acceptability of Measure Properties (including all 2a, 2b, and 2d)

- The general constructs are reasonable. Specific considerations include:
    - The combination of ED visits and hospital admissions has reasonable face validity, and the data presented support their combination.
    - The codes used in the numerator appear reasonable.
    - The eligibility screen for the denominator is complex. The inclusion of events in the prior year PLUS events in the months of the measurement year prior to the measurement month is unusual. The developers do not really explain why a child in January of the measurement year would have the prior 12 months to achieve the criteria for ""identifiable asthma"" while a child in November of the measurement year would have 22 mos (12 in the prior year, and the 10 months before November in the measurement year). (If I have misunderstood the approach, my apologies.) I am not sure that it introduces bias that would affect the comparison of health care entities on this measure.
    - I am puzzled by the inclusion of children >5 on the basis of 1 ambulatory visit plus an ED visit in the measurement month- only for children in this age group.
    - The inclusion of visits with ""bronchitis"" is unusual as a criterion for identifiable asthma. I do not see data on how including or excluding these diagnoses would change the denominator in a given health plan or population.
    - The expert panel process is described well. No rationale is given for why short acting beat-agonists are excluded in the denominator of ""identifiable asthma"". "
- Would seek additional clarification from the developers as to whether all inclusion criteria are assess within the past month. For example, a hospitalizations with a principal of secondary discharge diagnosis of asthma counts only if it happened in the last month? If so this measure cohort is clearly focused on a population of patients with active asthma symptoms or recent exacerbations. This seems logical as these are the patients at greatest risk of using the ED and for whom providers should be actively engaged in prevention or early management to avoid ED visits.
- "The reliability testing appears reasonable. The developers tested this measure in a large Medicaid claims environment. The use of zero inflated poisson models makes intuitive sense as described, but this does not seem to be acceptable by the NQF. This should be reviewed by a statistician. Evidence on the reliability of the data elements is also specified in terms of how data in medical charts are used to reliably code the diagnoses. There is not a robust discussion of the impact of including secondary vs. primary diagnosis codes in the numerator definition. I would like to understand better why the ZIP method is not seen as contributing to reliability leading to the preliminary assessment of the NQF staff as ""Inadequate"".

- Some reliability testing is provided under the validity items. However, they do not present a true test re-test validity test of the measure score at the plan or county level. I would encourage them to provide this to the committee
- "Validity testing was done only in one state's Medicaid data. While it would have been better if this had been across several states, or included commercial plans as well, this is a fairly large and diverse population in which to test this measure.
- The validity of the data elements is well-supported by the literature, though the developers do not really provide evidence for how changes in specifications (not using short-acting beta agonists) might affect validity. I agree that the developers do not explicitly make a claim of face validity of the score as a measure of quality, but it is implied and I think a reasonable conclusion. The findings by race/ethnicity and urban residence are all very consistent with studies in the literature, further supporting the validity."
- My primary concern is over the short enrollment period requirement of 3 months for inclusion in the denominator. According to the authors, this cohort accounts for 20% of the denominator. 3 months of enrollment (when you factor in that it takes around 2 full weeks for insurance card to arrive to home of new enrollee; and another 30 days or so for PCP selection; it does not seem valid to consider 3 months of enrollment as sufficient time for at least one physician visit to occur (which would either be a sick or well-child visit), to include asthma diagnosis, and expect asthma to be managed well-enough to prevent ED visit. seems like an unreasonable expectation. Have the authors compared their outcome data with and without the 20% who were enrolled for only 3 consecutive months?
- Measure score validity testing is not provided. Developers should be encouraged to include a systematic rating of the validity of the measure score from their TEP if available.
- The developers describe the impact of lack of pharmacy data for determining eligibility for a measurement month. The developers provide an argument for why this will not affect the measure very much.
- The authors did not discuss if they allowed for a three month run-out of the administrative data to ensure data completeness for claims. How much time elapsed after the end of the reporting period before the data were collected from the state or health plans? Need to hear more about data completeness and what percent of claims were complete (paid).
- "Exclusions appropriate with minor impact on cohort. Risk adjustment -- I agree with the developer's approach to stratify the measure by age group and to provide stratified results to illuminate disparities. The exclusion of children with chronic lung disease is likely sufficient and additional risk-adjustment not needed. However, I would have preferred to see some analysis to ensure that there is no biased distribution of chronically ill asthma patients across health plans. Meaningful differences: adequately demonstrated by developers. Comparability and missingness do not seem to be relevant for this measure.

Criterion 3. Feasibility

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 claims.
- The developer reports there are no fees.

Question for the Committee:

o *Are the required data elements routinely generated and used during care delivery?*

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

Previous review by Pulmonary Committee: Pass - Feasibility: H-15; M-6; L-0; I-0

Committee pre-evaluation comments

Criteria 3: Feasibility

- The feasibility of the basic measure is adequate, since it relies on data that is available in claims data. As the developers note, some of the stratification variables (e.g. race/ethnicity) may not be available and may vary in how reliable they in different systems. The developers argue that the measure will push health systems to improve data collection and reporting. This is possible, but not assured.

- Highly feasible as dependent on information available in hospital billing/claims. However, stratification would depend on accuracy of demographic information at the plan or state level which may be poor and highly variable. I would have like to have see some more detailed description of quality and quality assessment of these data.

Criterion 4: Usability and Use

4. Usability and 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.

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
- This is a new measure so current use is not required.
- The developer is working on specific plans for dissemination and use.
- The developer is discussing application and use of this measure with New York State Medicaid.
- The developer plans for the measure to be used for an accountability application within three years of NQF endorsement and public reporting within six years of initial endorsement.

Improvement results
- As a new measure, the developer does not present progress on improvement.
- The developer states a variety of stakeholders would benefit from this measure, e.g., clinicians, health systems, state and healthcare agencies, researchers, etc.

Unexpected findings (positive or negative) during implementation
N/A

Potential harms
- The developer reports no unintended negative consequences to individual or populations during testing.
- The developer reports possible unintended/negative consequences and recommends against the following:
  - Comparing individual health care professionals.
  - A single hospital comparison because this measure is intended to measure system performance not the hospital performance.
  - Measuring anything other than large practices or integrated delivery systems that own their own risk and manage inpatient and outpatient care or that have access to all payer data sources.

Vetting of the measure
- None provided

Feedback:
- No feedback provided on QPS. MAP has not reviewed this measure for inclusion in any federal program.

Question for the Committee:
- *Do the benefits of the measure outweigh any potential unintended consequences?*

Preliminary rating for usability and use: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**
Previous review by Pulmonary Committee: Pass - Usability and Use: H-4; M-11; L-5; I-1
Not recommended on Overall Vote for Endorsement: Y-3; N-15

Committee pre-evaluation comments

Criteria 4: Usability and Use

- The measure is not currently in use. The basic measure should be useable, but use of some of the stratifications may have additional barriers at least in the short to medium term.

Criterion 5: Related and Competing Measures

Related or competing measures

The developer did not include information on any of the related or competing measures. However, NQF staff identified the following measures that may be related and/or competing.

- 0047: Asthma: Pharmacologic Therapy for Persistent Asthma
- 0728: Asthma Admission Rate (PDI 14)
- 1800: Asthma Medication Ratio
- 2414: Pediatric Lower Respiratory Infection Readmission Measure
- 2816: Appropriateness of Emergency Department Visits for Children and Adolescents with Identifiable Asthma (*submitted by the same developer for review in this project)*

Harmonization

No information provided

Endorsement + Designation

The "Endorsement +" designation identifies measures that exceed NQF's endorsement criteria in several key areas. After a Committee recommends a measure for endorsement, it will then consider whether the measure also meets the "Endorsement +" criteria.

This measure is a candidate for the "Endorsement +" designation IF the Committee determines that it: meets evidence for measure focus without an exception; is reliable, as demonstrated by score-level testing; is valid, as demonstrated by score-level testing (not via face validity only); and has been vetted by those being measured or other users.

Eligible for Endorsement + designation: ☐ **Yes** ☒ **No**

RATIONALE IF NOT ELIGIBLE: This measure is not eligible for Endorsement+ because it has face validity testing only and has not been vetted by those being measured or other users.

Pre-meeting public and member comments

- None

NATIONAL QUALITY FORUM—Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:** Click here to enter NQF number
Measure Title: Rate of Emergency Department Visit Use for Children Managed for Identifiable Asthma: Visits per 100 Child-years
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: 12/14/2016

Instructions

- *Complete 1a.1 and 1a.12 for all measures.*
- *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:

- Health outcome: [3] a rationale supports the relationship of the health outcome to processes or structures of care. Applies to patient-reported outcomes (PRO), including health-related quality of life/functional status, symptom/symptom burden, experience with care, health-related behavior.
- 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.

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 U.S. Preventive Services Task Force (USPSTF) grading definitions and methods, or Grading of Recommendations, Assessment, Development and Evaluation (GRADE) guidelines.

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

☒ Health outcome: ED asthma visits for children with identifiable asthma

☐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.12 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.

Emergency department (ED) visits are often linked to the management of a child's asthma. Emergency Department Asthma was the topic assigned to CAPQuaM for measurement. ED visits for children with asthma is an outcome measure of intrinsic value. It represents utilization of an expensive service and constitutes a burden on children and their families. Two literature reviews as well as focused reviews that we have done to supplement the extensive review of the literature confirms the importance of an integrated approach to managing the health care of children with asthma. There is abundant evidence that ED visits are common, may be reduced through improved primary care or community-based interventions, and demonstrate disparities (1-11, 12-19). Asthma is generally recognized to be an ambulatory care sensitive condition. Nonetheless, we perceive and our panel articulated that the rate for ED visits ought not to be 0. So while in general a lower rate represents preferable care, too low a rate could indicate insufficient access to emergency room services. Our overarching conceptual framework that extends beyond this measure is shown in the evidence form.

~~Our measure benefits from a formal development process, CAPQuaM's 360 degree method, which is described in more detail in the measure testing form. The measure and its specifications result from a formal development process for this measure incorporated stakeholder input including a parent focus group, meeting with The Mount Sinai Pediatrics Department's Parent Advisory Council, interviews with primary care clinicians and ED physicians, the CAPQuaM's multidisciplinary scientific team, which includes investigators, a steering committee and a senior advisory board of nationally prominent figures. The measure also benefits from a national multidisciplinary Expert Panel which utilized a RAND type modified Delphi method to guide our specifications.~~

When epidemiologists describe how frequently something occurs the preferred measure is typically an incidence density, or rate. In contrast to a risk or proportion, the incidence density has as its denominator a measure of the extent of potential exposure in the population, expressed in people-years. This measure represents an advance in the measurement of healthcare performance for children: it incorporates this formulation both to enhance its interpretation (because it has a specific epidemiological meaning) and to limit distortion if sick children move in or out of eligibility for the measure. (20)

Further clinical evidence of gaps are demonstrated in the description by NHLBI's NAELPP guideline, cited in the evidence form, Schatz and colleagues study describing the relationship between asthma control and asthma exacerbations in managed care (21), and Fuhlbrigge et al's confirmation that medications can work to reduce ED visits for asthma but are used sub optimally (22). When children with asthma experience adequate management of chronic conditions and have access to coordinated care, a reduction in hospital rates is likely to occur. (23) Children who are linked to continuous care utilize less overall care, including ED care. (23)

The following diagram presents an overview of how CAPQuaM conceptualizes asthma ED visits for children with asthma.

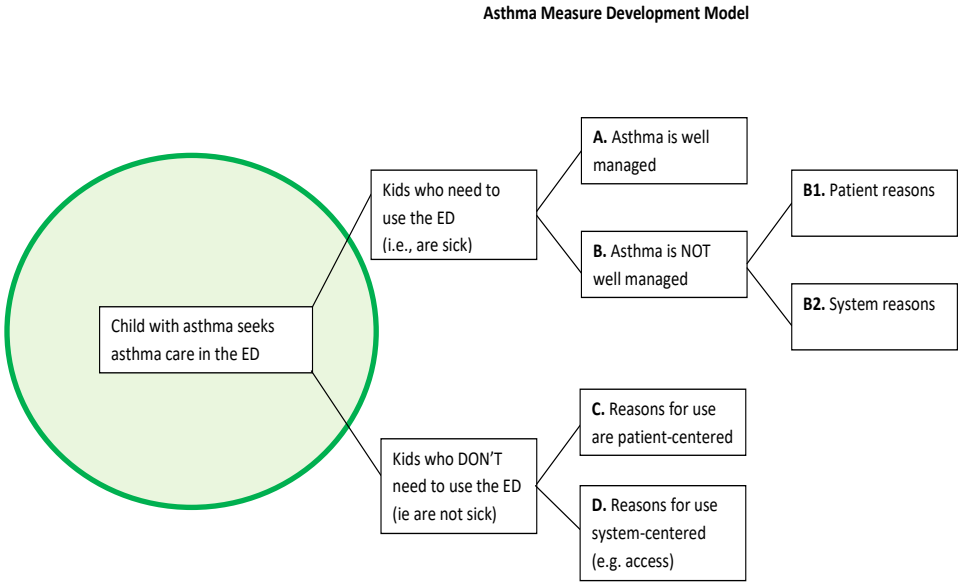


Figure Notes: The green circle highlights that this measure identifies which children who present to the emergency room should be considered to represent an ED visit for a child who is being managed for identifiable asthma.

References

1. Adams, R.J., B.J. Smith, and R.E. Ruffin, Factors associated with hospital admissions and repeat emergency department visits for adults with asthma. Thorax, 2000. 55(7): p. 566-73.
2. Bahadori, K., et al., Economic burden of asthma: a systematic review. BMC Pulm Med, 2009. 9: p. 24.
3. Weiss, K.B., P.J. Gergen, and T.A. Hodgson, An economic evaluation of asthma in the United States. N Engl J Med, 1992. 326(13): p. 862-6.
4. Coventry, J.A., M.S. Weston, and P.M. Collins, Emergency room encounters of pediatric patients with asthma: cost comparisons with other treatment settings. JAmbul Care Manage, 1996. 19(2): p. 9-21.
5. Association, A.L. American Lung Association Fact Sheet on Asthma and Children. October 2012 [cited 2013 May 20]; Available from: http://www.lung.org/lung-disease/asthma/resources/facts-and-figures/asthmachildren-fact-sheet.html#5.
6. Fuhrman, C., et al., Hospitalizations for asthma in children are linked to undertreatment and insufficient asthma education. J Asthma, 2011. 48(6): p. 565-71.
7. Sawicki, G.S., et al., Uncontrolled asthma in a commercially insured population from 2002 to 2007: trends, predictors, and costs. J Asthma, 2010. 47(5): p. 574-80.
8. Manice, M., Exploring the relationship between parental shared decision-making practices and acute asthma exacerbations among children age 0-17. 2013, Icahn School of Medicine at Mount Sinai: New York, NY.
9. Cerdan, N.S., et al., Asthma severity in children and the quality of life of their parents. Appl Nurs Res, 2012. 25(3): p. 131-7.
10. Fiese, B., et al., Family climate of routine asthma care: associating perceived burden and mother-child interaction patterns to child well-being. Fam Process, 2008. 47(1): p. 63-79.
11. Okelo, S.O., et al., Emotional quality-of-life and outcomes in adolescents with asthma. J Pediatr, 2004. 145(4): p. 523-9.
12. Talreja, N., et al., Modifiable factors associated with severe asthma exacerbations in urban patients. Ann Allergy Asthma Immunol, 2012. 109(2): p.128-32.

13. Auger, K.A., et al., Medical home quality and readmission risk for children hospitalized with asthma exacerbations. Pediatrics, 2013. 131(1): p. 64-70.
14. Ducharme, F.M., et al., Written action plan in pediatric emergency room improves asthma prescribing, adherence, and control. Am J Respir Crit Care Med, 2011.183(2): p. 195-203.
15. Farber, H.J., Optimizing maintenance therapy in pediatric asthma. Curr Opin Pulm Med, 2010. 16(1): p. 25-30.
16. Smith, S.R., D.B. Wakefield, and M.M. Cloutier, Relationship between pediatric primary provider visits and acute asthma ED visits. Pediatr Pulmonol, 2007.42(11): p. 1041-7.
17. Self, T.H., et al., Reducing emergency department visits and hospitalizations in African American and Hispanic patients with asthma: a 15-year review. J Asthma, 2005. 42(10): p. 807-12.
18. Lara, M., et al., Reducing quality-of-care disparities in childhood asthma: La Red de Asma Infantil intervention in San Juan, Puerto Rico. Pediatrics, 2013. 131 Suppl 1: p. S26-37.
19. Oraka, E., et al., Racial and ethnic disparities in current asthma and emergency department visits: findings from the national health interview survey, 2001-2010. J Asthma, 2013. 50(5): p. 488-96.
20. Rothman, K.J., Modern Epidemiology. 3rd edition ed. 2008, Philadelphia, PA: Lippincott, Williams and Wilkins.
21. Schatz M, Zeiger RS, Yang ST, et al. "Relationship of asthma control to asthma exacerbations using surrogate markers within a managed care database." Am J Managed Care. 16(5):327-333, 2010.
22. Fuhlbrigge A, Carey VJ, Adams RJ. "Evaluation of asthma prescription measures and health system performance based on emergency department utilization." Med Care 42(5):1-7, 2004.
23. Cooley W, McAllister J, Sherrieb K, Kuhlthau K. *Improved outcomes associated with medical home implementation in pediatric primary care.* Pediatrics, 2009. **124**(1): p. 358-364.

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

1a.2 FOR OUTCOME MEASURES including PATIENT REPORTED OUTCOMES- State the rationale supporting the relationship between the health outcome (or PRO) to at least one healthcare structure, process (e.g., intervention, or service).

An abundant literature supports both that emergency department visits and hospitalizations are considered undesirable outcomes for asthma and that at a population level these undesirable outcomes can be reduced by better clinical management, including medication management, the use of asthma action plans, and effective and continuous primary care, among other things. Asthma is considered to be an ambulatory care sensitive condition further reinforcing the consensus in the field that utilization of ED visits and/or hospitalizations are generally (at the population level) preventable when managed in an ambulatory setting within our current knowledge.

1. Accessible high quality primary care reduces the need for ED visits by decreasing the number of children who have acute breakthrough episodes requiring the ED or inpatient setting.
2. Accessible high quality primary care reduces the need for ED visits by decreasing the number of children who come to the ED for asthma care better performed in the office setting.

As ED visits and/or hospitalizations can represent significant cost for families and for the system, asthma is the single most prevalent diagnosis leading to ED visits for children in the USA, urgent asthma visits to the ED can be disruptive for families, and both ED visits and hospitalizations are not free of iatrogenic and nosocomial risk, these outcomes have intrinsic importance.

To support this rationale there are four specific systematic reviews of the evidence cited. Detailed information on each can be found in the Appendix Table 6-9. Highlights are below:

(1) National Heart, Lung, and Blook Institute, National Institutes of Health (NHLBI/NIH) Asthma Guideline 2007

Quick Reference Guide: Asthma control focuses on two domains: 1) reducing impairment --- the frequency and intensity of symptoms… and 2) reducing risk – the likelihood of future asthma attacks… [later described as "prevent exacerbations]

At the population level ED visits and hospitalizations represent failures of asthma control.

Asthma Guidelines:
- **Following science-based guidelines *works***
- Not only do they have the potential to improve a patient's *quality* of life; they can potentially *save a life*.

National asthma guidelines have been updated: In 2007, the National Asthma Education and Prevention Program (NAEPP), coordinated by the National Heart, Lung, and Blood Institute (NHLBI), released its third set of clinical practice guidelines for asthma. The Expert Panel Report 3—Guidelines for the Diagnosis and Management of Asthma (EPR-3) reflects the latest scientific advances in asthma drawn from a systematic review of the published medical literature by an NAEPP-convened expert panel. It describes a range of generally accepted best-practice approaches for making clinical decisions about asthma care.

The EPR-3 emphasizes the importance of asthma control and focuses on two domains—current impairment and future risk—by which to assess asthma severity (for initiating therapy) and asthma control (for ongoing monitoring). EPR-3 also includes an expanded section on childhood asthma (with an additional age group), new guidance on medications, new recommendations on patient education in settings beyond the physician's office, and new advice for controlling environmental exposures that can cause asthma symptoms.

Asthma can be controlled
Scientific evidence clearly shows that most people could control their asthma by following current asthma clinical practice guidelines. With proper care, people who have asthma can stay active, sleep through the night, and avoid having their lives disrupted by asthma attacks.

As a general rule, patients with well-controlled asthma should have:
- Few, if any, asthma symptoms.
- Few, if any, awakenings during the night caused by asthma symptoms.
- No need to take time off from school or work due to asthma.
- Few or no limits on full participation in physical activities.
- No emergency department visits.
- No hospital stays.
- Few or no side effects from asthma medicines.

KEYPOINTS: OVERVIEW OF MEASURES OF ASTHMA ASSESSMENT AND MONITORING (pg. 36)
- The functions of assessment and monitoring are closely linked to the concepts of severity, control, and responsiveness to treatment:
    - Severity: the intrinsic intensity of the disease process. Severity is measured most easily and directly in a patient not receiving long-term-control therapy.
    - Control: the degree to which the manifestations of asthma (symptoms, functional impairments, and risks of untoward events) are minimized and the goals of therapy are met.
    - Responsiveness: the ease with which asthma control is achieved by therapy.
- Both severity and control include the domains of current impairment and future risk:
    - Impairment: frequency and intensity of symptoms and functional limitations the patient is experiencing or has recently experienced
    - Risk: the likelihood of either asthma exacerbations, progressive decline in lung function (or, for children, reduced lung growth), or risk of adverse effects from medication

KEY DIFFERENCES FROM 1997 AND 2002 EXPERT PANEL REPORTS (pg. 37)

- The key elements of assessment and monitoring are refined to include the separate, but related, concepts of severity, control, and responsiveness to treatment. Classifying severity is emphasized for initiating therapy; assessing control is emphasized for monitoring and adjusting therapy. Asthma severity and control are defined in terms of two domains: impairment and risk.
- The distinction between the domains of impairment and risk for assessing asthma severity and control emphasizes the need to consider separately asthma's effects on quality of life and functional capacity on an ongoing basis (i.e., in the present) and the risks it presents for adverse events in the future, such as exacerbations and progressive loss of pulmonary function. These domains of asthma may respond differentially to treatment.

… p.38: An important point linking asthma severity, control, and responsiveness is that the goals are identical for all levels of baseline asthma severity. A patient who has severe persistent asthma compared to a patient who has mild persistent asthma, or a patient who is less responsive to therapy may require more intensive intervention to achieve well-controlled asthma; however, the goals are the same: in well-controlled asthma, the manifestations of asthma are minimized by therapeutic intervention.

… page 41 regarding identification asthma, one key factor is:
The Expert Panel recommends that the clinician trying to establish a diagnosis of asthma should determine that (EPR–2 1997):

- **Episodic symptoms of airflow obstruction are present.**

This is consistent with how we defined identifiable asthma…

Page 63: It is important to evaluate the frequency, rate of onset, severity, and causes of exacerbations…severe exacerbations leading to ED visits and hospitalizations (Adams et al. 2000; Eisner et al. 2001; Ford et al. 2001; Lieu et al. 1998).

(2) Interventions to Modify Health Care Provider Adherence to Asthma Guidelines: A Systematic Review

Demonstrates several tools are effective in enhancing the quality of care and reduce undesirable outcomes.

(3) Cochran Database of Systematic Reviews: Intermittent versus daily inhaled corticosteroids for persistent asthma in children and adults (Review)

Different approaches to treatment achieve different outcomes in children and adults (Daily achieves better asthma control than intermittent inhaled corticosteroids)

(4) Quality of Care for Childhood Asthma: Estimtating Impact and Implications

Identified multiple gaps in asthma care quality. Key outcomes identified include hospitalizations and emergency department visits. Identified large racial disparities in use of inhaled corticosteroids

In addition to the work cited above, we conducted a scoping review as follows:

We identified key constructs of asthma ED use measures for consideration. We created a table of these constructs in technical and lay language, and listed research questions for the review to answer. Our contractor (a national accrediting body experienced in measure development), prepared for us a literature review in 2 stages and we

supplemented this with targeted reviews as needed to answer specific questions that arose during the measure development process.

The construct table (Appendix; pg 26) was used to guide the review and was the basis for the first round of review. Following the table, we include a list of questions for focused review (Appendix: pg 39) that guided round 2 of the review, which resulted in a detailed summary of 91 articles from the peer-reviewed literature. In addition to this review, the CAPQuaM scientific team conducted an ad hoc series of reviews to answer specific questions such as the reliability of administrative data to identify asthma, and the value of expert panels and the RAND/UCLA appropriateness method. The CAPQuaM degree 360 method starts with a topic area and the measures emerge during the process, in this case necessitating the specified ad hoc reviews.

We searched peer reviewed and gray literature from 1985-2014 over the course of these reviews. Literature was summarized for our expert panel, which met in late 2013.

Our approach to developing this measure stems from a vibrant and scientifically sound tradition regarding measuring performance. We discuss herein research involving the soundness administrative data to identify children with asthma. This is a generally accepted and standard approach with acceptable reliability.

Brook and Davies [1] trace the early history of quality measurement and remind us of the importance of medical chart audit as an approach to quality measurement. Lohr and Brook at RAND and Roos in Manitoba, Canada pioneered the use of electronically-available administrative data (generated by routine health care operations, such as billings) as proxies for health care processes. Administrative data carefully used reduces burden of quality measurement. [2-6]

As the National Committee for Quality Assurance (NCQA) developed the Healthcare Employee Data Information Set (HEDIS) as the de facto measurement system for managed care, attention turned to the use of administrative data for routine performance measurement.

We have used rigorous and transparent methods [14] to assemble a national expert panel that included pediatricians, family physicians, pediatric and general emergency room specialists, a pediatric pulmonologist and a pediatric allergist from practices and medical schools around the country. This work was conducted in collaboration with national clinical societies (AAP, AAFP) and CAPQuaM's diverse other partner organizations, including NY State DoH/Medicaid. NCQA is an important technical consultant and partner. The specific criteria that we operationalize in this measure were all rated by the expert panel with a median score of 8 or 9 on a 9 point scale (9 high) to develop inclusion and exclusion criteria, variables for stratification and so forth. The use of Expert Panels has been demonstrated to be useful in measure development and health care evaluation, including for children.

The literature has demonstrated the reliability of claims data for assessing asthma. Though they have their limitations, these data types have been shown in multiple studies to be a reliable source of information for population level quality measurement. They are currently used for all of the analogous measures of which we are aware, including the former Core Measure and the NCQA measure considering children with persistent asthma.

The use of two years of data to validate the diagnosis of asthma has been found to produce substantial agreement with patient surveys and improves performance over the use of one year of data (28). Others have reported that using administrative databases to identify asthma is both sensitive and specific as compared to review of the primary care physician's office chart (29).

Select additional references documenting other aspects of performance gap, and supporting our process and data sources are also noted (7-13, 15-35).

1. Brook, R.H. and A. Davies-Avery, *Mechanisms for Assuring Quality of US Medical Care Services: Past, Present, and Future*. 1977: Santa Monica, CA.
2. Brook, R.H. and K.N. Lohr, *Efficacy, effectiveness, variations, and quality. Boundary-crossing research.* Med Care, 1985. **23**(5): p. 710-22.
3. Lohr, K.N., *Use of insurance claims data in measuring quality of care.* Int J Technol Assess Health Care, 1990. **6**(2): p. 263-71.
4. Roos, L.L., et al., *Using Administrative Data Banks for Research and Evaluation: A Case Study.* Evaluative Research 1979. **236**(3).
5. Roos, L.L., Jr., J.P. Nicol, and S.M. Cageorge, *Using administrative data for longitudinal research: comparisons with primary data collection.* J Chronic Dis,1987. **40**(1): p. 41-9.
6. Roos, L.L., Jr., et al., *How good are the data? Reliability of one health care data bank.* Med Care, 1982. **20**(3): p. 266-76.
7. Dresser, M.V., et al., *Clinical quality measurement. Comparing chart review and automated methodologies.* Med Care, 1997. **35**(6): p. 539-52.
8. Newton, K.M., et al., *The use of automated data to identify complications and comorbidities of diabetes: a validation study.* J Clin Epidemiol, 1999. **52**(3): p.
199-207.
9. Thompson, B.L., et al., *Measuring clinical performance: comparison and validity of telephone survey and administrative data.* Health Serv Res, 2001. **36**(4): p.
813-25.
10. Angier, H., et al., *Variation in outcomes of quality measurement by data source.* Pediatrics, 2014. **133**(6): p. e1676-82.
11. Weiskopf, N.G. and C. Weng, *Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research.* Journal of the American Medical Informatics Association, 2013. **20**(1): p. 144-151.
12. Pawlson, L.G., S.H. Scholle, and A. Powers, *Comparison of administrative-only versus administrative plus chart review data for reporting HEDIS hybrid measures.* Am J Manag Care, 2007. **13**(10): p. 553-8.
13. NCQA. *National Committee for Quality Assurance*. [cited 2014 7/30/14];
Available from: http://www.ncqa.org/.
14. Brook, R.H., et al., *A method for the detailed assessment of the appropriateness of medical technologies.* International journal of technology assessment in health care, 1986. **2**(01): p. 53-63.
15. DeAngelis, C., P. Fosarelli, and A.K. Duggan, *Use of the emergency department by children enrolled in a primary care clinic.* Pediatr Emerg Care, 1985. **1**(2): p.61-5.
16. Talreja, N., et al., *Modifiable factors associated with severe asthma exacerbations in urban patients.* Ann Allergy Asthma Immunol, 2012. **109**(2): p.128-32.
17. Auger, K.A., et al., *Medical home quality and readmission risk for children hospitalized with asthma exacerbations.* Pediatrics, 2013. **131**(1): p. 64-70.
18. Ducharme, F.M., et al., *Written action plan in pediatric emergency room improves asthma prescribing, adherence, and control.* Am J Respir Crit Care Med, 2011. **183**(2): p. 195-203.
19. Farber, H.J., *Optimizing maintenance therapy in pediatric asthma.* Curr Opin Pulm Med, 2010. **16**(1): p. 25-30.
20. Smith, S.R., D.B. Wakefield, and M.M. Cloutier, *Relationship between pediatric primary provider visits and acute asthma ED visits.* Pediatr Pulmonol, 2007. **42**(11): p. 1041-7.
21. Parchman, M.L. and S. Culler, *Primary care physicians and avoidable hospitalizations.* J Fam Pract, 1994. **39**(2): p. 123-8.
22. Prevention, C.f.D.C.a. *Home-based Multi-trigger, Multi-component interventions*. 2013 [cited 2013 May 20]; Available from: http://www.cdc.gov/asthma/interventions.htm.
23. Price, J.H., et al., *Racial/ethnic disparities in chronic diseases of youths and access to health care in the United States.* Biomed Res Int, 2013. **2013**: p. 1-12; Available in open access http://dx.doi.org/10.1155/2013/787616

24. Homer, C.J., et al., *Does quality of care affect rates of hospitalization for childhood asthma?* Pediatrics, 1996. **98**(1): p. 18-23.
25. Finkelstein, J.A., et al., *Quality of care for preschool children with asthma: the role of social factors and practice setting.* Pediatrics, 1995. **95**(3): p. 389-94.
26. Flores, G., et al., *Keeping children out of hospitals: parents' and physicians' perspectives on how pediatric hospitalizations for ambulatory care-sensitive conditions can be avoided.* Pediatrics, 2003. **112**(5): p. 1021-30.
27. Knudson, A., et al., *Disparities in pediatric asthma hospitalizations.* J Public Health Manag Pract, 2009. **15**(3): p. 232-7.
28. Bindman, A.B., et al., *Preventable hospitalizations and access to health care.* Jama, 1995. **274**(4): p. 305-11.
29. National Heart, Lung and Blood Institute. *Asthma Guidelines*. 2011 February 2011 [cited 2014 7/30/2014].
30. *A Matter of Urgency: Reducing Emergency Department Overuse* in *A NEHI Research Brief* 2010, New England Healthcare Institute.
31. Martin, B.C., *Emergency medicine versus primary care: a case study of three prevalent, costly, and non-emergent diagnoses at a community teaching hospital.* J Health Care Finance, 2000. **27**(2): p. 51-65.
32. Finkelstein, J.A., et al., *Comparing asthma care for Medicaid and non-Medicaid children in a health maintenance organization.* Arch Pediatr Adolesc Med, 2000. **154**(6): p. 563-8.
33. Owens, P.L., et al., *Care of children and adolescents in U.S. hospitals*, in *HCUP Fact Book No. 4*. Agency for Healthcare Research and Quality: Rockville: MD.
34. Pearson, W.S., et al., *State-based Medicaid costs for pediatric asthma emergency department visits.* Prev Chronic Dis, 2014. **11**: p. E108.
35. Taubman, S.L., et al., *Medicaid Increases Emergency-Department Use: Evidence from Oregon's Health Insurance Experiment.* Science, 2014. **343**(6168): p. 263-268

1a.3. SYSTEMATIC REVIEW(SR) OF THE EVIDENCE (for INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURES) 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)

☐ Clinical Practice Guideline recommendation (with evidence review)

☐ US Preventive Services Task Force Recommendation

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

☐ Other

_____

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.

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

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

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

nqf_evidence_attachment_Asthma_1_12_05_16_lk.docx

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

Please update any changes in the evidence attachment in red. Do not remove any existing information. If there have been any changes to evidence, the Committee will consider the new evidence. If there is no new evidence, no updating of the evidence information is needed.

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 PRO-PM (e.g. HRQoL/functional status, symptom/burden, experience with care, health-related behaviors), provide evidence that the target population values the measured PRO and finds it meaningful. (Describe how and from whom their input was obtained.)

IF a COMPOSITE (e.g., combination of component measure scores, all-or-none, any-or-none), SKIP this question and provide rationale for composite in question 1c.3 on the composite tab.

Asthma is a critical problem with racial and ethnic disparities and varies by urbanicity. Adherence to the National Asthma Education and Prevention Programs (NAEPP) Guidelines improves outcomes. [1-32]. We have elsewhere provided other articles, studies, and summaries of evidence to document that ED visits and hospitalizations are typically outcome measures of choice when assessing asthma control.

ED visits for asthma in children are common and expensive. They may result from poor quality of care delivered (failure to adhere to guidelines) as well as from insufficient access to primary care. Asthma is the leading diagnosis leading to urgent care/emergent care provided in emergency departments for children. It is among the most common chronic diseases in children and expenses for asthma care are in the billions of dollars annually. Further, CMS and AHRQ assigned us this measure. In addition to data and citations provided, the team has analyzed 2007 and 2011 waves of the National Survey of Children's Health and confirmed that this parent reported measure both identified a high prevalence of asthma nationwide and significant consequences in terms of parent reported child health for children who have asthma.

Our analysis of National Survey of Children's Health [33] data (NSCH, 2011/12), estimates that 10.3 million children in the U.S. have been told that they have asthma. Of these children 7.6 million live in more urban areas that are characterized as metropolitan statistical areas (MSAs), an asthma prevalence rate of 15.4%. Table 1 shows that asthma is very consequential for health.

Table 1. Impacts of Asthma for Children Age 2-17, NSCH 2011/12
Parent/caregiver reports child's health status is excellent or very good

	2 - 5 years	6 – 11 years	12 – 17 years	Total
All Children living in Metropolitan Statistical Areas				
Asthma	59.8 %	69.6 %	74.3 %	70.1 %
No asthma	87.8 %	85.3 %	85.1 %	85.9 %
Overall	84.9 %	82.8 %	83.1 %	83.4 %
Difference	-28.0 %	-15.7 %	-10.8 %	-15.8 %
Children living in MSAs with Asthma				
All Children	59.8 %	69.6 %	74.3 %	70.1 %
Black or Latino	52.1 %	64.1 %	66.4 %	62.9 %

Not Black/Latino	66.5 %	74.6 %	80.4 %	76.1 %	
Difference		-14.4 %	-10.5 %	-14.0 %	-13.2 %

We find overall a 15.8% drop in the proportion of parents who report their child's health as very good or excellent among those who have asthma, and almost twice that in younger children. Because 2 of our networks are in the greater NYC area, these data highlight children who live in more urban areas. Outside of urban areas both prevalence and gap between those with and without asthma are slightly higher (each ~17%). Effective delivery of guideline-based care can reduce the gap and decrease consequences of uncontrolled asthma, such as emergency room use and hospitalizations; better asthma care is beneficial and needed across the spectrum of children and primary care settings.[34-40] We find compelling evidence that the failure to effectively deliver guideline-based care contributes significantly to the lower health ratings for children with asthma, including for the 3.4 million urban Black and Hispanic children (age 2-17 years) with asthma. About 60% of these children are low income and have public insurance. We further are persuaded by evidence that quality of life and the quality of asthma management are associated specifically with such factors as family satisfaction with the nature of shared decision making.[41]

Citations for data demonstrating high priority

1. PCORI. PCORI Funding Annoouncement: Treatment Options for Afircan Americans and Hispanics/Latinos with Uncontrolled Asthma. 2013 [cited 2013 September 18]; Available from: http://pcori.org/assets/2013/06/PCORI-Asthma-PFA-061813.pdf.
2. Marcano-Belisario, J., Greenfield G, Huckvale K, Gunn LH, Car J, Apps for asthma self-management: a systematic assessment of content and tools. Cochrane Database Syst. Rev., 2012(8).
3. Health, O.o.M. Asthma and African Americans. [Fact Sheet]. 2012 [cited 2013 August 28]; Available from: http://minorityhealth.hhs.gov/templates/content.aspx?ID=6170.
4. Health, O.o.M. Asthma and Hispanic Americans. [Fact Sheet]. 2012 [cited 2013 August 28]; Available from: http://minorityhealth.hhs.gov/templates/content.aspx?ID=6173.
5. Wennergren, G., Strannegard I, Asthma hospitalizations continue to decrease in schoolchildren but hospitalization rates for wheezing illnesses remain high in young children. Acta Paediatr, 2002. 91(11): p. 1239-1245.
6. Wisnivesky, J., Lorenzo J, Lyn-Cook R, et al., Barriers to adherence to asthma management guidelines among inner-city primary care providers. Annals of Allergy, Asthma & Immunology, 2008. 101(3): p. 264-270.
7. DiSantostefano, R., Davis K, Yancey S, Crim C, Ecologic analysis of asthma-related events and dispensing of inhaled corticosteroid- and salmeterol-containing products. Ann Allergy Asthma Immunol, 2008. 100(6): p. 558-565.
8. Crocker, D., Kinyota S, et al., Effectiveness of home-based, multi-trigger, multicomponent interventions with an environmental focus for reducing asthma morbidity: a community guide systematic review. Am J Prev Med, 2011. 41(2): p. S5-S32.
9. Gustafson, D., Wise M, et al., The effects of combining Web-based eHealth with telephone nurse case management for pediatric asthma control: a randomized controlled trial. J Med Internet Res., 2012. 14(4): p. e101.
10. Program, T.N.A.E.a.P., Expert Panel Report 3: Guidelines for the Diagnosis and Management of Asthma. J Allergy Clin Immunol, 2007. 120(5 Suppl): p. S94-138.
11. Celano, M., Holsey CN, Kobrynski LJ, Home-based family intervention for low-income children with asthma: a randomized controlled pilot study. J Fam Psychol., 2012. 26(2): p. 171-178.
12. Zuniga, G., Kirk S, et al., The impact of asthma health education for parents of children attending head start centers. J Community Health, 2012. 37(6): p. 1296-1300.
13. Blaakman, S., Tremblay PJ, Halterman JS, Fagnano M, Borrelli B, Implementation of a community-based secondhand smoke reduction intervention for caregivers of urban children with asthma: process evaluation, successes and challenges. Health Educ Res, 2013. 28(1): p. 141-152.
14. Martin, M., Catrambone CD, et al., Improving asthma self-efficacy: developing and testing a pilot community-based asthma intervention for African American adults. J Allergy Clin Immunol, 2009. 123(1): p. 153-159.
15. Seid, M., D´Amico E, Varni JW, et al., The In Vivo Adherence Intervention For at Risk Adolescents With Asthma: Report of a Randomized Pilot Trial. Journal of Pediatric Psychology, 2012. 37(4): p. 390-403.
16. Edwards, J., INSPIRE curriculum delivered in a faith-based setting. Fam Community Health., 2010. 33(2): p. 117-122.PRINCIPAL INVESTIGATOR
17. Press, V., Pappalardo AA, et al., Interventions to improve outcomes for minority adults with asthma: a systematic review. J Gen Intern Med, 2012. 27(8): p. 1001-1015.
18. DeJongh, T., Gurol-Urganci I, Vodopivec-Jamsek V, Car J, Atun R, Mobile phone messaging for facilitating self-management of long-term illnesses. Cochrane Database Syst. Rev., 2012.
19. James, T., Fine M, Monitoring Asthma Control Using Claims Data And Patient-Reported Outcomes Measures. P.T., 2008. 33(8): p. 454-466.

20. Bender, B., Overcoming barriers to nonadherence in asthma treatment. J Allergy Clin Immunol, 2002. 109(6): p. 554-559.
21. Okelo, S., Eakin M, et al., The Pediatric Asthma Control and Communication Instrument asthma questionnaire: for use in diverse children of all ages. J Allergy Clin Immunol, 2013. 132(1): p. 55-62.
22. Benavides, S., Rodriquez JC, Maniscalco-Feichtl M, Pharmacist involvement in improving asthma outcomes in various healthcare settings: 1997 to present. Ann Pharmacother. , 2009. 43(1): p. 85-97.
23. Drotar, D., Physician behavior in the care of pediatric chronic illness: association with health outcomes and treatment adherence. J Dev Behav Pediatr., 2009. 30(33): p. 246-254.
24. Weinstein, A., The potential of asthma adherence management to enhance asthma guidelines. Ann Allergy Asthma Immunol, 2011. 106(4): p. 283-291.
25. Agency, E.P. President's Task Force on Environmental Health Risks and Safety Risks to Children:Coordinated Federal Action Plan to Reduce Racial and Ethnic Asthma Disparities. 2012 [cited 2013 August 28]; Available from: http://www.epa.gov/childrenstaskforce/actionplan_one_pager.pdf.
26. Fisher, E., Strunk RC, et al., A Randomized Controlled Evaluation of the Effect of Community Health Workers on Hospitalization for Asthma. Arch Pediatr Adolesc Med, 2009. 163(3): p. 225-232.
27. Cazzola, M., Matera M, Safety of long-acting beta2-agonists in the treatment of asthma. Ther Adv Respir, 2007. 1(1): p. 35-46.
28. Halterman, J., Fagnano M, et al., The school-based preventive asthma care trial: results of a pilot study. J Pediatr, 2012. 28(1): p. 1109-1115.
29. Association, A.L. Trends in Asthma Mortbidity and Mortality. 2012.
30. Cloutier, M., Hall C, Wakefield D, Bailit H, Use of asthma guidelines by primary care providers to reduce hospitalizations and emergency department visits in poor, minority, urban children. J Pediatr, 2005. 146(5): p. 591-597.
31. Prevention, C.f.D.C.a., Vital Signs: Asthma Prevalence, Disease Characteristics, and Self-Management Education -- United States, 2001--2009. 2011: p. 547-552.
32. Clinical Guidelines for the Diagnosis, Evaluation and Management of Adults and Children with Asthma, in New York State Department of Health and Mental Hygiene. 2008. Lawrence C. Kleinman, MD, MPH, FAAP
33. Available from: www.childhealthdata.org.
34. Diette, G.B., et al., Consistency of care with national guidelines for children with asthma in managed care. The Journal of Pediatrics, 2001. 138(1): p. 59-64.
35. Adams, R.J., et al., Impact of Inhaled Antiinflammatory Therapy on Hospitalization and Emergency Department Visits for Children With Asthma. Pediatrics, 2001. 107(4): p. 706-711.
36. Finkelstein, J.A., et al., Underuse of controller medications among medicaid-insured children with asthma. Archives of Pediatrics & Adolescent Medicine, 2002. 156(6): p. 562-567.
37. Finkelstein, J.A., et al., Self-Reported Physician Practices for Children With Asthma: Are National Guidelines Followed? Pediatrics, 2000. 106(Supplement 3): p. 886-896.
38. Bell, L.M., et al., Electronic Health Record–Based Decision Support to Improve Asthma Care: A Cluster-Randomized Trial. Pediatrics, 2010. 125(4): p. e770-e777.
39. Lob, S.H., et al., Promoting Best-Care Practices in Childhood Asthma: Quality Improvement in Community Health Centers. Pediatrics, 2011. 128(1): p. 20-28.
40. Scott, L., et al., Achieving and maintaining asthma control in inner-city children. Journal of Allergy and Clinical Immunology, 2011. 128(1): p. 56-63.
41. Gandhi, P., et al., Exploring factors influencing asthma control and asthma-specific health-related quality of life among children. Respiratory Research, 2013. 14(1): p. 1-10.

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 (4b) under Usability and Use.*

The overall rate of ED visits for asthma in NY State Medicaid Managed Care in 2012 is 20.65 per 100 child-years. The tables in the testing form break this down by age and race. The appendix contains additional data from the prior year including demonstrating expected seasonal variations in rate. Given our findings and our methods, although we consider this measure to be specified for a year we have demonstrated its validity to identify or compare asthma ED rate on a month-by-month basis.

The Appendix includes more data as indicated:

Page 2 Table 1. Month by Month Data, Stratified. New York State Medicaid Managed Care, 2012
Figure 2. Asthma ED Visits By Age and Month.
Page 3 Figure 3. ED Visits per 100 Child-years by Age and Urbanicity
Figure 4. ED Visits per 100 child-years by Age and County Poverty Quartile
Page 4 Figure 5. ED Visits per 100 Child Years by Age and Race/Ethnicity

Page 5 Table 3. ED Visits per 100 Child-years by Age and Quartile of Poverty
Table 4. ED Visits per 100 Child-years by Age and Urbanicity
Table 5. ED Visits per 100 Child-years by Age and Quartile of Poverty

Furthermore, within the NY State Medicaid data, a Poisson Regression Analysis and a ZIP analysis demonstrated significant differences by health plan, while controlling for Black race or Hispanic Ethnicity and Age group.

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.**

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 (4b) under Usability and Use.*
On a yearly and a monthly basis we can demonstrate differences in the data by age, urbanicity, race/ethnicity, and level of poverty. Such differences are also evident in other cross tabulations, for example, the rate for children 2-4 in large metropolitan areas is 52.6 visits per 100 child-years compared to those in small metropolitan areas with 26.2, in micropolitan areas with 18.3 and in rural areas with 12.3. Similar magnitudes of differences were seen in other age groups, although the patterns were not all identical. Racial and ethnic differences were notable: for children ages 2-4, the rate in non-Hispanic Whites was 18.4 visits per 100 child-years, in Asians 19.3, in Hispanics 53.9 and in non-Hispanic Blacks 74. Although less dramatic, similar patterns were observed in all age groups. Overall, the rate for different races ordered by varying magnitude as illustrated between Black and White children, 41.99 and 14.79, respectively. The rate for Hispanic children was intermediate at 31.91 visits per 100 child-years. Charts and graphs are shown in the Appendix Tables 2-5 and Figures 2-5. Other disparities data has been cited elsewhere in terms of asthma control and outcomes.

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**

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):*

De.6. **Cross Cutting Areas** *(check all the areas that apply):*
«crosscutting_area»

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

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.)*

n/a

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:** Asthma_1_12_06_16-636167097129812000.xlsx

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

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

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).
The numerator estimates the number of emergency department (ED) visits for asthma among children being managed for asthma. Since most hospitalizations for asthma are from the ED and many ED visits that result in hospitalization are not captured in encounter data, a numerator event may be identified either as an ED visit or as a hospitalization.

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).
Numerator elements include the date and count of all emergency visits or hospitalizations with a primary or secondary diagnosis of asthma in a child who was eligible in the month being assessed. ED visits and hospitalizations should be identified as a visit that is associated with codes found in S.2b for identifiable asthma.

An ED visit that results in hospitalization must be counted as a single numerator event. In other words, for each individual in the denominator for the specified month, consider evidence of hospitalization that is on the same day or one day after an ED visit to represent one discrete event. Consecutive days of hospitalization are considered to represent one hospitalization.

S.6. **Denominator Statement** *(Brief, narrative description of the target population being measured)*
The denominator represents the person time experience among eligible children with identifiable asthma. Assessment of eligibility is determined for each child monthly. The total number of child months experienced is summed and divided by 1200 to achieve the units of 100 child years for the denominator.

Assessing eligibility for the denominator requires 2 years of data, the reporting year and the 12 month period before the reporting year. (See Appendix 1, Figure 1)

S.7. **Denominator Details** *(All information required to identify and calculate the target population/denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b.)*
IF an OUTCOME MEASURE, describe how the target population is identified. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

The denominator is the sum total of the number of months that children meet all eligibility criteria divided by 1200. This calculation yields the denominator in terms of '100 child years', which is the equivalent of 100 children with identifiable asthma in the plan for 1 year each.

We consider children to be managed for identifiable asthma to meet two criteria simultaneously:
1) They have been enrolled for three consecutive months including the month being assessed, and
2) They have evidence of claims sufficient to meet the eligibility criteria for identifiable asthma.

The analysis should be conducted on a month by month basis as described herein: Within the group of children who meet the criteria for identifiable asthma, identify and maintain a unique patient identifier, age, and all stratification variables. We call the time frame during which eligibility is established to be the Assessment Period.

For each month of the Reporting Year, determine eligibility for each patient, as of the last day of the month prior to the reporting month. This illustration assumes that the Reporting Year is 2011. When assessing January 2011, consider all of Calendar Year 2010 as the Assessment Period for assessing the presence or absences of identifiable asthma. For February, 2011 the Assessment Period includes all of calendar year 2010 AND January 2011. Repeat this progression monthly so that for December, 2011 identifiable asthma one would identify children with identifiable asthma using an Assessment Period from January 2010 through November 2011. For each month, assess whether the continuous enrollment criterion is met prior to including the month in the denominator. For example, for January 2011, the child must have been enrolled in November and December, 2010 (plus January 2011). Another example, for December 2011, to be eligible the child must have been enrolled in October 2011 and November 2011, as well as December.

Please see Appendix: Figure 1 and codes used for definitions (Sb.2). These are considered INTEGRAL to these specifications and are not optional.

Identifiable asthma is present when there is evidence as specified for any of the following:
a. Prior hospitalization with asthma as primary or secondary diagnosis
b. Other qualifying events after the fifth birthday (age is age at occurrence):
  i. One or more prior ambulatory visits with asthma as the primary diagnosis AND a subsequent ED visit in the Reporting Month, OR
  ii. Two or more ambulatory visits with asthma as a diagnosis, OR
  iii. One ambulatory visit with asthma as a diagnosis AND at least one asthma-related prescription, OR
  iv. Two or more ambulatory visits with a diagnosis of bronchitis
c. Other qualifying events, any age:
  i. Three or more ambulatory visits with diagnosis of asthma and/or bronchitis, OR
  ii. Two or more ambulatory visits with a diagnosis of asthma and/or bronchitis AND one or more asthma- related prescriptions.

As noted in the specifications, asthma-related medicine means long-acting beta-agonist (alone or in combination) or inhaled corticosteroid (alone or in combination), anti-asthmatic combinations, methylxanthines (alone or in combination), and/or mast cell stabilizers.

Please note that in order to promote better harmonization, we start with the current HEDIS asthma medication list. From that list, in accordance with our expert panel recommendations we eliminate medications in the following 2 categories: leukotriene modifiers, short-acting inhaled beta-agonists.

We further exclude indacaterol, a recently approved long acting beta agonist that is indicated in the US only for the treatment of COPD. As indicated elsewhere, COPD is an exclusion criterion for this measure. These specifications anticipate that NCQA will update the medication list from time to time and with the stated exclusions updated lists may be substituted for the list linked herein. The table used for testing is labeled Table AMR-A: Asthma Controller and Reliever Medications, and can be found at http://www.ncqa.org/HEDISQualityMeasurement/HEDISMeasures/HEDIS2015/HEDIS2015NDCLicense/HEDIS2015FinalNDCLists.aspx (last accessed September 12, 2015).

If pharmacy data are not available, the measure should be reported with notation that pharmacy data were not used for the assessment of eligibility. This avoids eliminating from the measure those facilities with no link to pharmacies. Our testing reveals that only a small proportion of patients are excluded by not including pharmacy data to establish eligibility.

The presence of identifiable asthma (see S.2b and above) is established each month from administrative data using the specified algorithm. (Appendix: Figure 1 and this section's narrative)

All events in the administrative data should be associated with a date of service.

S.8. **Denominator Exclusions** *(Brief narrative description of exclusions from the target population)*
Children with specified concurrent or pre-existing diagnosis and children who have not been consecutively enrolled in the reporting plan for at least three months, as specified in the details section.

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.)*
Children with concurrent or pre-existing: Chronic Obstructive Pulmonary Disease (COPD) diagnosis, Cystic Fibrosis diagnosis, or Emphysema diagnosis.

Children who have not been consecutively enrolled in the reporting plan for at least two months prior to the index reporting month and for the reporting month (a total of three consecutive months ending in the reporting month).

For entities that use AHRQ's Clinical Classifications Software, please note that it is important to apply the exclusion after identifying visits that satisfy CCS class 128.

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.)*
This measure calls for stratification by age group, by race/ethnicity, and by age group and race/ethnicity. Several additional stratifications are recommended but optional. These may be required by the accountability entity or reported by the reporting entity. These variables include rurality/urbanicity and county level of poverty.

Age groups are 2-5, 6-11, 12-18, and 19-20, each inclusive. (reporting entity should specify whether to use age at month of qualifying event or age on first day of reporting year)

Race/ethnicity should incude White non-Hispanic, Black non-Hispanic, and Hispanic as well as other groups as requested by the accountability entity and consistent with current HHS usage.

For social demographic stratification: identify County equivalent of child's residence. If County and State or FIPS code are not in the administrative data, the zip codes can be linked to County indirectly, using the Missouri Census Data Center (http://mcdc.missouri.edu/). These data will link to County or County equivalents as used in various states.

i.Identify the Urban Influence Code (1) or UIC for the county of child's residence. (2013 urban influence codes available at: http://www.ers.usda.gov/data-products/urban-influence- codes.aspx#.UZUvG2cVoj8).

ii.Identify the Level of Poverty in the child's county of residence. The percent of all residents in poverty by county or county equivalent are available from the US Department of Agriculture at http://www.ers.usda.gov/data-products/county-level-data-sets/download- data.aspx. Our stratification standards are based on 2011 US population data that we have analyzed with SAS 9.3. Using child's state and county of residence (or equivalent) or FIPS code, use the variable PCTPOVALL_2011 to categorize into one of 5 Strata:

a.Lowest Quartile of Poverty if percent in poverty is <=12.5%
b.Second Quartile of Poverty if percent in poverty is >12.5% and <=16.5%
c.Third Quartile of poverty if percent in poverty is >16.5% and <=20.7%
d.First Upper Quartile (75th-90th) if percent in poverty is >20.7% and <=25.7%
e.Second Upper Quartile (>90th percentile)

iii.Categorize age by age at the last day of the month that ends the assessment period. Aggregate into age categories 2-4, ages 5 through 11, ages 12-18, ages 19-21.

iv.Categorize Race/Ethnicity as Hispanic, Non-Hispanic White, Non- Hispanic Black, Non-Hispanic Asian/Pacific Islander, and Non-Hispanic Other

v.Categorize Insurance Type as Private (Commercial), Public, None or Other

vi.Categorize benefit type as HMO, PPO, FFS, PCCM, or Other

S.11. **Risk Adjustment Type** (Select type. Provide specifications for risk stratification in measure testing attachment)
Other
If other: In order to allow for more granular comparisons this measure is specified to be stratified. Stratification for risk adjustment of this measure would not be justified by the literature. Although epidemiological findings support our stratification schema, no biological evidence exists to support intrinsic correlation of ED rates with stratification variables.

S.12. **Type of score:**
Rate/proportion
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 = Lower score

S.14. **Calculation Algorithm/Measure Logic** (*Diagram or describe the calculation of the measure score as an ordered sequence of steps including identifying the target population; exclusions; cases meeting the target process, condition, event, or outcome; time period for data, aggregating data; risk adjustment; etc.)*
Step 1: Measure person-time eligible for each patient and record by month.
a. For each month in the reporting year, identify all children ages 2 – 21 years who meet the criteria for Identifiable asthma during the assessment period. The assessment period is defined as the year prior to the reporting year plus all months in the reporting year prior to the reporting month.

Identify and maintain a unique patient identifier and all stratification variables.

To illustrate: if the goal is to report for January 2011, first one would identify children with Identifiable asthma using the criteria, and analyze all of calendar year 2010 when doing so. Continuous enrollment criterion requires that the child was enrolled in November and December of 2010, as well as January 2011. This total represents the number of person-months (child-months) for January.

Next, for February: one would identify children with Identifiable asthma using the criteria, and analyze all of calendar year 2010 AND January 2011 when doing so. Continuous enrollment criterion requires that the child was enrolled in December 2010 and January 2011, as well as February 2011. This is the number of person-months (child-months) for February.

Repeat this progression monthly so that for December, one would identify children with Identifiable asthma and analyze all of calendar year 2010 AND January through November 2011 when doing so. Continuous enrollment criterion requires that the child was enrolled in October 2011 and November 2011, as well as December 2011. This is the number of person-months (child-months) for December.

b. Sum all months that are eligible from the reporting year. This sum is the denominator in people-months. Divide by 1200. This is denominator in 100 people-years. This is the denominator for the year.

Step 2: Month by month, considering the definitions above, identify the number of discrete numerator events that occur in children eligible in that specific month:
a.Identify the number and date of ED visits with asthma as a primary or secondary diagnosis among those children who are eligible for that reporting month.
b.Identify the number and date of inpatient hospitalizations with asthma as a primary or secondary diagnosis among those children who are eligible for that reporting month.
c.Identify the number of discrete numerator events. Consecutive days with inpatient hospital codes are considered one hospitalization. Hospitalizations on day of or day after ED visit are NOT considered discrete from the ED visit.
d.Sum the number of numerator events across the year.

e.Maintain stratification variables and unique identifiers.

Step 3. Calculate rate as Numerator / Denominator. While this measure is specified for the year, it has also been validated to demonstrate seasonality using monthly rates.

Step 4. Calculate stratification variables as specified in S.12.

Step 5. Repeat by strata. Within age strata repeat by other specified strata. Perform other cross tabulations as requested by the accountability entity. Eliminate any strata with less than 40 person-months in any month's denominator OR less than 1000 person-months for the year.

Appendix 1: Figure 1 illustrates the calculation of person-time and is considered fundamental to this calculation algorithm.

When data cannot be obtained from any source:
- If critical for calculation – delete patient from consideration for that reporting month
- If non-critical for calculation – include patient
Critical data include encounter data for the reporting month and some period of time in the assessment period. In order to report stratifications age and race/ethnicity are considered critical. Pharmacy data are not considered critical

S.15. **Sampling** *(If measure is based on a sample, provide instructions for obtaining the sample and guidance on minimum sample size.)*
IF a PRO-PM, identify whether (and how) proxy responses are allowed.
n/a

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.)*
IF a PRO-PM, specify calculation of response rates to be reported with performance measure results.
n/a

S.17. **Data Source** *(Check ONLY the sources for which the measure is SPECIFIED AND TESTED).*
If other, please describe in S.18.
Claims (Only), Claims (Other)

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 is collected.)*
IF a PRO-PM, identify the specific PROM(s); and standard methods, modes, and languages of administration.
n/a

S.19. **Data Source or Collection Instrument** *(available at measure-specific Web page URL identified in S.1 OR in attached appendix at A.1)*
No data collection instrument provided

S.20. **Level of Analysis** *(Check ONLY the levels of analysis for which the measure is SPECIFIED AND TESTED)*
Health Plan, Population : Community, County or City, Population : Regional and State

S.21. **Care Setting** *(Check ONLY the settings for which the measure is SPECIFIED AND TESTED)*
Emergency Department, Hospital, Hospital : Acute Care Facility, Other
If other: This measure incorporates data from the ambulatory, ED, and hospital settings to describe performance at the level of the plan or the community.

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.)*
n/a

2. Validity – See attached Measure Testing Submission Form
nqf_testing_attachment_Asthma_1_12_07_16_lk_v2.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. (Do not remove prior testing information – include date of new information in red.)

2.2 <u>For maintenance of endorsement</u>
Has additional empirical validity testing of the measure score been conducted? If yes, please provide results in the Testing attachment. (Do not remove prior testing information – include date of new information in red.)

2.3 <u>For maintenance of endorsement</u>
Risk adjustment: For outcome, resource use, cost, and some process measures, risk-adjustment that includes SDS factors is no longer prohibited during the SDS Trial Period (2015-2016). Please update sections 1.8, 2a2, 2b2, 2b4, and 2b6 in the Testing attachment and S.14 and S.15 in the online submission form in accordance with the requirements for the SDS Trial Period. NOTE: These sections must be updated even if SDS factors are not included in the risk-adjustment strategy. If yes, and your testing attachment does not have the additional questions for the SDS Trial please add these questions to your testing attachment:

What were the patient-level sociodemographic (SDS) variables that were available and analyzed in the data or sample used? For example, patient-reported data (e.g., income, education, language), proxy variables when SDS data are not collected from each patient (e.g. census tract), or patient community characteristics (e.g. percent vacant housing, crime rate).

Describe the conceptual/clinical and statistical methods and criteria used to select patient factors (clinical factors or sociodemographic 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)

What were the statistical results of the analyses used to select risk factors?

Describe the analyses and interpretation resulting in the decision to select SDS 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)
No - This measure is not risk-adjusted

**Note: New information added since the previous review is in blue text

NATIONAL QUALITY FORUM—Measure Testing (subcriteria 2a2, 2b2-2b7)

Measure Number (*if previously endorsed*)**:** Click here to enter NQF number

Measure Title: Rate of Emergency Department Visit Use for Children Managed for Identifiable Asthma: Visits per 100 Child-years

Date of Submission: 12/14/2016

Type of Measure:

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

Instructions

- 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.
- **For all measures, sections 1, 2a2, 2b2, 2b3, and 2b5 must be completed.**
- **For outcome and resource use measures**, section **2b4** also must be completed.
- If specified for **multiple data sources/sets of specificaitons** (e.g., claims and EHRs), section **2b6** also must be completed.
- 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 (2b2-2b6) must be in this form. An appendix for *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.
- Maximum of 20 pages (*incuding questions/instructions;* minimum font size 11 pt; do not change margins). ***Contact NQF staff if more pages are needed.***
- Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.
- For information on the most updated guidance on how to address sociodemographic variables and testing in this form refer to the release notes for version 6.6 of the Measure Testing Attachment.

Note: The information provided in this form is intended to aid the Steering 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 **PRO-PMs and composite performance measures**, reliability should be demonstrated for the computed performance score.

2b2. 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 **PRO-PMs and composite performance measures**, validity should be demonstrated for the computed performance score.

2b3. Exclusions are supported by the clinical evidence; otherwise, they are supported by evidence of sufficient frequency of occurrence so that results are distorted without the exclusion; [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]

2b4. 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 sociodemographic 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.

2b5. 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.

2b6. If multiple data sources/methods are specified, there is demonstration they produce comparable results.

2b7. For **eMeasures, composites, and PRO-PMs** (or other measures susceptible to missing data), 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.
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.23***)	**Measure Tested with Data From:**
☐ abstracted from paper record	☐ abstracted from paper record
☒ administrative claims	☒ administrative claims
☐ clinical database/registry	☐ clinical database/registry
☐ abstracted from electronic health record	☐ abstracted from electronic health record
☐ eMeasure (HQMF) implemented in EHRs	☐ eMeasure (HQMF) implemented in EHRs
☐ other: Click here to describe	☐ other: Click here to describe

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*).

New York State Medicaid claims data.

Also, our work builds off of work performed by our CAPQuaM partner and steering committee member, NCQA. For specific data reliability and signal to noise analyses, we incorporate by reference (and will present more selectively) NCQA data relevant to their submission for NQF –endorsed asthma related measures:

- Use of Appropriate Medications for People with Asthma (ASM) – 0036 (we understand this is no longer being maintained as of 2015, but it was endorsed and the data were accepted.)
- Medication Management for People With Asthma (MMA) – 1799
- Asthma Medication Ratio (AMR) – 1800

We note that 1799 and 1800 are not directly applicable because they were tested at the score level. However, the scores were dependent upon definitions which use the same data element level as our measure and thus provide indirect evidence of the capacity of a measure using such data elements to produce valid scores.

The analyses above provide information regarding the capacity to use administrative data to identify the applicable denominator population. There is nearly complete overlap of the denominator codes and there is overlap of the denominator elements. Where codes differ it is specific to decisions made by the CAPQuaM expert panel which was aware of the NCQA measures.

1.3. What are the dates of the data used in testing? 2010 - 2012

1.4. What levels of analysis were tested? (*testing must be provided for <u>all</u> 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.26***)	**Measure Tested at Level of:**
☐ individual clinician	☐ individual clinician
☐ group/practice	☐ group/practice
☐ hospital/facility/agency	☐ hospital/facility/agency
☒ health plan	☒ health plan
☒ other: Population, State, Region, County, Integrated delivery system	☒ other: Population, State, Region, County, Integrated delivery system

1.5. How many and which <u>measured entities</u> 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*)

NY State Medicaid Managed Care claims data, including claims from all MCO's that are contracted for Medicaid care by our partner, the NY State Department of Health.

For our primary analysis of MCO's we included both full (8) and partially (10) capitated plans, each of which had at least 900 children who contributed time to the denominator of the measure.

For our primary analysis of county of residence, we used 45 counties that contributed at least 1000 months of person time to the denominator.

The numbers we present are from reporting year 2012, include children from counties in nine urban influence codes and in counties poverty level 1-3. NY State does not have any counties in the lowest 25% of poverty or with UIC of 10-12. New York has more than 60 counties and numerous health plan vendors. Analysis in Year 2011 provided very similar data.

Foundational analyses for this measure were performed and previously reported by NCQA considering *nine health plans covering a variety of geographic areas within the United States that were asked to provide a complete administrative data file consisting of any member in their commercial and Medicaid product lines for anyone that had a diagnosis code for asthma during the calendar years of 2009-2010. The complete member-level administrative file used for analysis included a total of more than 82,000 health plan members with asthma*.

1.6. How many and which <u>patients</u> 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*)

All children 0-21 with records in the 2011 in the 2011 NYS Medicaid Managed Care (MMC) administrative database and all pediatric patients meeting the criteria for identifiable asthma in the 2012 NYS Medicaid Managed Care administrative database.

There were 192,722 children with identifiable asthma in the managed care (plan level) analysis, 211,703 in the county analysis, and 212,432 overall in MMC.

19,903 children experienced 30,382 qualifying emergency department visits in the reporting year. 1806 visits were in young adults age 19 or 20. In 2011, the median number of visits per child with an ED visit was 1, the 75th percentile was 2, and the 90th percentile was 3. One percent had 6 or more visits.

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.

Data source 1 (Chart):
Prior to initial specification of the measure we contracted for a survey of quality managers representing more than a dozen hospitals to assess data availability and the ease and feasibility of abstraction of data relevant for asthma and other CAPQuaM measure, including 10 hospitals that responded regarding asthma-specific data elements. Our survey found that availability in the medical record of age, (date of birth), race, ethnicity, date and site of visit, documentation of primary or secondary diagnosis of asthma, hospitalization, and payment source were routinely available in the chart and "Not Difficult to Collect." Our chart review of 1200 medical records for ED visits in a single institution performed for validation of a sister measure on appropriateness confirmed the availability of these data. This validates the capacity to obtain such data form the medical record and the primary occurrence of the data in the chart so that coders have the clinical information required to population ICD-9/10, CPT, and Revenue codes that comprise the administrative data that are the preferred data source for these measures.

Data Source 2 (Administrative):
Assessment of the capacity to identify the eligible population and qualifying events was performed in NY State Medicaid data in both 2011 and 2012 reporting years.

For MCO analysis we analyzed both with the 18 plans that had 900 or more children contributing to the denominator and with the 20 plans that contributed at least 1000 months of person time to the denominator and found no meaningful differences in the analyses. We present details from the analysis of the 18 plans.

For county level analysis we included those 45 counties that contributed at least 1000 months of person time to the denominator.

Other analyses included all children with identifiable asthma.

Data source 3 (explicit criteria):
Our construct for the CAPQuaM measure was defined by the multidisciplinary national expert panel using a RAND type modified Delphi process, which produced a set of explicit criteria that were both substantive and addressed specification details, such as what combination of administrative codes could be used to identify a child with asthma, other inclusion and exclusion criteria for the measure, and preferences regarding how to report and stratify the measure.

The panel initially used the term persistent asthma to describe asthma that was pre-existing and should have been recognized as asthma by the health care system prior to the timing of the ED visit. This construct was renamed by our stakeholder group to be identifiable asthma to avoid confusion with other uses of the term persistent asthma. The construct was intended to be more inclusive than HEDIS' persistent asthma diagnosis, while still removing from consideration those whose asthma was unlikely to have been actively managed at the time.

Data Source 4 (National Survey of Children's Health)
We validate the construct of identifiable asthma comparing it to two other constructs:
HEDIS' definition of persistent asthma, which should have been more restrictive than 'identifiable asthma'; and the National Survey of Children's Health's question regarding if the caregiver had ever been told by a doctor or nurse that the child had asthma, which should have been less restrictive than 'identifiable asthma.' The former analysis was conducted in Medicaid 2011 and the latter in the most recent NSCH data.

Holding steady the continuous enrollment criterion at 12 months, HEDIS criteria identified a rate of persistent asthma of 3.1% with the CAPQuaM criteria identifying identifiable asthma at a rate of 8.6%. As expected, identifiable asthma was between 2 and 3 times more permissive than the intentionally restrictive persistent asthma. We analyzed NSCH data to estimate a population rate of asthma in NY State Medicaid child population to be between 15 - 16%, indicating that our criteria did provide a meaningful filter as we had intended.

Reducing the continuous enrollment period down to three months as was suggested by members of our steering committee increases the number of children eligible for the measure by several tens of thousands while still restricting the measure to those who had received sufficient care for asthma to be identified, and requiring continuous enrollment for attribution to the extent felt important by our multi-stakeholder group. This inclusiveness help to counter risks of churning that are particularly prominent in the Medicaid population. This analysis was conducted in the NYS Medicaid data.

Data Source 5: HEDIS
Assessment of data elements for identifying a population with asthma and asthma scores was performed by NCQA in nine geographically diverse managed care plans. We considered the HEDIS data for measures 1799, 1800 and 0036. We cite 1799 and 1800 not as specific evidence of score level performance of our measure, but as evidence that measures that rely on the same administrative data elements for their denominator have the capacity to distinguish signal from noise at a very high level.

1.8 What were the patient-level sociodemographic (SDS) variables that were available and analyzed in the data or sample used? For example, patient-reported data (e.g., income, education, language), proxy variables when SDS data are not collected from each patient (e.g. census tract), or patient community characteristics (e.g. percent vacant housing, crime rate).

Race, ethnicity, zip code/county of residence, level of poverty in the county of caregiver residence, and urban influence code for the county of caregiver residence for the NY State analysis. Within the Medicaid data, we looked at eligibility category.

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*)

See section 2b2 for validity testing of data elements.

See 2b5 for testing that demonstrates sufficient precision and reliability of the performance measure score. Please note as this is a rate the binomial signal to noise ratio analysis is not appropriate for this measure. Similarly since measured variation between entities is expected sometimes to be meaningful and other times not, the key performance attribute is to demonstrate the capacity in real use to identify which of multiple comparisons are statistically significant. We discuss our findings in this context in 2b5.2.

In addition to the analyses presented below, we also conducted the analysis using a zero inflated poisson approach.

With NY State Medicaid we conducted analyses that demonstrate the measure's capacity to distinguish among health plans. The standard approach to measuring reliability is inappropriate as the measure is a rate and not a binomial. The appropriate model is either a Poisson, model (which is discussed in 2b5), a hurdle model or a Zero inflated Poisson (ZIP). Hurdle requires additional assumptions that model two processes, and is more sensitive. ZIP misses out on capturing some of the plans' impact on whether a child makes it to the ER, but models the rate very well. We performed both with similar results and report on the ZIP as the more conservative approach (it under attributes the impact of the plan).

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*)

ZIP Models: Using Proc HPFMM with a log link, a Poisson distribution and an offset equal to the log of the number of months the child had asthma in the plan, the model was highly significant (p<.0001) incorporating specified age groups and plans as categorical variables. Comparing to a randomly selected index plan, 14 of 17 plans had statistically significant differences in performance with the median and modal p-value being <0.001. Non-significant plans' p-values=0.08, 0.16 and 0.88.The model is able to differentiate distinct performance levels. Results were similar when we performed the models considering only plans, after stratifying for age group. Because of low numbers in the 18-21 yr old group across plans, fewer were significant, but findings suggest that the measure is sensitive to real differences given adequate sample sizes.

Ages 2-4: 15 plans of 17 are significant (p<0.05). Additional are 0.06 and 0.21.
Ages 5-11: 14 plans of 17 are significant (p<0.05). Additional are 0.37, 0.21, and 0.70.
Ages 12-18: 13 plans of 17 are significant (p<0.05). Additional are 0.11, 0.06, 0.26, and 0.43.
Ages 19-21: 7 that were significant (p<0.01). In general the sample size was sufficient to assess some plan's performance for this group.

ZIP models also showed that even after controlling for age groups: Urban counties have different performance than rural counties; Large urban counties are distinct in performance from all others; Small urban counties are different from suburban counties and rural counties, although the smaller numbers in rural counties contributes to a P-value of 0.07; Performance in suburban and rural counties are generally similar. New York State does not have extremely rural counties; ED utilization of Blacks is significantly different from Whites (p<0.01); ED utilization of Hispanics is significantly different from Whites (p<0.01); ED utilization of Blacks and Hispanics are significantly different from one another (p<0.01).
These data contribute evidence to support use of the measure, adding both to the data on reliability (as plan to plan differences were meaningful) and validity (in that the models performed as predicted and consistent with current knowledge regarding variations associated with race, ethnicity, and urbanicity).

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?*)

The ZIP models reinforce what is described below and add robustness to our interpretation that the performance measure scores demonstrate reliability with a high degree of certainty and confidence.

_____

2b2. VALIDITY TESTING

2b2.1. 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*)

2b2.2. 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)*

Analysis of 2011 and 2012 data provided similar findings, including the pattern of month to month variations and the variability associated with various stratification variables, demonstrating test-retest reliability. Further the identification of predicted seasonal changes within a defined population is a more difficult challenge for distinction (or signal to noise) than comparison in distinct populations and confirms in the analysis theoretically sound and predicted differences among groups of children. We found evidence of theoretically sound and predicted differences whether we categorized on race (Blacks were highest rate), type of health plan (fully capitated pans had lower rates than partially capitated plans), poverty (higher poverty had higher rates), or urbanicity (large urban areas had the highest rates), and age group (school age had the lowest rates). This leads us to have high confidence that the performance measure scores are a valid and reliable indicator of quality and of the underlying construct of undesirable outcomes in asthma.

Observed differences between counties and between health plans were similar regardless of whether or not we controlled for age group, race, ethnicity in the analysis.

Please see descriptions of both NCQA and CAPQuaM testing above in 1.2-1.7.

The literature also supports the use of claims data to identify the presence of asthma. We use administrative data to identify the age of the child, various stratification variables and the presence of asthma, as well as the presence of an asthma ED visit or hospitalization. These are routinely used to support billing by CMS, Medicaid, and private insurers and are routinely used in quality measurement.

There is moderate agreement (kappa = 0.45 – 0.50) when comparing administrative data regarding the presence of constructs such as recent asthma attacks, use of asthma medications, attack or medication, attack and medication, using 1 year of administrative claims data. The agreement improves from 0.55 to 0.60 when using two years of data as this measure does. (8). We expect that these kappas would be significantly higher were the analyses restricted to children with disease that met our construct criteria for identifiable asthma.

We followed a peer-reviewed systematic process for measure development that incorporated a literature review, an expert panel process, and a multi-stakeholder process that included the input and review of the measure by potential users in the development process. More details about this process are available upon request.

Explicit criteria were developed using a variation of the two-round modified Delphi process RAND/UCLA Appropriateness Method with a multidisciplinary and geographically diverse expert panel comprised of both clinicians and researchers. Identifiable asthma was based on panel findings and appropriateness criteria included for this measure were those that were both available in the chart and highly rated. The general reliability of this approach is well established. [9, 10] It has been applied successfully to pediatric services previously. [11-13] The expert panel further validated the measure subsequent to development via an email poll.

Development included a series of alpha tests to refine specifications by conducting iterative analyses in New York State Medicaid data. Conclusions from alpha tests include:

1) The reporting period and the assessment period could not overlap completely, leading to use of 2 years of data as shown in the specifications' diagram. The optimal approach was to divide the reporting year into 12 reporting months. ED events in that month are eligible for the numerator if persistent asthma criteria have been satisfied (combining the look-back year and all prior months in the reporting year) and the child has been continuously enrolled for the two months immediately prior to the reporting month. The optimal building block unit for the denominator is in child-months, which is rolled up to child-years;

2) Using both revenue codes and CPT codes increased our sensitivity meaningfully, a choice validated by consultation with coding and billing experts and confirmed by analyzing the NY State data;

3) NY State Medicaid data and national survey data (HCUP) converged to demonstrate the importance of including hospitalizations as numerator events even when the underlying construct is ED visits. This is consistent with policies of many payers to request providers not to submit both ED and hospital claims for the same day. Error is far less by considering both ED visits and hospitalizations as numerator events, than by not including hospitalizations. (See poster that follows in the next section)

4) The expert panel only wanted numerator events for which the children were already known to the accountable entity as having asthma and established definitions for such "identifiable asthma". Alpha testing in NY State Medicaid demonstrated the expected results:

    a. Holding steady the continuous enrollment criterion at 12 months, HEDIS criteria identified a rate of persistent asthma of 3.1%, the CAPQuaM criteria identifying identifiable asthma at a rate of 8.6%. This more inclusive approach was our goal.

b. More than 25% of children with any asthma claim are not included in the denominator, indicating that this is a meaningful filter. Confirming this, the observed rate of 8.3% in the denominator. 8.3 is just over half of what we found when analyzing NSCH data to identify an expected rate of NY State Medicaid children whose caregivers would report that they every been told the child had asthma.

c. Relaxing the continuous enrollment period to 3 months was suggested by members of our stakeholder steering committee. Doing so increased the eligible number by more than 20,000 while still restricting the measure to those who had received sufficient care for asthma to be identified, and requiring continuous enrollment for attribution to the extent felt important by our multi-stakeholder group.

The use of Expert Panels has been demonstrated to be useful in measure development and health care evaluation, including for children. [14]

The definitions were specified to allow their use with data elements that ought to be available in electronic form to a responsible entity, such as a health plan or state Medicaid program. Potential exceptions to this are elements such as ZIP code of residence and race and ethnicity of the child. We have data from a feasibility study we conducted with a contractor that surveyed quality departments at more than a dozen hospitals across three measure sets. 10 hospitals responded to the asthma-specific questionnaire. We found that these data elements are generally available in the chart, although the definition of race and ethnicity, as well as how it is determined, may vary by institution. Nonetheless, the CHIPRA legislation (2009), which has funded the development of this measure, directs for measures to be capable of identifying disparities and we have specified it to be so, despite concerns about reliability in the collection and assessment of race and ethnicity by health-care-providing institutions and practices. In this case, we need to drive performance through measurement, as it is foundational to the legislative and executive branch sources of our funding.

former Medicaid core measure that we were tasked with enhancing was a simple risk, with asthma patients defined in the measurement year as having primary or secondary diagnosis for any service, and ED visits defined as CPT-code-identified ED visits with asthma as the primary diagnosis. The numerator for the Core Measure includes all patients with at least one ED visit for asthma as asthmatic events, whether or not the patient was known to be an asthmatic before the event. Further, numerator events alone could qualify children for inclusion in the denominator. Our partners in the New York State Medicaid program have described this characteristic as highly undesirable and the CAPQuaM team agreed, prompting our month–by-month approach to analysis. We enhanced the validity of this measure by deflating competing concepts and clearly specifying it as an interpretable epidemiological rate (incidence density). Enhancements include: we set a threshold of utilization below which a child is not considered to have given the health care system an opportunity to have identified the child as an asthmatic; we restrict the measure to those children who meet this threshold before the ED visit occurs and we are measuring and incidence density or rate and not a risk, allowing us to count each ED visit in the numerator and person-time in the denominator. While the median number of visits among those with visits is 1, more than one-quarter of children in New York State Medicaid Managed Care with an ED visit have a second visit. A few outliers contribute more than 10 Ed visits per child. The rate measure allows us to provide a better estimate of the number of undesirable outcomes, rather than the number of children with undesirable outcomes.

As a rate, one child can contribute to the numerator many times. It also is self-adjusting for children who enter or leave the eligible population since children contribute to the denominator independently for each month that they are eligible. It also assures that ages can be calculated to the month rather than to the year, if the reporting entity requests this level of detail. Our analysis uses the age on the first day of the Reporting Year.

To enhance the meaningfulness of the measure, we have included a two-month continuous enrollment requirement prior to the reporting month. Since the child must also be eligible for the reporting month, this becomes a three-month continuous enrollment requirement. In doing this, we sought to strike a delicate balance between developing a meaningful accountability measure and eliminating children because of problems of churning, which have been well

documented by researchers (15). This balance was achieved in close collaboration with our colleagues at NY State Medicaid.

The development team's goal was to develop an ICD10 code set that was fully consistent with the intent of the original measure. Our process began by performing general equivalency mapping using the forward mapping from www.icd9data.com. We then did a de novo review of the CMS ICD 10 CM set to seek to identify codes that might be appropriate for asthma. We reviewed potential codes identified by both sources and developed a new list of codes appropriate for inclusion criteria and a new list of codes appropriate for exclusion criteria. Drs. Kleinman and Sharma reviewed the lists independently and then achieved consensus in a conference call review and discussion. The guidance for the intended constructs for both ICD9 and ICD10 coding were the findings from a RAND style modified Delphi panel that incorporated 9 national experts over the course of the measure development process.

1. Dresser, M.V., et al., *Clinical quality measurement. Comparing chart review and automated methodologies.* Med Care, 1997. **35**(6): p. 539-52.

2. Newton, K.M., et al., *The use of automated data to identify complications and comorbidities of diabetes: a validation study.* J Clin Epidemiol, 1999. **52**(3): p.

   199-207.

3. Thompson, B.L., et al., *Measuring clinical performance: comparison and validity of telephone survey and administrative data.* Health Serv Res, 2001. **36**(4): p.

   813-25.

4. Angier, H., et al., *Variation in outcomes of quality measurement by data source.*

   Pediatrics, 2014. **133**(6): p. e1676-82.

5. Weiskopf, N.G. and C. Weng, *Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research.* Journal of the American Medical Informatics Association, 2013. **20**(1): p. 144-151.

6. Pawlson, L.G., S.H. Scholle, and A. Powers, *Comparison of administrative-only versus administrative plus chart review data for reporting HEDIS hybrid measures.* Am J Manag Care, 2007. **13**(10): p. 553-8.

7. NCQA. *National Committee for Quality Assurance*. [cited 2014 7/30/14]; Available from: http://www.ncqa.org/

8. Huzel, L, et al. Diagnosing Asthma: The fit between survey and administrative database. Can Respir J. 2002 Nov-Dec;9(6):407-12.

9. Fitch, K., et al., *The RAND/UCLA Appropriateness Method User's Manual*. 2001 RAND.

10. Kosecoff, J., et al., *The appropriateness of using a medical procedure. Is information in the medical record valid?* Med Care, 1987. **25**(3): p. 196-201.

11. Kleinman, L.C., et al., *The medical appropriateness of tympanostomy tubes proposed for children younger than 16 years in the United States.* Jama, 1994. **271**(16): p. 1250-5.

12. Kleinman, L.C., E.A. Boyd, and J.C. Heritage, *Adherence to prescribed explicit criteria during utilization review. An analysis of communications between attending and reviewing physicians.* Jama, 1997. **278**(6): p. 497-501.

13. Keyhani, S., et al., *Overuse of tympanostomy tubes in New York metropolitan area: evidence from five hospital cohort.* Bmj, 2008. **337**: p. a1607.

14. Brook, R.H., et al., *A method for the detailed assessment of the appropriateness of medical technologies.* International journal of technology assessment in health care, 1986. **2**(01): p. 53-63.

15. Fairbrother, G., et al., Churning in Medicaid managed care and its effect on accountability. J Health Care Poor Underserved, 2004. 15(1): p. 30-41.

2b2.3. What were the statistical results from validity testing? (*e.g., correlation; t-test*)

Performance Measure Score:

We found statistically meaningful differences in predicted directions when we used categories such as race/ethnicity, age group, month of year, type of managed care, level of poverty, and urbanicity. These differences were robust to whether we considered them individually or in a common model. The model for testing as a Poisson regression with a log-link function in SAS (Proc GenMod), using the number of ED visits as the outcome and the natural log of the months of exposure as the offset.

When we tested for plan to plan differences findings were similar. If we randomly selected a plan we typically found more than half differing from it with a P<0.05, typically less than 0.01. If we picked extreme plans, virtually all were different. We found similar findings when using county rather than plan. (See section 2b5.2). Statistical differences were robust to inclusion of the age group and race/ethnicity. Stratified by age group, similar findings were found, with somewhat fewer plan-plan and county to county differences found in the older age group, which had smaller numbers of children.

Data Elements and Expert Process

For the foundational NCQA work (Measures 1799, 1800, 0036), NCQA's field test retested a number of previously validated criteria for identifying an eligible population with persistent asthma using administrative claims data. Using the dataset provided, NCQA examined several different scenarios to determine the effects of different specification criteria on this particular population. This information was combined with multiple years of HEDIS data collection of this measure to examine the reliability of collecting this measure through administrative claims.

From NCQA's submissions: *Reliability was estimated on the HEDIS 2011 submissions (2010 data) using the beta-binomial model. Beta-binomial is a better fit when estimating the reliability of simple pass/fail rate measures as is the case with most HEDIS® health plan measures. The beta-binomial model assumes the plan score is a binomial random variable conditional on the plan´s true value that comes from the beta distribution. The beta distribution is usually defined by two parameters, alpha and beta. Alpha and beta can be thought of as intermediate calculations to get to the needed variance estimates. The beta distribution can be symmetric, skewed or even U-shaped.*

Reliability used here is the ratio of signal to noise. The signal in this case is the proportion of the variability in measured performance that can be explained by real differences in performance. A reliability of zero implies that all the variability in a measure is attributable to measurement error. A reliability of one implies that all the variability is attributable to real differences in performance. The higher the reliability score, the greater is the confidence with which one can distinguish the performance of one plan from another. A reliability score greater than or equal to 0.7 is considered very good.

For the foundational NCQA work, NCQA's field test retested a number of previously validated criteria for identifying an eligible population with persistent asthma using administrative claims data. Using the dataset provided, NCQA examined several different scenarios to determine the effects of different specification criteria on this particular population. This information was combined with multiple years of HEDIS data collection of this measure to examine the reliability of collecting this measure through administrative claims. They report that score level reliability of the HEDIS 2011 submissions (2010 data) was assessed using the beta-binomial model. Beta-binomial is a better fit when estimating the reliability of simple pass/fail rate measures as is the case with most HEDIS® health plan measures.

For the NCQA analysis reliability was reported as the ratio of signal to noise. The higher the reliability score, the greater is the confidence with which one can distinguish the performance of one plan from another. A reliability score greater than or equal to 0.7 is considered very good.

We note that 1799 and 1800 are not directly applicable because they were tested at the score level. However, the scores were dependent upon definitions which use the same data element level as our measure and thus provide indirect evidence of the capacity of a measure using such data elements to produce valid scores.

Thus we cite them not as specific evidence of our score level performance of the submitted measure, but as evidence that the HEDIS measures that rely on the same administrative data elements for their denominator have the capacity to distinguish signal to noise at a very high level. While the evidence is indirect it is dispositive. That is, we assert that had the data elements been inadequate it would result in non-differential misclassification error which is a major bias towards the null thus introducing noise and reducing signal. That this does not happen to an appreciable degree specifically implies that the data elements function well – indeed this could be one rationale for why NQF allows the use of performance score level analysis in the first place. These findings provide strong indirect evidence of the validity of our approach to capturing the measure's denominator.

There is nearly complete overlap of the denominator codes and there is overlap of the denominator elements. Where codes differ it is specific to decisions made by the CAPQuaM expert panel which was aware of the NCQA measures. Review of the medication lists for 0036 reveal that all medication used by the submitted CAPQuaM measure are also in the HEDIS measure. The CAPQuaM measure excludes specifically short acting beta agonists and leukotriene inhibitors at the specific direction of the CAPQuaM expert panel. We also specify exclude indacaterol from the list of "asthma specific medications" since it is a long acting beta agonist which is only indicated in the USA for treatment of COPD, which is a specific exclusion criterion for this measure.

Further, we identify asthma visits and medications using the same data that an insurance company or Medicaid would use for payment, including ICD codes, CPT codes, and revenue codes. We have had conversations with expert coders and New York State Department of Health Office of Health Insurance Programs to confirm our choices. Our literature review found that while there is moderate agreement (kappa = 0.45 - 0.50) when comparing administrative data regarding the presence of constructs such as recent asthma attacks, use of asthma medications, attack or medication, attack and medication, using 1 year of administrative claims data to parent report, the agreement improves from 0.55 to 0.60 when using two years of data.(1) We expect that these kappas would be significantly higher were the analyses restricted to children with disease that met our construct criteria for identifiable asthma.

The literature further supports our work. ICD-9 and ICD-10 codes for asthma on patients' medical charts typically match claims data. ICD-9-CM administrative data have been validated using various methodologies for various purposes (2-10). As examples: Jollis et. al. compared insurance claims data to the clinical database data to identify patients using ICD-9-CM codes for selected diagnoses and found that when all diagnoses were included, overall kappa agreement was .75 (2). Lee et. al. compared heart failure diagnoses identified in ICD-9 to the Framingham clinical criteria as the gold standard and found a positive predictive value of 94.3% (3). Muhajarine et. al. compared self-reported heart health survey data to physician claims from a database registry and found an overall agreement for hypertension of 81.7% indicating moderate to high agreement(4).Quan et. al. tested administrative discharge data to chart data for recording of comorbidity information using a Charlson index for measurement. Overall agreement of the Charlson index was good between databases but decreased as burden of comorbidity increased. Despite the differences, the Charlson index score derived from the administrative data had an identical ability of predicting in-hospital mortality to the score derived from chart data (5). Weiner and colleagues advocate a broad use of administrative data for monitoring quality and our uses fall within their recommendations (6). Romano and Mark assessed the sensitivity and reliability of coding for common diagnoses and procedures using California discharge abstracts and found in 7 of 8 comorbidity categories, sensitivity exceeded 85% (7). Weingart et. al. used administrative data, specifically a complications screening algorithm to identify inpatient complications using physician judgment as the gold standard and found flagged complications in 68.4% of surgical cases and 27.2% of medical cases (8). Yasmeen et. al. examined the sensitivity and positive predictive value to validate the coding of obstetric diagnoses and procedures in hospital-reported data using the medical record as the gold standard and found that surgical procedures and birth deliveries were accurately reported with sensitivities and PPVs exceeding 90% (9). Quam et.al. found that claims data that includes diagnostic and pharmacy data yields a high level of concordance with the medical record and survey data in the identification of a specific medical condition (10). Studies have shown high sensitivity of 72% and specificity of 95% for high risk conditions with overall accuracy of 90% obtained from administrative billing data among children with high-risk conditions including asthma which made up 87%

of the high risk conditions (11), and high predictive value among adolescents and adults with asthma (12). Twiggs et. al. found that the combined use of both medical and pharmaceutical claims was more effective in identifying asthmatics than either one by itself (13). HEDIS criteria using administrative data support peer reviewed research, for example in patients with persistent asthma based on HEDIS criteria in five Medicaid programs (Colorado, Georgia, Indiana, New Jersey, Washington) using ICD-9-CM code 493.x to measure filling prescriptions of asthma control medication and the ratio of controller medication to the total number of medication prescriptions filled within one year (14). Fowles and colleagues report sensitivity and specificity of claims compared with ambulatory medical records to identify asthma was 0.82 and 0.99, respectively. Sensitivity of .82 using claims was higher than sensitivity using self-report at .64 (15). Wilchesky compared chart abstraction to diagnoses obtained from administrative database: asthma claims were highly specific, Sp= 96.76 (95%CI 96.5, 97.0). Although sensitivity for most conditions was below 60%, sensitivity was enhanced when all claims for services were assessed, as we propose to do (16). Bronstein et al found that 88.3% of diagnoses asthma on claims agreed with medical record, with a negative predictive value of 0.85 and a positive predictive value of 0.88.They conclude that claims are generally an accurate indicator of the content of a patient encounter. (17) Steinwachs et al. compared billed claims to medical records based on date of visit and diagnosis, on average, 90% of billed visits were documented in the medical record, for asthma there was 90.9 percent of billed visits in record on same date and 82.8 percent of billed visits with same diagnosis in record on same date. (18) Quan et al documented the validity of ICD-9-CM and ICD-10 coding systems in coding clinical information and found that ICD-10 data was generally comparable with that of ICD-9-CM data in recording clinical information (19). Regarding our capacity to identify exclusions, Quan et al found that claims had a PPV of 91.9, and a negative predictive value of 92.6, with *k* of 0.65 (substantial agreement[1]) compared to chart review for chronic pulmonary disease . ICD 10 performed similarly in this study (19).

From a public health perspective, asthma surveillance systems in several states, including Maine, North Carolina, Connecticut and Michigan, have shown the feasibility of using administrative data to identify children having asthma, based on primary and secondary diagnosis codes reported on inpatient and outpatient claims. In addition to identifying asthma, important demographic data such as gender, race/ethnicity, program of enrollment and county of residence (urbanicity) can be used to assess associations between utilization services for asthma, including ED visits or hospitalizations, and demographic characteristics. Risk factor information from administrative data can be used to target educational programs, clinical assessments, and treatment programs (20-23).

Researchers also classified children with evidence of persistent asthma using HEDIS criteria, (24). Another study showed the usefulness of ICD9 493.x to identify asthma for a quality measure using Maryland Medicaid Claims data (25). Like our measure, those researchers excluded children with a diagnosis of cystic fibrosis (ICD9 277) (25). Schneeweiss commented that misclassification errors from claims data are asymmetric, with specificity typically exceeding 95% and sensitivity often less (26). Such a pattern makes it unlikely that an accountable entity would be held accountable for patients that do not actually have asthma.

As noted in 1.67 above, as part of an alpha test for our measure we used a contractor to survey more than a dozen hospitals across three CAPQuaM measure sets. Responses from 10 hospitals were specific to asthma. We found that variables including date of birth, race, ethnicity, county of residence, primary and secondary diagnosis of asthma in the ED, hospitalizations, payment source, and others were reported to be readily available and easy to access within the medical record.

In light of the literature review and our alpha test, we attest that the data elements for the measure match those assessed in the literature and our alpha test, with most being supported by both the literature review and the alpha test. We further note that our data element use is consistent with health care industry standards.

Validity of Measuring ED Visits and Exemplar Panel Findings

A national expert panel was convened and applied the RAND/UCLA appropriateness method to reviewing the constructs underlying this measure. The 9 member panel also supported the measure itself without objection.

[1] The *k* value indicates a near perfect agreement (*k*: 0.81-1.0 between coded data and chart review data), substantial agreement (*k*: 0.61-0.80), moderate agreement (*k*: 0.41-0.60), and fair agreement (*k:* 0.21-0.40).

As the constructs of this measure are defined via the expert panel process and this is an innovative approach to measuring undesirable asthma outcomes, there is no gold standard or statistical analysis. As an outcome measure, no association with process needs to be tested, although the NHLBI guideline discusses ED visits and hospitalizations as undesirable and potentially preventable outcomes. We used Median Scores from the panel ratings of at least 7 to identify desirable constructs for the measure.

Some interesting exemplar ratings are shown below, with some key findings bolded:

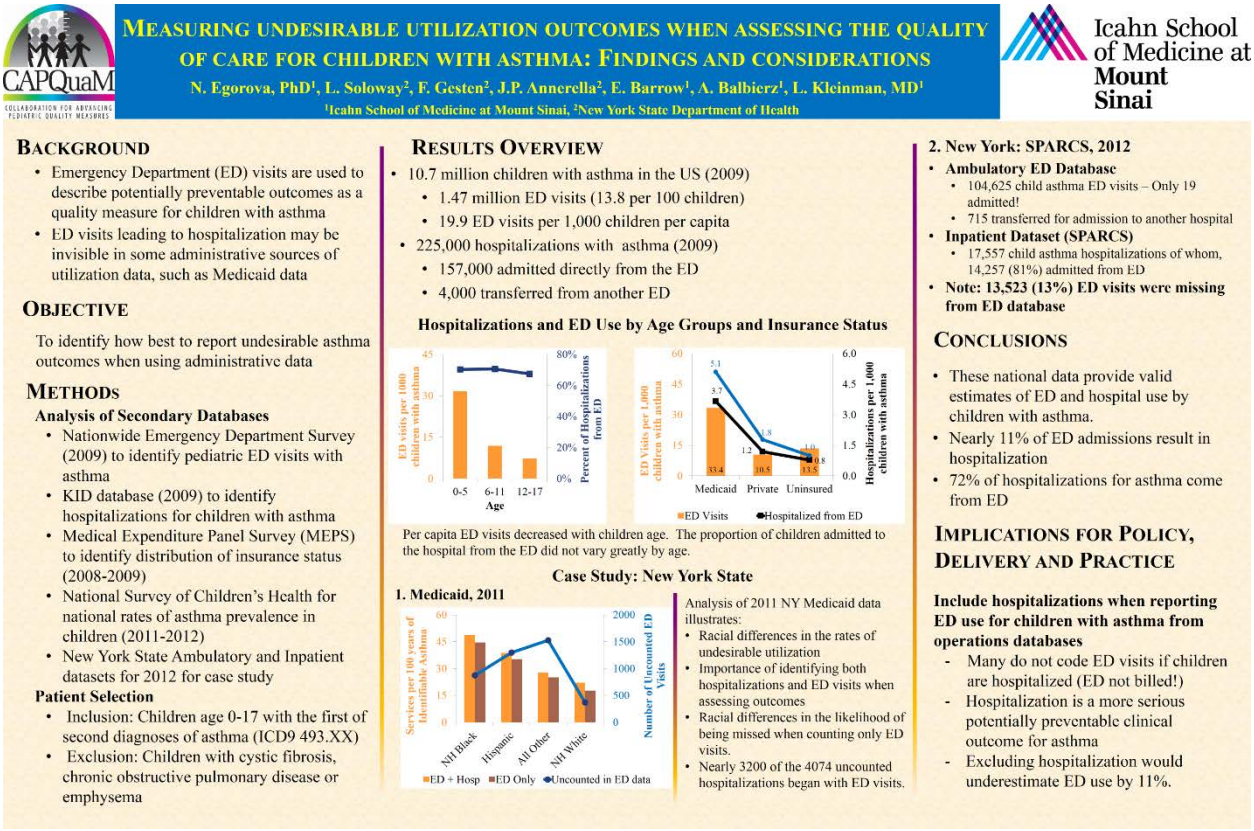
Scenario	MED
In general, this measure is intended to describe care for children who have asthma and identifiable since before the ED visit.	9
Asthma is established by a single prior hospital admission with asthma as the primary discharge diagnosis	9
A single admission is not sufficient to establish the presence of asthma.	1
In children after their 5th birthday, Asthma is established by a single prior ED visit with asthma as the primary discharge diagnosis	8
In children after their 5th birthday, Asthma is established by a single prior ED visit with asthma as the secondary discharge diagnosis	7
In children after their 5th birthday, Asthma is established by a single prior ED visit with asthma as any discharge diagnosis	6
In children after their 5th birthday, Asthma is established by 2 or more outpatient visits with asthma as a diagnosis.	9
In children after their 5th birthday, Asthma is not established until 4 or more outpatient visits with asthma as a diagnosis	2
Asthma related medication use helps to establish the presence of asthma.	**8**
Prescription for leukotriene inhibitors are typically asthma related..	**5**
Prescriptions for long acting beta 2 agonists are typically asthma related.	9
Prescriptions for inhaled steroids are typically asthma related.	8
Oral steroid bursts are typically asthma related.	5
In order to establish a diagnosis of asthma, a child should experience a total of at least 2 asthma related events such as outpatient visits for asthma and or asthma related prescriptions, one of which must be an outpatient visit	8
Filled prescriptions should not be considered when establishing the presence of asthma	1
Children with a diagnosis of COPD with chronic aspiration should be excluded from this measure.	**9**
Children with a diagnosis of COPD should be excluded from this measure.	**9**
Children with a diagnosis of cystic fibrosis should be excluded from this measure.	**9**
Children with a diagnosis of emphysema and chronic aspiration should be excluded.	**9**
Children with a diagnosis of emphysema should be excluded	**9**
The time frame for establishing a diagnosis of asthma extends before the reporting year.	**9**
This measure should include children over 2	8
The upper age limit for this measure should be children until their 20th birthday	4
The upper age limit for this measure should be children until their 21st birthday	**9**
For reporting purposes, adolescents 19-21 should be grouped with adolescents under 18.	5
For the purposes of this measure, only ED visits with asthma as the primary diagnosis are eligible for inclusion.	3
For the purposes of this measure, only ED visits with asthma as the primary or secondary diagnosis are eligible for inclusion.	**8**
For the purposes of this measure, all ED visits with asthma as a diagnosis are eligible for inclusion.	5
For the purposes of this measure, a treatment for asthma must be provided or prescribed in order for the ED visit to be eligible for inclusion.	3
In children prior to their 5th birthday, Asthma is established by 3 or more outpatient visits with asthma as a diagnosis	**9**
In children prior to their 5th birthday, Asthma is not established until 4 or more outpatient visits with asthma as a diagnosis	3

Our approach to identifiable asthma was validated by comparing the prevalence of identifiable asthma to the number of children with NY asthma claims, and to the prevalence estimate expected via analysis of the National Survey of Children's Health and to the prevalence of children with preventable asthma as defined by the NCQA's asthma measures. We sought to have a measure that would be much more inclusive than the persistent asthma criteria but still filtered with a threshold requirement. Indeed our findings supported this with more than 25% of all children with asthma claims eliminated by our definition, a denominator that was about 50% of the estimated survey-reported

lifetime incidence of asthma and 2.8 fold the number of children included than the NCQA criteria. We note that the NSCH survey prevalence exceeded the single claim approach.

Our own research looking at NY State Medicaid and national all payer data (see poster below, which was presented at peer-reviewed AcademyHealth national meeting) is consistent with expert and other recommendations that to identify all ED visits, one also needs to include hospitalizations for asthma as potential indicators of an otherwise unrecognized ED visit, which we have done and incorporated into the specifications.

In NY State using ED visits alone would miss about 13% of ED visits, nationally about 11%. The inclusion of hospitalizations will overestimate the number of ED visits by between 4 and 5 percent. As many of these hospitalizations are for acute exacerbations, the construct of undesirable utilization outcomes would include them, so that while the estimate is likely to be a bit high for ED visits, it is a fair estimate of asthma outcomes. Our approach to avoid de-duplication and double counting of an ED visit and its associated admission as two numerator events is specified (admission on same or next day in the same institution) in a manner that will slightly underestimate numerator events, thus compensating in part for the overestimation of ED events that may occur by including hospitalization.



REFERENCES

(1) Huzel, L. et al. Diagnosing Asthma: The fit between survey and administrative database. Canada Resp. Journal 2002.

(2) Jollis, J.G., M. Ancukiewicz, E.R. DeLong, D.B. Pryor, L.H. Muhlbaier, and D.B. Mark. 1993. ''Discordance of Databases Designed for Claims Payment versus Clinical Information Systems. Implications for Outcomes Research.'' Annals of Internal Medicine 119: 844–50.

(3) Lee, D.S., L. Donovan, P.C. Austin, Y. Gong, P.P. Liu, J.L. Rouleau, and J.V. Tu. 2005. ''Comparison of Coding of Heart Failure and Comorbidities in Administrative and Clinical Data for Use in Outcomes Research.'' Medical Care 43: 182–8.

(4) Muhajarine, N., C. Mustard, L.L. Roos, T.K. Young, and D.E. Gelskey. 1997. ''Comparison of Survey and Physician Claims Data for Detecting Hypertension.'' Journal of Clinical Epidemiology 50: 711–8.

(5) Quan, H., G. Parson, and W. Ghali. 2002. ''Validity of Information on Comorbidity Derived from ICD-9-CM Administrative Data.'' Medical Care 40: 675–85.

(6) Weiner, J.P., Powe,. R., Steinwachs, D.M., & Dent, G. (1990). Applying insurance claims data to assess quality of care: A compilation of potential indicators. Quality Review Bulletin, 16, 424-438.
(7) Romano, P.S., and D.H. Mark. 1994. ''Bias in the Coding of Hospital Discharge Data and Its Implications for Quality Assessment.'' Medical Care 32: 81–90.
(8) Weingart, S.N., L.I. Iezzoni, R.B. Davis, R.H. Palmer, M. Cahalane, M.B. Hamel, K. Mukamal, R.S. Phillips, D.T. Davies Jr., and N.J. Banks. 2000. ''Use of Administrative Data to Find Substandard Care: Validation of the Complications Screening Program.'' Medical Care 38: 796–806.
(9) Yasmeen, S., P.S. Romano, M.E. Schembri, J.M. Keyzer, and W.M. Gilbert. 2006. ''Accuracy of Obstetric Diagnoses and Procedures in Hospital Discharge Data.'' American Journal of Obstetrics and Gynecology 194: 992–1001.
(10) Quam, L., Ellis, L. B. M., Venus, P., Clouse, J., Taylor, C. G., & Leatherman, S. (1993). Using claims data for epidemiologic research: the concordance of claims-based criteria with the medical record and patient survey for identifying a hypertensive population. Medical Care, 31(6), 498-507.
(11) Daley MF, Barrow J, Pearson K, Crane LA, Gao D, Stevenson JM, et al. Identification and recall of children with chronic medical conditions for influenza vaccination. Pediatrics 2004;113(1 Pt 1):e26-33.
(12) Vollmer WM, O'Connor EA, Heumann M, Frazier EA, Breen V, Villnave J, Buist AS. Searching multiple clinical information systems for longer time periods found more prevalent cases of asthma. J Clin Epidemiol 2004;57:392-7.
(13) Twiggs, J. E., Fifield, J., Apter, A. J., Jackson, E. A., & Cushman, R. A. (2002). Stratifying medical and pharmaceutical administrative claims as a method to identify pediatric asthma patients in a Medicaid managed care organization. Journal of Clinical Epidemiology, 55(9), 938-944.
(14) Samnaliev, M., Baxter, J. D., & Clark, R. E. (2009). Comparative evaluation of two asthma care quality measures among Medicaid beneficiaries. Chest, 135(5), 1193-1196.
(15) Fowles, J. B., Fowler, E. J., & Craft, C. (1998). Validation of claims diagnoses and self-reported conditions compared with medical records for selected chronic diseases. Journal of Ambulatory Care Management, 21(1), 24-34.
(16) Wilchesky, M., Tamblyna, R. M., & Huang, A. (2004). Validation of diagnostic codes within medical services claims. Journal of Clinical Epidemiology, 57, 131-141.
(17) Bronstein, J. M., Santer, L., & Johnson, V. (2000). The use of Medicaid claims as a supplementary source of information on quality of asthma care. Journal for Healthcare Quality, 22(6), 13-18.
(18) Steinwachs, D. M., Stuart, M. E., Scholle, S., Starfield, B., Fox, M. H., & Weiner, J. P. (1998). A comparison of ambulatory Medicaid claims to medical records: A reliability assessment. American Journal of Medical Quality, 13(2), 63-69.
(19) Quan, H., Li, B., Saunders, L. D., Parsons, G. A., Nilsson, C. I., Alibhai, A., et al. (2008). Assessing validity of icd-9-cm and icd-10 administrative data in recording clinical conditions in a unique duallycoded database. HSR: Health Services Research, 43(4), 1424.
(20) Deprez RD, Asdigian NL, Oliver LC, Anderson N, Caldwell E, Baggott LA. Development of a prototype system for statewide asthma surveillance. Am J Public Health 2002;92:1946-51.
(21) Buescher PA, Jones-Vessey K. Using Medicaid data to estimate state- and county-level prevalence of asthma among low-income children. Matern Child Health J 1999;3:211-6.
(22) Children's Health Council. Asthma affects nearly one in ten children enrolled in HUSKY Part A in 1999–2000. Hartford (CT): Children's Health Council; 2001.
(23) Dombkowski, K. J., Wasilevich, E. A., & Lyon-Callo, S. K. (2005). Pediatric asthma surveillance: Using Medicaid claims. Public Health Reports, 120, 515.
(24) National Committee for Quality Assurance. Use of appropriate medications for people with asthma. HEDIS 2003, volume 2: technical specifications. Washington: NCQA; 2003.
(25) Stuart, B., Singhal, P. K., Magder, L. S., & Zuckerman, I. H. (2003). How robust are health plan quality indicators to data loss? A monte carlo simulation study of pediatric asthma treatment. HSR: Health Services Research, 38(6), 1547.
(26) Schneeweiss, S., & Avorn, J. (2005). A review of uses of health care utilization databases for epidemiologic research on therapeutics. Journal of Clinical Epidemiology, 58(4), 323-337.

2b2.4. 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?*)

Please see the section above. The face validity of our expert panel, the test-test reliability, the critical importance of having a standard, reliable, and valid approach to measuring the rate of asthma ED visits all support this measure.

We interpret our measure to be a valid estimate of the rate of ED visits and an even better estimate of undesirable outcomes from asthma.

Our interpretation is that administrative data are reliable for identifying asthma, and that year to year test retest reliability seems to indicate similar patterns of performance when identifying ED visits for asthma, reinforcing the reliability of our operational definitions for identifying eligible children. Our specification provide a sensitive and face valid approach to identifying an unbiased sample of children with ED visits (ensuring we don't bias the results towards the inappropriate by missing those with hospitalization).

Most databases contain consistent elements, are available in a timely manner, provide information about large numbers of individuals, and are relatively inexpensive to obtain and use. Validity of many databases has been established, and their strengths and weaknesses relative to data abstracted from medical records and obtained via survey have been documented (30). Administrative data are supported, if not encouraged by federal agencies, such as NIH, AHRQ, HCFA, and the VA. The Centers for Medicare & Medicaid Services has made clear to the participating AHRQ-CMS CHIPRA Centers of Excellence funded to develop measures in the Pediatric Quality Measures Program that it places a premium on feasibility when assessing those measures that it will most highly recommend to states to complete. The sources of data for the existing measure and other similar measures are typically based upon administrative data as well, providing consensual validation for using administrative data as the primary data source.

_____

2b3. EXCLUSIONS ANALYSIS

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

Exclusions are clinical and specifically guided by the explicit criteria developed by the expert panel.

2b3.1. 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*)

Denominator Exclusions: Children with concurrent or pre-existing: Chronic Obstructive Pulmonary Disease (COPD) diagnosis (ICD-9 Code: 496), Cystic Fibrosis diagnosis (ICD-9 code 277.0, 277.01. 277.02, 277.03, 277.09), or Emphysema diagnosis (ICD-9 code 492xx). Children who have not been consecutively enrolled in the reporting plan for at least two months prior to the index reporting month, as well as the index reporting month itself.

There are no numerator exclusions.

Exclusions were only included if they were endorsed by the expert panel. In studying the denominator we found that a very few percent of potentially eligible children (<=2.5%) were excluded by clinical diagnoses. The use of three months of continuous enrollment was recommended by our multi-stakeholder consortium and avoids the exclusion of more than 20% of otherwise eligible children from the population with identifiable asthma compared to a 12 month requirement.

2b3.2. 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*)

In order to develop a sample of approximately 125,000 children with asthma in our initial field test (that required a 12 month continuous enrollment criterion), we excluded 212 with COPD, 650 with cystic Fibrosis and 482 with emphysema

(those children were not mutually exclusive, in other words, children may have been excluded for more than one reason so the total number of exclusions was at least 212 and less than the sum of the three diagnoses (between 1.6% and 2.5% of otherwise eligible children).

Had we used a 12 month continuous enrollment criterion, we would have excluded more than 20% of otherwise eligible children.

2b3.3. 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*)

Exclusions are clinical and represent construct validity rather than statistical considerations.

The exclusions are purposeful and not statistical, and are based upon the findings of the expert panel. Noise is likely to be reduced by the exclusion of key diagnoses. Longer continuous enrollment requirements would harm validity since large number of children with real symptoms who are established and being managed for asthma would have been excluded. The 3 month continuous enrollment requirement is also conceptual, requiring the children to be under the management of the health plan in order to ask the plan to accept accountability.

2b4. 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.

2b4.1. 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** 1 **risk categories**
☐ **Other,** Click here to enter description

2b4.1.1 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.

2b4.2. If an outcome or resource use component measure is not risk adjusted or stratified, provide rationale and analyses to demonstrate that controlling for differences in patient characteristics (case mix) is not needed to achieve fair comparisons across measured entities.

Specifications for this measure requires stratification and reporting by age group only and also within age group by race/ethnicity. Several additional stratifications are optional but may be requested by the accountability entity or provided by the accountable entity. These variables include rurality/urbanicity and county level of poverty. Our findings suggest that risk adjustment is not critical for interpreting the results or for validity, but that stratification is informative to help to promote like to like comparisons and allow for plans to demonstrate how they do on specified subgroups. Such voluntary stratification specified in the measure helps to mitigate against the potential for misinterpretation and unintended consequences.

Within age group, we specify a number of stratifications as we have done for all of our CAPQuaM PQMP measure. Absent clear biological evidence that ED visits should be more likely in any of the sub categories we have chosen not to adjust but to report both topline and stratified results within age groups.

We used stratification to allow for a granular understanding of performance. Biological data and national guidelines agree and do not support risk adjustment to control for patient characteristics on the variables of interest. The Pediatric Quality Measures Program which funded development of this measure requests that measures be specified to be able to identify disparities and differences by a variety of characteristics and this measure does that.

The NIH NHLBI NAEPP guideline notes that goals of care and definition of successful management are the same regardless of baseline presentation. Hence clinical risk adjustment is not appropriate.

As indicated by the NHLBI guideline (http://www.nhlbi.nih.gov/files/docs/guidelines/asthgdln.pdf page 38)

> ***"An important point linking asthma severity, control, and responsiveness is that the goals are identical for all levels of baseline asthma severity.*** *A patient who has severe persistent asthma compared to a patient who has mild persistent asthma, or a patient who is less responsive to therapy may require more intensive intervention to achieve well-controlled asthma; however, the goals are the same: in well-controlled asthma, the manifestations of asthma are minimized by therapeutic intervention."*

For reasons other than controlling for case mix, we specify this measure to be stratified by age group and race/ethnicity as well as providing a top line analysis. Without such stratifications, racial and ethnic disparities (which have been found to be prevalent in children with asthma) might go unnoticed. The CHIPRA legislation that funded the development of these measures asks for the capacity to identify such disparities to be included in the measure specifications.

Specifications for further stratifications, such as by rurality/urbanicity and by county level of poverty are provided, in the event such stratification is requested by the accountability entity or desired by the reporting entity.

2b4.3. Describe the conceptual/clinical and statistical methods and criteria used to select patient factors (clinical factors or sociodemographic 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*)

Consistent with the Disparities Working Group of the Pediatric Quality Measurement Program, CAPQuaM has chosen an approach to not risk adjust for outcomes as being most appropriate to measuring actual performance. Nonetheless, we honor the parameters in the legislation funding the PQMP and also recognize the interest of various stakeholders in comparing like-to-like: hence we have specified key stratifications for analysis and presentation. The accountability entity has the option to request the granularity of stratification that suits its needs beyond age strata and race/ethnicity.

The conceptual model is that of CAPQuaM that includes that in pediatrics age is a key predictor and stratification is valuable. We were asked by AHRQ and CMS to include other constructs and we have manifest them as specified, such as race/ethnicity, poverty level in the caregivers county of residence, rurality/urbanicity on the caregiver's county of residence, insurance type and plan type, when variable. We have not added a stratum for children with special health care needs since asthmatics going to the emergency room are highly likely to belong in this category.

2b4.4a. What were the statistical results of the analyses used to select risk factors?

n/a

2b4.4b. Describe the analyses and interpretation resulting in the decision to select SDS 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)

2b4.5. 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*)

Provide the statistical results from testing the approach to controlling for differences in patient characteristics (case mix) below.

If stratified, skip to 2b4.9

2b4.6. Statistical Risk Model Discrimination Statistics (*e.g., c-statistic, R-squared*)**:**

2b4.7. Statistical Risk Model Calibration Statistics (*e.g., Hosmer-Lemeshow statistic*):

2b4.8. Statistical Risk Model Calibration – Risk decile plots or calibration curves:

2b4.9. Results of Risk Stratification Analysis:

Data presented in chart and figure show asthma outcomes stratified by age and race/ethnicity. Additional analyses showed meaningful stratifications by time of year, county level of poverty, and rurality/urbanicity.

Rate by Age and Race/Ethnicity (per 100 child-years)

Age Group	Race/Ethnicity	Rate	N
2 to 6 years	Non-Hispanic Not-Black	19.7	28,559
	Non-Hispanic Black	47.6	11,305
	Hispanic	32.7	22,524
7 to 12 years	Non-Hispanic Not-Black	12.3	34,766
	Non-Hispanic Black	27.9	16,825
	Hispanic	20.5	30,391
13 to 18 years	Non-Hispanic Not-Black	11.7	23,587
	Non-Hispanic Black	22.0	11,240
	Hispanic	15.1	18,251
19 or 20 years	Non-Hispanic Not-Black	16.9	8,088
	Non-Hispanic Black	31.1	2,797
	Hispanic	20.3	4,096

Rate by Age Group (per 100 child-years)

Age Group	Rate	N
2 to 6 years	29.7	62,388
7 to 12 years	18.7	81,982
13 to 18 years	15.1	53,078
19 or 20 years	20.5	14,981

Rate by Race/Ethnicity

(per 100 child-years)		
Race/Ethnicity	**Rate**	**N**
Non-Hispanic Not-Black	14.4	95,002
Non-Hispanic Black	22.3	75,262
Hispanic	31.0	42,168

All of these are statistically significant by chi square analysis, p well below 0.05.

2b4.10. 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*)

We acknowledge the association of the stratification variables with the rate of asthma ED visits but have not found evidence justifying such differences as either acceptable or un-modifiable by health care. Indeed there is evidence that primary care, adherence to guidelines, and other healthcare interventions can reduce or eliminated the impact of these factors. Federal guidelines quoted above support this perspective. We have expanded our requirement for stratification to create a "Both-And" presentation of the results and to enhance a more granular interpretation of the findings from this measure and to reduce the likelihood of misinterpretation or unintended consequences from measurement.

2b4.11. 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*)

_____

2b5. IDENTIFICATION OF STATISTICALLY SIGNIFICANT & MEANINGFUL DIFFERENCES IN PERFORMANCE

2b5.1. 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)*

Contingency table analysis with chi-square and using t-statistics were coherent and each illustrated the presence of statistical differences. Additionally, Poisson Regression analyses indicate significant differences by health plans and by counties, whether or not controlling for age group and race/ethnicity. As is desirable, the number of significant differences demonstrated depend upon which of the entities is used as a comparator in the analysis. Using a typical plan/county results in a smaller proportion of significant differences than selecting a more extreme plan/county.

2b5.2. 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*)

We have described ZIP models and their results in the sections of 2a2. Here we describe a complementary analysis using Poisson.

Differences between major subgroups were statistically significant, including race/ethnicity, age group, level of poverty in the county, and level of urbanicity (urban, suburban, rural). All one way and two way (within age stratum) chi square analyses and t test analyses were $p<0.05$.

We further note that the measure was sensitive enough to demonstrate face validity with statistically significant differences from month to month and season to season as expected for this outcome. Within the NY State Medicaid

data, differences were also found by eligibility category, which in this case can serve as proxy for health plan. Chi Square of the rate difference (7.7 ED visits per 100 children) between those qualifying for cash assistance versus those qualifying because of SSI was 32.07 with one degree of freedom meaningfully exceeds the critical value of 10.828 for p<.001. This demonstrates excellent capacity to distinguish between health plans.

Specifically, within the NY State Medicaid data, there are 22 plans identified. Using a poisson regression analyses, we found that 18 Managed Care Organization plans that included at least 900 children contributing time to the denominator yielded statistically significant differences among plans and among counties, whether or not we controlled for age group and/or race and ethnicity and/or urbanicity. This is true also when we analyzed stratified by age group.

8 of these plans were fully managed and 10 were partially managed plans. Partially managed plans had a statistically higher rate than fully managed HMO plans. .Among the 18 plans, the mean rate is 15.7, with a standard deviation of 6.0. A SAS summary of the rate distribution is shown below.

Quantiles (Definition 5)	
Level	Quantile
100% Max	30.03732
99%	30.03732
95%	30.03732
90%	28.70196
75% Q3	17.62075
50% Median	13.65690
25% Q1	11.67487
10%	9.19899
5%	9.07094
1%	9.07094
0% Min	9.07094

The exemplar table below shows when a randomly selected plan is chosen as the comparator in a model that controls for Black race, Hispanic ethnicity, and age group. Highlighted rows are statistically different from the index plan, demonstrating the power to for signal to be interpreted over noise, even when controlling for the above factors. Uncontrolled analyses produced analogous results.

Algorithm converged.

Analysis Of Maximum Likelihood Parameter Estimates							
Parameter		DF	Estimate	Standard Error	Wald 95% Confidence Limits	Wald Chi-Square	Pr > ChiSq
Intercept		1	-4.4112	0.0303	-4.4704 -4.3519	21264.0	<.0001
plan	1	1	-0.1376	0.0294	-0.1952 -0.0800	21.94	<.0001
plan	2	1	-0.1701	0.0386	-0.2458 -0.0944	19.39	<.0001
plan	3	1	0.1310	0.0274	0.0772 0.1848	22.76	<.0001
plan	4	1	-0.4571	0.0730	-0.6001 -0.3140	39.23	<.0001
plan	5	1	-0.4727	0.0479	-0.5665 -0.3789	97.56	<.0001
plan	6	1	0.0402	0.0265	-0.0117 0.0921	2.31	0.1287
plan	7	1	0.4882	0.0218	0.4455 0.5310	501.20	<.0001
plan	8	1	-0.0725	0.0733	-0.2162 0.0713	0.98	0.3232
plan	10	1	-0.2006	0.0551	-0.3086 -0.0925	13.22	0.0003
plan	11	1	-0.3049	0.0508	-0.4043 -0.2054	36.07	<.0001
plan	12	1	-0.3607	0.1098	-0.5760 -0.1455	10.79	0.0010
plan	13	1	-0.2146	0.0629	-0.3380 -0.0913	11.63	0.0006
plan	14	1	-0.3115	0.0722	-0.4529 -0.1700	18.62	<.0001
plan	15	1	-0.6056	0.0375	-0.6791 -0.5320	260.63	<.0001
plan	16	1	-0.5285	0.0587	-0.6435 -0.4136	81.15	<.0001
plan	17	1	0.4491	0.0213	0.4074 0.4908	445.49	<.0001
plan	21	1	-0.4093	0.0807	-0.5675 -0.2511	25.72	<.0001
plan	22	0	0.0000	0.0000	0.0000 0.0000	.	.
black		1	0.6985	0.0160	0.6671 0.7299	1902.69	<.0001
Hisp		1	0.2927	0.0154	0.2625 0.3229	360.50	<.0001
agegrp	1	1	0.3165	0.0268	0.2639 0.3691	139.06	<.0001
agegrp	2	1	-0.2037	0.0268	-0.2563 -0.1511	57.64	<.0001
agegrp	3	1	-0.4135	0.0285	-0.4694 -0.3576	210.25	<.0001
agegrp	4	0	0.0000	0.0000	0.0000 0.0000	.	.
Scale		0	1.0000	0.0000	1.0000 1.0000		

Note: The scale parameter was held fixed.

County by county results also varied significantly and were aggregated by urban influence code and level of poverty and continued to show capacity to show meaningful differences. The distribution of findings across 45 counties is shown below (unadjusted) (45 counties with at least 1000 months contributed to the denominator) the mean is 10.9 (std. 7.9).

Quantiles (Definition 5)	
Level	Quantile
100% Max	53.33333
99%	53.33333
95%	19.13721
90%	16.81866
75% Q3	12.68946
50% Median	9.02256
25% Q1	7.55814
10%	4.56100
5%	2.60530
1%	1.64249
0% Min	1.64249

Using a randomly chosen rural county, there were a number of significant differences county to county. Selecting New York City as the comparator, all other counties differed significantly. This nuance in findings confirms the validity and specificity of differences as likely to be meaningful. We note that NYS analyses frequently contrast findings with those in NYC and meaningful differences are the rule and not the exception.

2b5.3. 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?*)

The measures are sensitive enough to detect meaningful differences as observed within a population (as with the seasonal and month to month variations described above) and across populations such as counties or plans, as we have demonstrated. They are sufficiently robust and precise to measure real differences and not to create artificial ones. We interpret our findings in aggregate to indicate that the signal to noise ratio is very strong for this measure. There is high certainty and confidence that the performance measure scores are both reliable and valid.

2b6. COMPARABILITY OF PERFORMANCE SCORES WHEN MORE THAN ONE SET OF SPECIFICATIONS
If only one set of specifications, this section can be skipped.

Note: *This item is directed to measures that are risk-adjusted (with or without SDS 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 SDS 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.***

2b6.1. 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)

2b6.2. 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*)

2b6.3. 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*)

2b7. MISSING DATA ANALYSIS AND MINIMIZING BIAS

2b7.1. 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*)

We use administrative claims to establish eligibility. No missing data analysis performed as we were using standard data sources that are contractually obligated to be provided to NYS Medicaid. We had a total of three children in our analysis of children with identifiable asthma who dropped out of the analyses because of any missing data element.

Our analyses found that the absence of pharmacy data would reduce only slightly (as we recall, less than 1%) the number of children identified as having identifiable asthma. This finding became apparent during alpha testing of our specifications and was incorporated into our specifications as a permissive allowance when pharmacy data were not available. We have not located the original analysis and hope for the NY State team to replicate the analysis by the time of the Committee meeting.

2b7.2. 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*)

Addressed elsewhere regarding data sources and definition of identifiable asthma, requirements for 3 months of continuous enrollment. The use of a composite requirement to establish eligibility reduces that likelihood of systematic error or dependence upon any specific data field. The use of complementary sources of identifying visits (CPT codes and revenue codes) accomplishes a similar goal.

2b7.3. 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*)

Generally N/A. Systems unable to integrate pharmacy data into the eligibility analysis would have a minimally higher risk population than those with pharmacy claims. The specifics of the definitions and the limited impact of pharmacy claims on eligibility combine to make the expected impact of this on the rate of ED visits to almost zero. They are included in the identification of denominator because our expert panel directed us to do so. More importantly, as cited above, the NHLBI guideline tells us that outcomes should not be adjusted for baseline risk, so this does not truly disadvantage a reporting entity according to the guideline. Further, unlike many asthma measures where the absence

of pharmacy data would systematically disadvantage identification of satisfactory performance, for this measure, pharmacy data is used only to complement other utilization data when determining eligibility. In theory there is likely to be a short period of time when a child would be identified using the pharmacy data and not by the other utilization criteria prior to meeting other criteria.

3. Feasibility

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

3a. Byproduct of Care Processes

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

3a.1. **Data Elements Generated as Byproduct of Care Processes.**
Coded by someone other than person obtaining original information (e.g., DRG, ICD-9 codes on claims), Other
If other: In rare instances, race/ethnicity and zip code may need to be abstracted from the chart.

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 a PRO-PM, consider implications for both individuals providing PRO data (patients, service recipients, respondents) and those whose performance is being measured.
Medicaid systems typically include race and ethnicity and zip code as defined electronic fields.

Emergency department visits that result in hospitalization are often not coded as ED visits in administrative data. The most valid estimate of ED visit rate requires use of both ED and hospitalizations as numerator events.

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)*.
None at present

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)
Not in use	

4a.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

4a.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?*)

We are awaiting NQF endorsement for use. There are no policies or actions of the developer/steward or accountable entities that would restrict access to performance results or impeded implementation.

The measure is not currently in use but its application is currently being explored by the UCSF Center of Excellence in the Pediatric Quality Measurement Program (Round 2).

The topic of ED asthma use was assigned to our measure development project in the Pediatric Quality Measures Program by CMS, by far the largest single third party payer for medical care for children in the US, and by AHRQ. This measure has received the imprimatur of the American Academy of Pediatrics as one of its high priority measures that emerged from their joint (with the American Board of Pediatrics) Measurement Alignment and Strategic Selection Work Group.

4a.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.*)

The measure is a straightforward approach to estimating the rate of Emergency Department visit Use for children managed for identifiable asthma. Our analyses in NY State Medicaid data confirmed feasibility, usability, and responsiveness of the measure to substantive constructs including race/ethnicity, age group and by health plans. We find these data and their consistency with expected findings to be persuasive that the measure is both valid and sensitive to real differences. Therefore, when this measure is endorsed by NQF, it will be applicable to a variety of settings and organizations.

As a part of our work with PQMP, we are working on specific plans for dissemination and use. Our plan for implementation includes submitting our application for measurement endorsement from the National Quality Forum. We are having conversations with NY State Medicaid (who was one of our partners in development) regarding the application and use of this measure. No time frames have been established.

Meeting the expected timeframes of NQF, CAPQuaM intends for the measure to be used for an accountability application within 3 years of initial endorsement and public reporting within six years of initial endorsement.

At this point in time, the submitted measure has received the imprimatur of the American Academy of Pediatrics as one of its high priority measures that emerged from their joint (with the American Board of Pediatrics) Measurement Alignment and Strategic Selection Work Group.

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.

4b. **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.
Rates of emergency department visit use for children managed for identifiable asthma is an important outcome measure with intrinsic value that helps ensure high-quality, efficient healthcare for individuals and populations. Not all ED visits for asthma are necessary, some cases require a different level of care for the clinical circumstance. Also, a significant proportion of visits potentially could have been prevented with better prior management. A variety of stakeholders benefit from this measure:

- Plans could provide clinicians this data to use to accurately identify patients who benefit from enhanced asthma care.

- Health systems can use this data to distinguish patients who have identifiable asthma, their demographics, and the care they receive and the associated costs. This information allows practices, groups and facilities to evaluate and compare treatment plans between practice sites, medical and other professional groups and between integrated or other delivery networks. This evidence-based evaluation promotes the adoption of more effective and efficient health systems.

- States and healthcare agencies can also use these measures to compare larger systems to test and evaluate treatment options, payment models (e.g. managed care, primary care case management), quality of health plans, costs and health outcomes. Findings can be stratified by state or regionally (e.g. urban, rural, health shortage regions) to understand policy, demographic and culture effects.

- The data also allows clinical, public health and epidemiology researchers to understand the type, level and cost of care patients are receiving related to their health outcomes, giving opportunity to compare between health plans, payment models and treatment options. It also gives a deeper understanding of how individual (e.g. sex, age, gender, race, social economic status, poverty) and community determinants (e.g. work environment, community benefits) affect the rate of ED visits over time. Furthermore, these measures gives researchers the tools to identify and reach out to patients and their families to understand their health culture and practices to ensure that health services offered will most likely be utilized.

4c. 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).

4c.1. **Please explain any unexpected findings (positive or negative) during implementation of this measure including unintended impacts on patients.**
There has not been any evidence of unintended negative consequences to individual or populations. There are no anticipated unintended consequences if measuring at the level of comparing states, geographic regions, payment models, or health plans.

Comparing individual health care professionals is not recommended as care is provided across practices may be necessary. Also, it is not appropriate for a single hospital comparison because it is measuring the system performance not the hospital performance. Lastly, although the measure can be used to compare practice sites, medical or other professional groups or integrated or other delivery networks, the measures are only recommended for large practices or integrated delivery systems that own their own risk and manage inpatient and outpatient care or that have access to all payer data sources

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

4d1.1. **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.

4d1.2. **Describe the process(es) involved, including when/how often results were provided, what data were provided, what educational/explanatory efforts were made, etc.**

4d2.1. **Summarize the feedback on measure performance and implementation from the measured entities and others described in 4d.1.**
Describe how feedback was obtained.

4d2.2. **Summarize the feedback obtained from those being measured.**

4d2.3. **Summarize the feedback obtained from other users**

4d.3. **Describe how the feedback described in 4d.2 has been considered when developing or revising the measure specifications or implementation, including whether the measure was modified and why or why not.**

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?
No

5a.2. **If the measure specifications are not completely harmonized, identify the differences, rationale, and impact on interpretability and data collection burden.**
Our definition of identifiable asthma is more inclusive than, for example, NCQA's persistent asthma construct. We use similar medication definitions as NCQA, except we exclude leukotriene inhibitors from asthma-related medications because our expert panel felt that these medications were used frequently for allergy patients and judged that the small gain in sensitivity of identifying children (considering all criteria) would be less than the loss in sensitivity and likelihood to include non-asthmatic

children with allergies. Our specifications have been validated by an expert panel in the context of a peer reviewed process commissioned by AHRQ and CMS to advance the field and science of pediatric quality measurement beyond the state represented in pre-existing measures. The specification of a person-time denominator allows for the measure to have a shorter requirement for continuous enrollment than other measures with less risk of bias than previous measures.

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 **Attachment:** Appendix_Asthma_rate_12_12_16.docx

Contact Information

Co.1 **Measure Steward (Intellectual Property Owner):** University Hospitals Cleveland Medical Center
Co.2 **Point of Contact:** Lawrence, Kleinman, drlarrykleinman@gmail.com, 216-286-6969-
Co.3 **Measure Developer if different from Measure Steward:** University Hospitals Cleveland Medical Center
Co.4 **Point of Contact:** Lawrence, Kleinman, drlarrykleinman@gmail.com, 216-286-6969-

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.
Role: Expert Panelists
Elizabeth Allen, M.D Ohio State University
Chitra Dinakar, M.D. Department of Pediatrics University of Missouri-Kansas City
Stephen Teach, M.D., M.P.H. Children's National Medical Center
Charles Macias M.D., M.P.H. Pediatrics, Baylor Texas Children's Hospital
Michael Cabana M.D., M.P.H. University of California, San Francisco

Barbara Yawn M.D., M.S. Olmstead Medical Center, University of Minnesota
Joan Connell M.D. University of North Dakota
Delaney Gracy M.D., M.P.H. The Children's Health Fund and Montefiore Children's Hospital
Sharlene Miner M.D. Emergency Medicine Inter Mountain Medical Center

ROLE: Steering Committee and Investigator
Wilson Pace, MD American Academy of Family Physicians – DARTNET Institute - University of Colorado
Lynn Olson, PhD American Academy of Pediatrics
Christina Bethell, PhD, MBA, MPH Child and Adolescent Health Measurement Initiative, Johns Hopkins University (Previous OHSU)
Elizabeth Howell, MD Icahn School of Medicine at Mount Sinai
Harold Kaplan, MD Icahn School of Medicine at Mount Sinai
Lawrence Kleinman, MD, MPH Icahn School of Medicine at Mount Sinai

Rebecca Anderson Mount Sinai Medical Center
Ira Nash, MD Northshore - Long Island Jewish Medical Center (previous Mount Sinai School of Medicine)
Eyal Shemesh, MD Icahn School of Medicine at Mount Sinai
Mary Barton, MD National Committee on Quality Assurance
Charles Homer, MD, MPH US Department of HHS (previous National Institute for Child Health Quality)
Marla Clayman, PhD American Institutes for Research (previous Northwestern University)
Foster Gesten, MD New York State Dept. of Health, Office of Health Insurance Programs
Jerod M. Loeb, PhD The Joint Commission
Robert Rehm National Committee for Quality Assurance
Steve Kairys, MD American Academy of Pediatrics/QuIIN
Erin DuPree, MD The Joint Commission (previous: Icahn School of Medicine at Mount Sinai)
David Baker The Joint Commission (Previous Northwestern)
Beverley Johnson, BSN* Institute for Patient- and Family-Centered Care
Doris Peter Consumers Union

ROLE: Senior Advisory Board Member and Investigator
Shoshanna Sofaer, DrPH American Institutes for Research (previous CUNY Baruch)
Lynne Richardson, MD Icahn School of Medicine at Mount Sinai
Ian Holzman, MD Icahn School of Medicine at Mount Sinai
Marilyn Kacica, MD New York State Dept. of Health, Division of Family Health

ROLE: Senior Advisory Board Member
Marc Lashley, MD Allied Pediatrics
Gary Mirkin, MD Allied Pediatrics
Barbara Kupferman, RN, CMCN, LNC AmeriChoice by United Healthcare
L. Gregory Pawlson, MD, MPH SE Healthcare QUALITY Consultants (previous Blue Cross Blue Shield)
John Santa, MD* (previous Consumers Union)
John Clarke, MD ECRI, PA Patient Safety Authority
Scott Breidbart, MD Empire Blue Cross Blue Shield/ Anthem
Jeff Terry, MBA GE Healthcare
Robert St. Peter, MD, MPH Kansas Health Institute
Ruth Stein, MD Montefiore Children's Hospital
Arthur Aufses, MD Icahn School of Medicine at Mount Sinai
Eric Rose, MD Icahn School of Medicine at Mount Sinai
Wendy Brennan, MS National Alliance on Mental Illness of New York City (NAMI - NYC Metro)
Laurel Pickering, MPH Northeast Business Group on Health
Tony Hope, BM BCh, PhD Oxford University, The Ethox Center
Martin Hatlie, JD Partnership 4 Pt. Safety, Consumers Advancing Patient Safety
Paul Wise, MD Stanford University
Mark Chassin, MD, MPH, MPP The Joint Commission
Ira Nash, MD Icahn School of Medicine at Mount Sinai

ROLE: Investigators/Key Staff
Rusty McLouth, MS AAFP
Kasey McCracken, MPH CAHMI
Elise Barrow, MPH Icahn School of Medicine at Mount Sinai
Kasey Coyne, BA Icahn School of Medicine at Mount Sinai
Natalia Egorova, PhD Icahn School of Medicine at Mount Sinai
Elizabeth Howell, MD Icahn School of Medicine at Mount Sinai
Harold Kaplan, MD Icahn School of Medicine at Mount Sinai
Melissa Manice, MPH Icahn School of Medicine at Mount Sinai
Joshua Oppenheimer, BA Icahn School of Medicine at Mount Sinai
Eve Oster, MD Pediatric Partners of Gwinnett (previous Icahn School of Medicine at Mount Sinai)
Daniel O'Sullivan, BS University of Virginia (previous Icahn School of Medicine at Mount Sinai)
Michael Parides, PhD Icahn School of Medicine at Mount Sinai

Barbara Rabin, MHA	Icahn School of Medicine at Mount Sinai
Carolyn Rosen, MD	Icahn School of Medicine at Mount Sinai
Melissa Saperstein, MSW	Icahn School of Medicine at Mount Sinai
Eyal Shemesh, MD	Icahn School of Medicine at Mount Sinai
Andrew Ting, MD	Icahn School of Medicine at Mount Sinai
Adam Vella, MD	Icahn School of Medicine at Mount Sinai
Virginia Walther, MSW	Icahn School of Medicine at Mount Sinai
Jennifer Zeitlin, SD	Icahn School of Medicine at Mount Sinai, French Institute of Health and Medical Research (INSERM)
Marianne McPherson, PhD, MS	NICHQ
Joseph Anarella, MPH	NYS DOH, OHIP
Lee Sanders, MD, MPH	Stanford University
Kasey Coyne	Northwestern
Victoria Wagner	New York State Health Department Office Quality Patient Safety
Laura Soloway	New York State Health Department Office Quality Patient Safety
Kevin Malloy	New York State Health Department Office Quality Patient Safety
Judy Stribling	Icahn School of Medicine at Mount Sinai
Linda Radecki	AAP
Keri Thiessen	AAP – QuIIN
Laura Morris	New York State Department of Health
Olivia Lindly	OHSU
Kasey McCracken	OHSU
Sandeep Sharma	Icahn School of Medicine at Mount Sinai
Ann Nevar	UH Rainbow Babies and Children's Hospital/ CAPQuaM
Suzanne Lo	UH Rainbow Babies and Children's Hospital/ CAPQuaM
Amy Balbierz	Icahn School of Medicine at Mount Sinai
Samantha Raymond	Icahn School of Medicine at Mount Sinai
Allisyn Vachon	Icahn School of Medicine at Mount Sinai
Shannon Weber	Icahn School of Medicine at Mount Sinai
ROLE: Contractors	
Mary Fialkowski, MBA	The Joint Commission
Linda Hanold, MHSA	The Joint Commission
Elvira Ryan, MBA, BSN, RN	The Joint Commission
Sharon Sprenger, RHIA, CPHQ, MPA	The Joint Commission
Ann Watt, MBA, RHIA	The Joint Commission
Amy Slott	The Joint Commission
Denise Krusenoski	The Joint Commission
Elvira Ryan	The Joint Commission
Tasha Mearday	The Joint Commission

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:**

Appendix 1: Tables and Figures

Table of Contents

Figure 1. Eligibility Algorithm Diagram

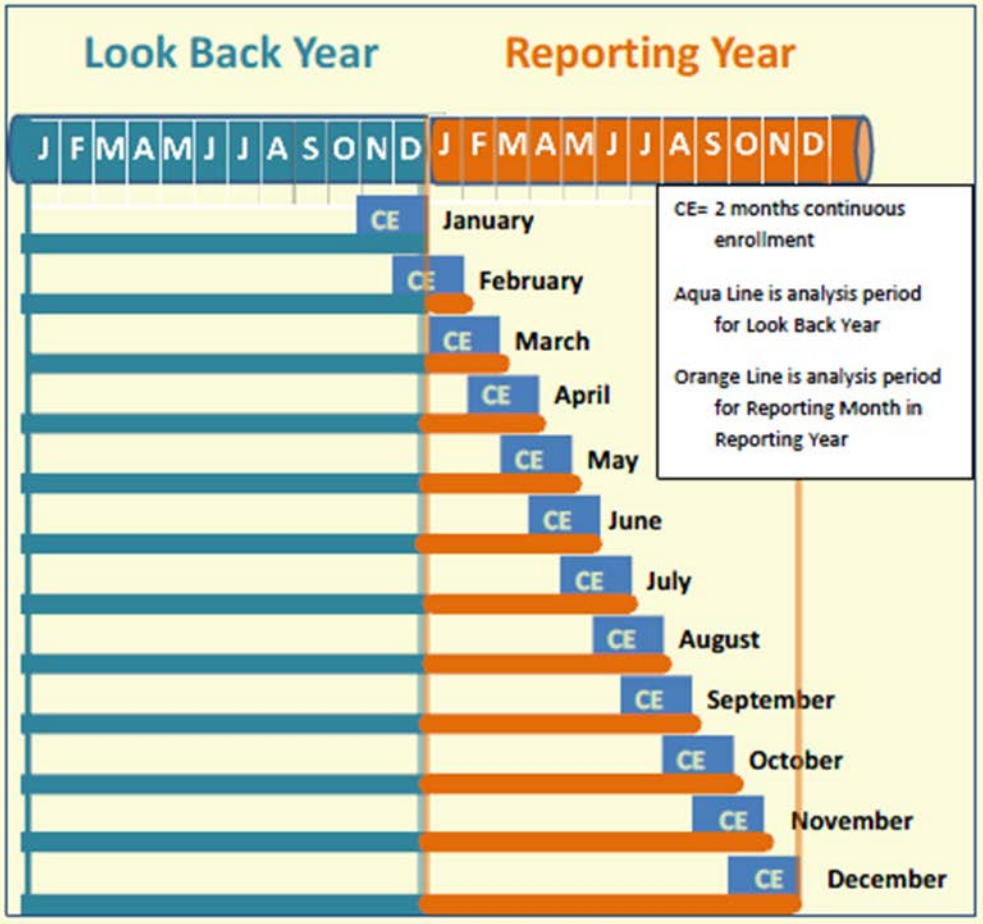


Diagram of Eligibility Requirements for Asthma Measure

Description of Diagram Components
Aqua: Look Back Year
Orange: Reporting Year
Blue: Continuous Enrollment (CE) period must be met to be eligible in a given reporting month
Black: Label of Reporting Month

Description of Eligibility Criteria Assessment
Eligibility Criteria are assessed month by month, using both the entire Look Back Year and the Reporting Year.

For any given reporting month, assess eligibility on 2 criteria. Eligible children are those that meet both of the following:

1) Has the child been continuously enrolled (CE) for the two months prior to the Reporting Month (3 months of continuous enrollment including the current Reporting Month)
2) Does the child meet the criteria for having been managed for persistent asthma in the assessment period (Aqua + Orange Lines)

Table 1. Month by Month Data, Stratified. New York State Medicaid Managed Care, 2012 In ED visits per 100 child-years.

Figure 2. Asthma ED Visits By Age and Month.

	2 - 5 y.o.	6-11 y.o.	12-18 y.o.	19-21 y.o.		Urban 1	Urban 2	Rural 1	Rural 2		Black	White	Hispanic
Jan	55								13.4	Jan	44.9	15.4	38.3
Feb	56								18.0	Feb	49.2	16.7	36.7
Mar	58								18.9	Mar	52.1	17.8	41.0
Apr	51								21.2	Apr	41.0	14.1	33.0
May	50								12.2	May	59.7	16.6	40.2
Jun	32.8	17.0	15.6	27.8	Jun	35.6	21.9	27.4	11.2	Jun	29.1	9.5	21.1
Jul	27.1	12.5	14.5	28.1	Jul	30.4	14.1	7.2	7.1	Jul	23.4	8.5	16.6
Aug	25.2	12.1	15.2	27.8	Aug	28.0	14.0	20.9	6.0	Aug	23.9	8.6	15.9
Sep	45.8	22.6	22.2	30.5	Sep	49.3	35.4	33.5	16.7	Sep	39.2	14.8	28.5
Oct	49.8	29.0	26.1	35.3	Oct	56.4	25.9	9.5	8.7	Oct	49.3	14.9	34.1
Nov	60.9	33.2	23.8	35.7	Nov	69.0	27.0	17.6	19.3	Nov	50.1	17.1	40.0
Dec	58.7	30.5	22.6	31.0	Dec	66.1	30.3	16.5	17.3	Dec	45.3	14.5	39.6

Asthma ED Visits By Age and Month

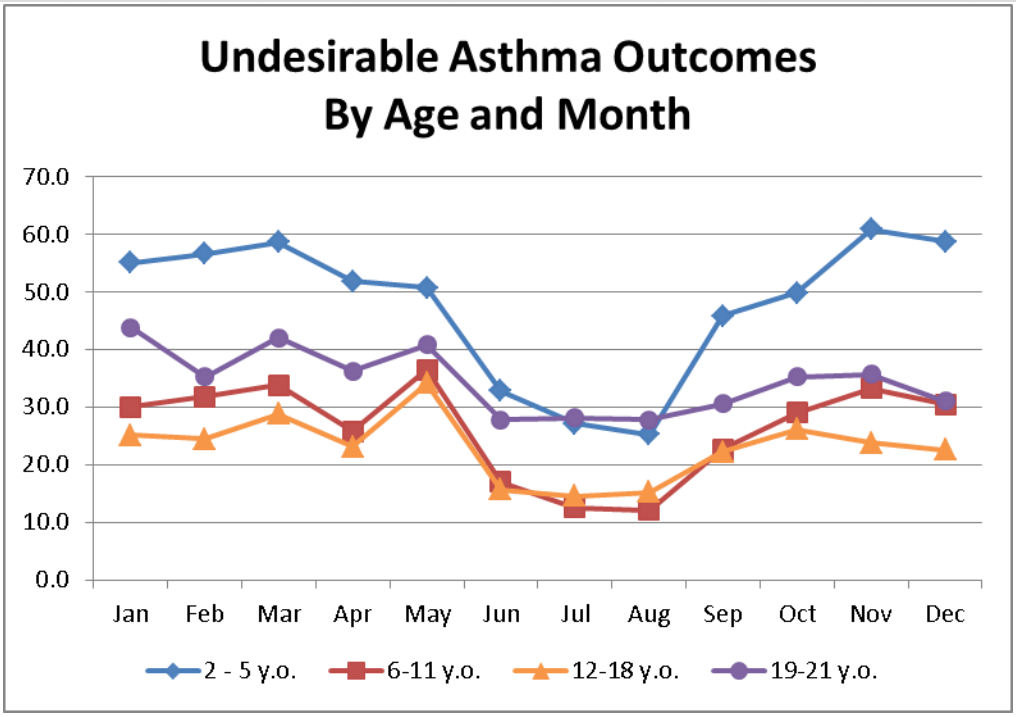


Figure 3. ED Visits per 100 Child-years by Age and Urbanicity

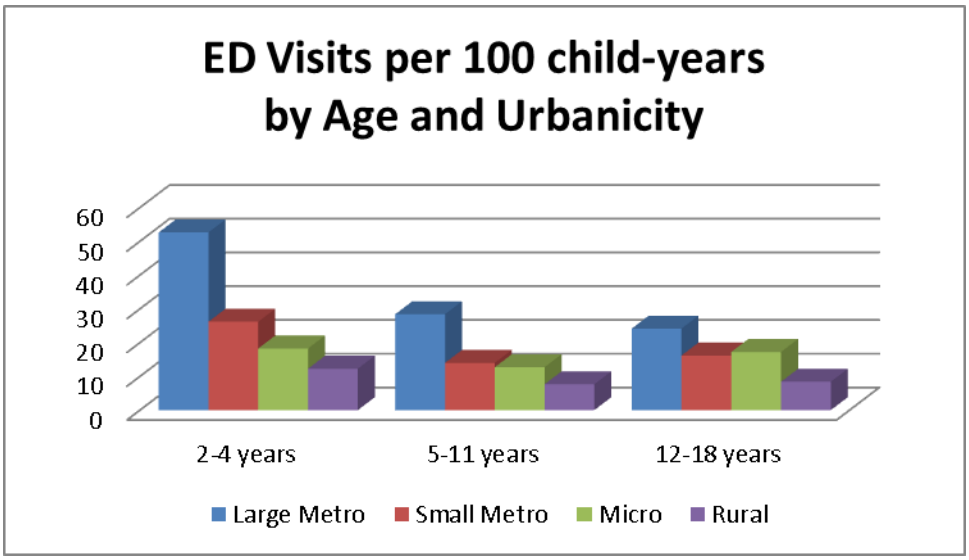


Figure 4. ED Visits per 100 child-years by Age and County Poverty quartile

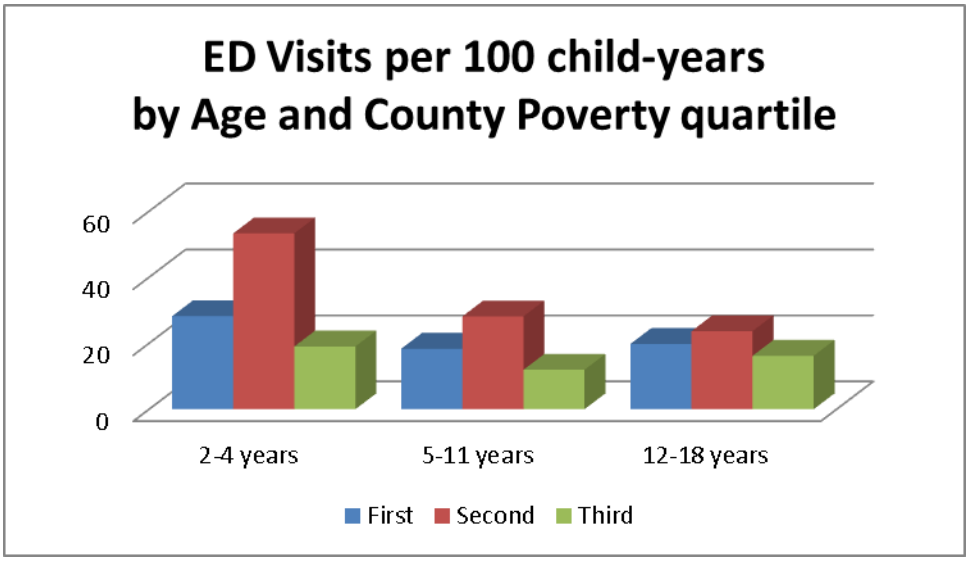


Figure 5. ED Visits per 100 Child Years by Age and Race/Ethnicity

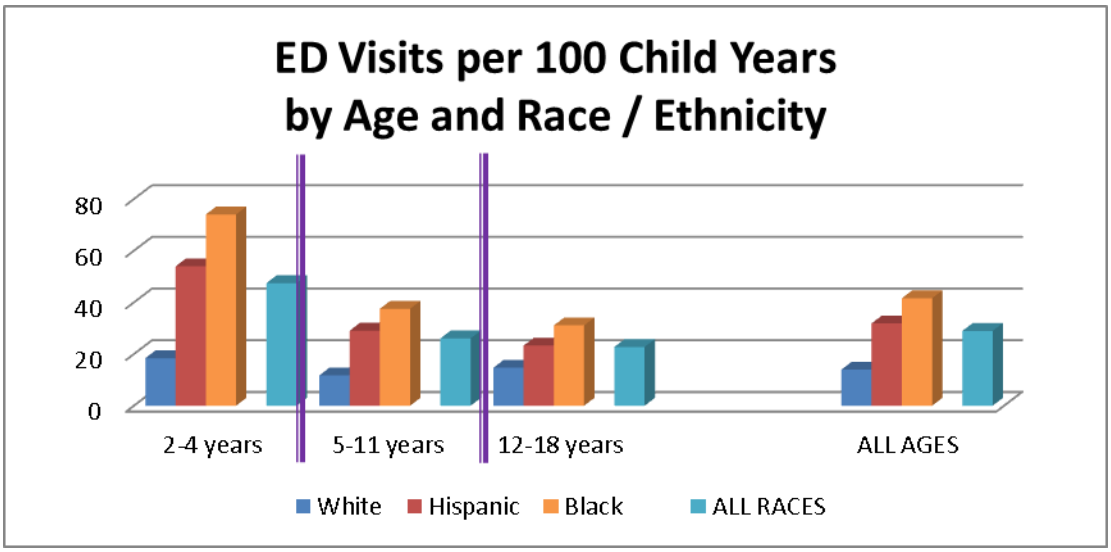


Table 3. ED Visits per 100 Child-years by Age and Quartile of Poverty

	White	Hispanic	Black		ALL RACES
2-4 years	18	54	74		47.44
5-11 years	12	29	38		26.03
12-18 years	15	23	31		22.74
ALL AGES	13.94	31.87	41.60		28.95

In visits per 100 child years.

Table 4. ED Visits per 100 Child-years by Age and Urbanicity

	Large Metro	Small Metro	Micro	Rural
2-4 years	52.6	26.2	18.3	12.3
5-11 years	28.4	14.0	12.7	7.6
12-18 years	24.1	16.2	17.2	8.4

Table 5. ED Visits per 100 Child-years by Age and Quartile of Poverty (First is lowest level of poverty)

Quartile of Poverty

	First	Second	Third
2-4 years	28	53	19
5-11 years	18	28	12
12-18 years	20	24	16

Table 6

Source of Systematic Review: • **Title**	

• **Author** • **Date** • **Citation, including page number** • **URL**	• National Heart, Lung, and Blook Institute, National Institutes of Health (NHLBI/NIH) Asthma Guideline 2007 www.nhlbi.nih.gov/guidelines/asthma (NAEPP Guideline) • http://www.nhlbi.nih.gov/health-pro/resources/lung/naci/asthma-info/asthma-guidelines.htm
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.	Quick Reference Guide: Asthma control focuses on two domains: 1)reducing impairment --- the frequency and intensity of symptoms… and 2) reducing risk – the likelihood of future asthma attacks… [later described as "prevent exacerbations] At the population level ED visits and hospitalizations represent failures of asthma control. Asthma Guidelines: - **Following science-based guidelines *works*** - Not only do they have the potential to improve a patient's *quality* of life; they can potentially *save a life*. **National asthma guidelines have been updated** In 2007, the National Asthma Education and Prevention Program (NAEPP), coordinated by the National Heart, Lung, and Blood Institute (NHLBI), released its third set of clinical practice guidelines for asthma. The Expert Panel Report 3—Guidelines for the Diagnosis and Management of Asthma (EPR-3) reflects the latest scientific advances in asthma drawn from a systematic review of the published medical literature by an NAEPP-convened expert panel. It describes a range of generally accepted best-practice approaches for making clinical decisions about asthma care. The EPR-3 emphasizes the importance of asthma control and focuses on two domains—current impairment and future risk—by which to assess asthma severity (for initiating therapy) and asthma control (for ongoing monitoring). EPR-3 also includes an expanded section on childhood asthma (with an additional age group), new guidance on medications, new recommendations on patient education in settings beyond the physician's office, and new advice for controlling environmental exposures that can cause asthma symptoms. **Asthma can be controlled** Scientific evidence clearly shows that most people could control their asthma by following current asthma clinical practice guidelines. With proper care, people who have asthma can stay active, sleep through the night, and avoid having their lives disrupted by asthma attacks. As a general rule, patients with well-controlled asthma should have: • Few, if any, asthma symptoms. • Few, if any, awakenings during the night caused by asthma symptoms. • No need to take time off from school or work due to asthma. • Few or no limits on full participation in physical activities. • No emergency department visits. • No hospital stays. • Few or no side effects from asthma medicines.

KEYPOINTS: OVERVIEW OF MEASURES OF ASTHMA ASSESSMENT AND MONITORING (pg. 36)

- The functions of assessment and monitoring are closely linked to the concepts of severity, control, and responsiveness to treatment:
    - Severity: the intrinsic intensity of the disease process. Severity is measured most easily and directly in a patient not receiving long-term-control therapy.
    - Control: the degree to which the manifestations of asthma (symptoms, functional impairments, and risks of untoward events) are minimized and the goals of therapy are met.
    - Responsiveness: the ease with which asthma control is achieved by therapy.
- Both severity and control include the domains of current impairment and future risk:
    - Impairment: frequency and intensity of symptoms and functional limitations the patient is experiencing or has recently experienced
    - Risk: the likelihood of either asthma exacerbations, progressive decline in lung function (or, for children, reduced lung growth), or risk of adverse effects from medication

KEY DIFFERENCES FROM 1997 AND 2002 EXPERT PANEL (pg. 37)
REPORTS

- The key elements of assessment and monitoring are refined to include the separate, but related, concepts of severity, control, and responsiveness to treatment. Classifying severity is emphasized for initiating therapy; assessing control is emphasized for monitoring and adjusting therapy. Asthma severity and control are defined in terms of two domains: impairment and risk.
- The distinction between the domains of impairment and risk for assessing asthma severity and control emphasizes the need to consider separately asthma's effects on quality of life and functional capacity on an ongoing basis (i.e., in the present) and the risks it presents for adverse events in the future, such as exacerbations and progressive loss of pulmonary function. These domains of asthma may respond differentially to treatment.

… p.38

An important point linking asthma severity, control, and responsiveness is that the goals are identical for all levels of baseline asthma severity. A patient who has severe persistent asthma compared to a patient who has mild persistent asthma, or a patient who is less responsive to therapy may require more intensive intervention to achieve well-controlled asthma; however, the goals are the same: in well-controlled asthma, the manifestations of asthma are minimized by therapeutic intervention.

	… page 41 regarding identification asthma, one key factor is: **The Expert Panel recommends that the clinician trying to establish a diagnosis of asthma should determine that (EPR–2 1997):** • **Episodic symptoms of airflow obstruction are present.** This is consistent with how we defined identifiable asthma… Page 63 It is important to evaluate the frequency, rate of onset, severity, and causes of exacerbations…severe exacerbations leading to ED visits and hospitalizations (Adams et al. 2000; Eisner et al. 2001; Ford et al. 2001; Lieu et al. 1998).
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	The National Asthma Education and Prevention Program (NAEPP) guidelines are the prevailing clinical recommendation for children with asthma. The Expert Panel Reports presenting clinical practice duielines for the diagnosis and management of asthma have organized recommendations for asthma care around four components considered essential to effective asthma management: - Measures of assessment and monitoring, obstained by objective tests, physical examination, patient history and patient report, to diagnose and assess the characteristics and severity of asthma and to monitor whether asthma control is achieved and maintained. - Education for partnership in asthma care - Control of environmental factors and comorbid conditions that affect asthma - Pharmacologic therapy This section of the report updates information on each of these four components based on the Expert Panel's review of the scientific literature. The sections that follow present specific clinical recommendations for managing asthma long term and for managing exacerbations that incorporate the four compoenents. Two evidence tables were prepared: (1) Predictors of Exacerbation: https://www.nhlbi.nih.gov/health-pro/guidelines/current/asthma-guidelines/evid_tbls/1_predexacer.pdf (2) Usefulness of Peak Flow Measurement: https://www.nhlbi.nih.gov/health-pro/guidelines/current/asthma-guidelines/evid_tbls/2_peakflowmeasure.pdf
Provide all other grades and definitions from the evidence grading system	Methodology for report: Overall Methods Used To Develop This Report **Background** In June 2004, the Science Base Committee of the NAEPP recommended to the NAEPP CC that its clinical practice guidelines for the diagnosis and management of asthma be updated. In September, under the leadership of Dr. Barbara Alving, M.D. (Chair of the NAEPP CC, and Acting Director of the NHLBI), a panel of experts was selected to update the clinical practice guidelines by using a systematic review of the scientific evidence for the treatment of asthma and consideration of literature on implementing the guidelines.

In October 2004, the Expert Panel assembled for its first meeting. Using EPR-2 1997 and EPR-Update 2002 as the framework, the Expert Panel organized the literature searches and subsequent report around the four essential components of asthma care, namely: (1) assessment and monitoring, (2) patient education, (3) control of factors contributing to asthma severity, and (4) pharmacologic treatment. Subtopics were developed for each of these four broad categories.

The steps used to develop this report include: (1) completing a comprehensive search of the literature; (2) conducting an indepth review of relevant abstracts and articles; (3) preparing evidence tables to assess the weight of current evidence with respect to past recommendations and new and unresolved issues; (4) conducting thoughtful discussion and interpretation of findings; (5) ranking strength of evidence underlying the current recommendations that are made; (6) updating text, tables, figures, and references of the existing guidelines with new findings from the evidence review; (7) circulating a draft of the updated guidelines through several layers of external review, as well as posting it on the NHLBI website for review and comment by the public and the NAEPP CC, and (8) preparing a final-report based on consideration of comments raised in the review cycle.

Preparation Of Evidence Tables

Evidence tables were prepared for selected topics. It was not feasible to generate evidence tables for every topic in the guidelines. Furthermore, many topics did not have a sufficient body of evidence or a sufficient number of high-quality studies to warrant the preparation of a table.

The Panel decided to prepare evidence tables on those topics for which an evidence table would be particularly useful to assess the weight of the evidence-e.g., topics with numerous articles, conflicting evidence, or which addressed questions raised frequently by clinicians. Summary findings on topics without evidence tables, however, also are included in the updated guidelines text.

Evidence tables were prepared with the assistance of a methodologist who served as a consultant to the Expert Panel. Within their respective committees, Expert Panel members selected the topics and articles for evidence tables. The evidence tables included all articles that received a "yes" vote from both the primary and secondary reviewer during the systematic literature review process. The methodologist abstracted the articles to the tables, using a template developed by the Expert Panel. The Expert Panel subsequently reviewed and approved the final evidence tables. A total of 20 tables, comprising 316 articles are included in the current update (see figure 1-1). Evidence tables are posted on the NHLBI Web site.

Ranking The Evidence

The Expert Panel agreed to specify the level of evidence used to justify the recommendations being made. Panel members only included

ranking of evidence for recommendations they made based on the scientific literature in the current evidence review. They did not assign evidence rankings to recommendations pulled through from the EPR-2 1997 on topics that are still important to the diagnosis and management of asthma but for which there was little new published literature. These "pull through" recommendations are designated by EPR-2 1997 in parentheses following the first mention of the recommendation. For recommendations that have been either revised or further substantiated on the basis of the evidence review conducted for the EPR-3: Full Report 2007, the level of evidence is indicated in the text in parentheses following first mention of the recommendation. The system used to describe the level of evidence is as follows (Jadad et al. 2000):

- **Evidence Category A:** Randomized controlled trials (RCTs), rich body of data. Evidence is from end points of well-designed RCTs that provide a consistent pattern of findings in the population for which the recommendation is made. Category A requires substantial numbers of studies involving substantial numbers of participants.

- **Evidence Category B:** RCTs, limited body of data. Evidence is from end points of intervention studies that include only a limited number of patients, post hoc or subgroup analysis of RCTs, or meta-analysis of RCTs. In general, category B pertains when few randomized trials exist; they are small in size, they were undertaken in a population that differs from the target population of the recommendation, or the results are somewhat inconsistent.

- **Evidence Category C:** Nonrandomized trials and observational studies. Evidence is from outcomes of uncontrolled or nonrandomized trials or from observational studies.

- **Evidence Category D:** Panel consensus judgment. This category is used only in cases where the provision of some guidance was deemed valuable, but the clinical literature addressing the subject was insufficient to justify placement in one of the other categories. The Panel consensus is based on clinical experience or knowledge that does not meet the criteria for categories A through C.

In addition to specifying the level of evidence supporting a recommendation, the Expert Panel agreed to indicate the strength of the recommendation. When a certain clinical practice "is recommended," this indicates a strong recommendation by the panel. When a certain clinical practice "should, or may, be considered," this indicates that the recommendation is less strong. This distinction is an effort to address nuances of using evidence ranking systems. For example, a recommendation for which clinical RCT data are not available (e.g., conducting a medical history for symptoms suggestive of asthma) may still be strongly supported by the Panel. Furthermore, the range of evidence that qualifies a definition of "B" or "C" is wide, and

the Expert Panel considered this range and the potential implications of a recommendation as they decided how strongly the recommendation should be presented.

Panel Discussion

The first opportunity for discussion of findings occurred within the "topic teams." Teams then presented a summary of their findings during a conference call to all members of their respective committee. A full discussion ensued on each topic, and the committee arrived at a consensus position. Teams then presented their findings and the committee position to the full Expert Panel at an in-person meeting, thereby engaging all Panel members in critical analysis of the evidence and interpretation of the data.

A series of conference calls for each of the 10 committees as well as four in-person Expert Panel meetings (held in October 2004, April 2005, December 2005, and May 2006) were scheduled to facilitate discussion of findings and to dovetail with the three cycles of literature review that occurred over the 18-month period. Potential conflicts of interest were disclosed at the initial meeting.

Report Preparation

Development of the EPR-3: Full Report 2007 was an iterative process of interpreting the evidence, drafting summary statements, and reviewing comments from the various external reviews before completing the final report. In the summer and fall of 2005, the various topic teams, through conference calls and subsequent electronic mail, began drafting their assigned sections of the report. Members of the respective committees reviewed and revised team drafts, also by using conference calls and electronic mail. During the calls, votes were taken to ensure agreement with final conclusions and recommendations. During the December 2005 meeting, Panel members reviewed and discussed all committee drafts.

During the May 2006 meeting, the Panel conducted a thorough review and discussion of the report and reached consensus on the recommendations. For controversial topics, votes were taken to ensure that each individual's opinion was considered. In July, using conference calls and electronic mail, the Panel completed a draft of the EPR-3: Full Report 2007 for submission in July/August to a panel of expert consultants for their review and comments. In response to their comments, a revised draft of the EPR-3: Full Report 2007 was developed and circulated in November to the NAEPP Guidelines Implementation Panel (GIP) for their comment. This draft was also posted on the NHLBI Web site for public comment in February 2007. The Expert Panel considered 721 comments from 140 reviewers. Edits were made to the documents, as appropriate, before the full EPR-3: Full Report 2007 was finalized and published. The EPR-3: Full Report 2007 will be used to develop clinical practice guidelines and practice-based tools as well as educational materials for patients and the public.

References

1. EPR. Expert panel report: guidelines for the diagnosis and management of asthma (EPR 1991). NIH Publication No. 91-

	3642. Bethesda, MD: U.S. Department of Health and Human Services; National Institutes of Health; National Heart, Lung, and Blood Institute; National Asthma Education and Prevention Program, 1991. 2. EPR-2. Expert panel report 2: guidelines for the diagnosis and management of asthma (EPR-2 1997). NIH Publication No. 97-4051. Bethesda, MD: U.S. Department of Health and Human Services; National Institutes of Health; National Heart, Lung, and Blood Institute; National Asthma Education and Prevention Program, 1997. 3. EPR-Update 2002. Expert panel report: guidelines for the diagnosis and management of asthma. Update on selected topics 2002 (EPR-Update 2002). NIH Publication No. 02 5074. Bethesda, MD: U.S. Department of Health and Human Services; National Institutes of Health; National Heart, Lung, and Blood Institute; National Asthma Education and Prevention Program, June 2003. 4. Jadad AR, Moher M, Browman GP, Booker L, Sigouin C, Fuentes M, Stevens R. Systematic reviews and meta-analyses on treatment of asthma: critical evaluation. BMJ 2000;320(7234):537-40. 5. NHIS. National health interview survey (NHIS 2005). Hyattsville, MD: National Center for Health Statistics (NCHS), Centers for Disease Control and Prevention, 2005. Available at http://www.cdc.gov/nchs/about/major/nhis/reports_2005.htm. Link to the evidence tables themselves: http://www.nhlbi.nih.gov/health-pro/guidelines/current/asthma-guidelines/evidence-tables
Grade assigned to the **recommendation** with definition of the grade	
Provide all other grades and definitions from the recommendation grading system	
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	**Systematic Evidence Review Overview** **Inclusion/Exclusion Criteria** The literature review was conducted in three cycles over an 18-month period (September 2004 to March 2006). Search strategies for the literature review initially were designed to cast a wide net but later were refined by using publication type limits and additional terms to produce results that more closely matched the framework of topics and subtopics selected by the Expert Panel. The searches included human studies with abstracts that were published in English in peer reviewed medical journals in the MEDLINE database. Two timeframes were used for the searches, dependent on topic: January 1, 2001, through March 15, 2006, for pharmacotherapy (medications), peak flow monitoring, and written action plans, because these topics were recently reviewed in the EPR-Update 2002; and January 1, 1997, through March 15, 2006, for all other topics, because these topics were last reviewed in the EPR-2 1997. **Search Strategies**

	Panel members identified, with input from a librarian, key text words for each of the four components of care. A separate search strategy was developed for each of the four components and various key subtopics when deemed appropriate. The key text words and Medical Subject Headings (MeSH) terms that were used to develop each search string are found in an appendix posted on the NHLBI Web site. **Literature Review Process** The systematic review covered a wide range of topics. Although the overarching framework for the review was based on the four essential components of asthma care, multiple subtopics were associated with each component. To organize a review of such an expanse, the Panel was divided into 10 committees, with about 4-7 reviewers in each (all reviewers were assigned to 2 or more committees). Within each committee, teams of two ("topic teams") were assigned as leads to cover specific topics. A system of independent review and vote by each of the two team reviewers was used at each step of the literature review process to identify studies to include in the guidelines update. The initial step in the literature review process was to screen titles from the searches for relevancy in updating content of the guidelines, followed by reviews of abstracts of the relevant titles to identify those studies meriting full-text review based on relevance to the guidelines and study quality. The combined number of titles screened from cycles 1, 2, and 3 was 15,444. The number of abstracts and articles reviewed for all three cycles was 4,747. Of these, 2,863 were voted to the abstract Keep list following the abstract-review step. A database of these abstracts is posted on the NHLBI Web site. Of these abstracts, 2,122 were advanced for full-text review, which resulted in 1,654 articles serving as a bibliography of references used to update the guidelines, available on the NHLBI Web site. Articles were selected from this bibliography for evidence tables and/or citation in the text. In addition, articles reporting new and particularly relevant findings and published after March 2006 were identified by Panel members during the writing period (March 2006-December 2006) and by comments received from the public review in February 2007.
Estimates of benefit and consistency across studies	In summary, the NAEPP "Expert Panel Report 3: Guidelines for the Diagnosis and Management of Asthma-Full Report 2007" represents the NAEPP's ongoing effort to keep recommendations for clinical practice up to date and based upon a systematic review of the best available scientific evidence by a Panel of experts, as well as peer review and critique by the collective expertise of external research/science consultants, the NAEPP CC members, guidelines implementation specialists, and public comment. The relationship between guidelines and clinical research is a dynamic one, and the NAEPP recognizes that the task of keeping guidelines' recommendations up to date is an increasing challenge. In 1991, many recommendations were based on expert opinion because there were only limited randomized clinical trials in adults, and almost none in children, that adequately tested clinical interventions grounded in research findings about the disease process in asthma. The large gaps in the literature defined pressing

	clinical research questions that have now been vigorously addressed by the scientific community, as the size of the literature reviewed for the current report attests. The NAEPP is grateful to all of the Expert Panel members for meeting the challenge with tremendous dedication and to Dr. William Busse for his outstanding leadership. The NAEPP would particularly like to acknowledge the contributions of Dr. Gail Shapiro, who served on NAEPP Expert Panels from 1991 until her death in August 2006. Dr. Shapiro provided valuable continuity to the Panel's deliberations while simultaneously offering a fresh perspective that was rooted in observations from her clinical practice and was supported and substantiated by her clinical research and indepth understanding of the literature. Dr. Shapiro had a passion for improving asthma care and an unwavering commitment to develop evidence-based recommendations that would also be practical. Dr. Shapiro inspired in others the essence of what NAEPP hopes to offer with this updated Expert Panel Report: a clear vision for clinicians and patients to work together to achieve asthma control.
What harms were identified?	
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	

Table 7

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	• Interventions to Modify Health Care Provider Adherence to Asthma Guidelines: A Systematic Review • Sande O. Okelo, Arlene M. Butz, Ritu Sharma, Gregory B. Diette, Samantha I. Pitts, Tracy M. King, Shauna T. Linn, Manisha Reuben, Yohalakshmi Chelladurai and Karen A. Robinson. • September 2013 • Systematic Review Okelo et al, Pediatrics 2013 132:3:S17-34 • http://pediatrics.aappublications.org/content/pediatrics/early/2013/08/20/peds.2013-0779.full.pdf
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.	Demonstrates several tools are effective in enhancing the quality of care and reduce undesirable outcomes.
Grade assigned to the **evidence** associated with the	

recommendation with the definition of the grade	
Provide all other grades and definitions from the evidence grading system	
Grade assigned to the **recommendation** with definition of the grade	
Provide all other grades and definitions from the recommendation grading system	
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	We followed the Agency for Healthcare Research and Quality Methods Guide for Effectiveness and Comparative Effectiveness Reviews (available at www. effectivehealth care.ahrq.gov/methods guide.cfm). Our protocol and the full report were subject to review. Data sources included Medline, Embase, Cochrane CENTRAL Register of Controlled Trials, Cumulative Index to Nursing and Allied Health Literature, Educational Resources Information Center, PsycINFO, and Research and Development Resource Base in Continuing Medical Education up to July 2012. Paired investigators independently assessed study eligibility. Investigators abstracted data sequentially and independently graded the evidence. RESULTS: Sixty-eight eligible studies were classified by intervention: decision support, organizational change, feedback and audit, clinical pharmacy support, education only, quality improvement/pay-forperformance, multicomponent, and information only. Half were randomized trials (n = 35). We identified 4217 unique citations of which 68 studies were eligible.
Estimates of benefit and consistency across studies	
What harms were identified?	
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	

Table 8

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	• Cochran Database of Systematic Reviews: Intermittent versus daily inhaled corticosteroids for persistent asthma in children and adults (Review) • Chauhan BF, Chartrand C, Ducharme FM • December 12, 2012 • Cochrane Database of Systematic ReviewsIntermittent versus daily inhaled corticosteroids forpersistent asthma in children and adults (Review)Chauhan BF, Chartrand C, Ducharme FMChauhan BF, Chartrand C, Ducharme FM.Intermittent versus daily inhaled corticosteroids for persistent asthma in children and adults.Cochrane Database of Systematic Reviews 2013, Issue 2. Art. No.: CD009611. • ~~Cochrane Review: Chauhan et al Cochrane Database Syst Rev 201212:CD009611~~ • https://www.ncbi.nlm.nih.gov/pubmed/23235678 • http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD009611.pub3/epdf
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.	Different approaches to treatment achieve different outcomes in children and adults (Daily achieves better asthma control than intermittent inhaled corticosteroids)
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	**Search methods** We searched the Cochrane A irways Group Specialised Register of trials (CAGR) and the ClinicalTrials.gov web site up to October 2012. **Selection criteria** We included randomised controlled trials (RCTs) that compared intermittent ICS versus daily ICS in children and adults with persistent asthma. No co-interventions were permitted other than rescue relievers and oral corticosteroids used during exacerbations. **Data collection and analysis** Two review authors independently assessed trials for inclusion, meth odological quality and e xtracted data. The primary e fficacy outcome was th e number of patients with one or more exacerbations requiring oral corticosteroids and the pr imary safety outcome was th e number of patients with serious adverse health events. Secondary outcomes included exacerbations, lung function tests, asthma control, adverse effects, withdrawal rates and inflammatory markers. Equivalence was assumed if the risk ratio (RR) estimate and its 95% confidence interval (CI) were between 0.9 and 1.1. Quality of the evidence was assessed using GRADE.
Provide all other grades and definitions from the evidence grading system	We identified trials from the Cochrane Airways Group Specialised Register of trials (CAGR), which is derived from systematic searches of bibliographic databases including the Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE, EMBASE, CINAHL, AMED, and PsycINFO, and handsearching of respiratory journals and meeting abstracts (see Appendix 1 for further details). All records in the CAGR coded as 'asthma' were

	searched using the following te r ms: (intermittent* or as-needed* or “as needed” or pr n or irregular* or occasional* or sporadic* or short-course*) and (daily* or regular* or routine*). We also conducted an advanced search of ClinicalTrials.gov using ’intermittent’ as keyword, ’asthma’ as condition and ’interventional studies’ as study ty pe. All databases were searched from their inception to October 2012 and th ere was no restriction on language of publication. The search for literature conducted until to December 2011 identified a total of 206 citations and abstracts through database searching and 26 citations from clinicaltrials.gov . Of th em, 16 full-text potential trials were reviewed and finally six trials (seven comparisons) were included for the meta-analysis (Figure 1). We updated the literature search in October 2012. There were 6 additional references, but no new included studies. This review summarises the best evidence available up to October 2012 derived from six trials (1211 patients with suspected or confirmed persistent asthma) of high methodological quality. The results pertain to children and adults with persistent asthma, and preschoolers with repeated wheezing suspected of persistent asthma. The systematic search to identify eligible trials and unpublished reports minimise the risk of inclusion bias. The outstanding collaboration of the authors/funders of six of the seven comparisons (Martinez 2011a; Martinez 2011b; Papi 2007; Papi 2009; Turpeinen 2008; Zeiger 2011) allowed us to obtain additional unpublished data and confirmation of methodological quality which strengthened the meta-analysis. Due to th epaucity of trials or the absence of events, 11 of 37 secondary outcomes could not be aggregated. While study authors reported enrolling patients with confirmed or suspected persistent asthma, the criteria used in paediatric trials (frequency of exacerbations with or without atopy, family history of asthma and eosinophilia) may have included an unknown propor tion of preschool children with intermittent viral-induced asthma that may have diluted the effect. The review is heavily weighted towards preschool- and school-aged children, with only two trials pertaining to adults. The long-term impact of intermittent versus daily ICS on lung growth, airway remodelling, bone mineralisation and adrenal function in children and lung function decline in adults beyond one year of follow-up remain to be addressed. **Quali ty of the evidence** The included trials were of high methodology and were generally at low r isk of bias. The confirmation of methodology by almost all authors or funders (with supportive evidence such as study protocols) and the provision of additional unpublished data allowed more precise estimates. The quality of evidence for our key outcomes reflects a lack of power from the studies that we included in the analysis (statistical imprecision) and variation in the different

	approaches used (indirectness). No potential biases were found in the review process.
Grade assigned to the **recommendation** with definition of the grade	
Provide all other grades and definitions from the recommendation grading system	
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	
Estimates of benefit and consistency across studies	
What harms were identified?	
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	

Table 9

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	• Quality of Care for Childhood Asthma: Estimtating Impact and Implications • Soeren Mattke, Francisco Martorell, Priya Sharma, Floyd Malveaux, Nicole Lurie • 2009 • S Mattke, et al. Quality of Care for Childhood Asthma: Estimating Impact and Implications. Pediatrics 123 Suppl 3, S199-S204. 3 2009. • https://www.ncbi.nlm.nih.gov/labs/articles/19221164/ • http://pediatrics.aappublications.org/content/pediatrics/123/Supplement_3/S199.full.pdf • ~~Systematic Review Mattke et al, Pediatrics 2009 123 S199-204~~
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.	Identified multiple gaps in asthma care quality. Key outcomes identified include hospitalizations and emergency department visits. Identified large racial disparities in use of inhaled corticosteroids
Grade assigned to the **evidence** associated with	

the recommendation with the definition of the grade	
Provide all other grades and definitions from the evidence grading system	
Grade assigned to the **recommendation** with definition of the grade	
Provide all other grades and definitions from the recommendation grading system	
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	We conducted a review of 164 relevant publications to consolidate the evidence on gaps in the quality of asthma care, the impact of those gaps, and the costs and benefits of closing those gaps. To identify relevant publications, a comprehensive review was performed of the English-language literature dating from 1995 to 2006, using the key words asthma, quality of care, treatment, care, therapy, disparities, inadequate, variability, differential, inequity, gap, variation, variance, medication, adherence, utilization, guideline adherence, disease management, impact, effect, morbidity, mortality, hospitalization, hospital admission, emergency room, emergency department, loss, absence, work, employment, workdays, school, attendance, absenteeism, and presenteeism. To obtain abstracts of articles published in peer-reviewed journals, the traditional health literature databases (ie, Medline, Embase, CINAHL, PsycINFO, Social Sciences Abstracts, and EconLit) were searched. Reference lists in relevant articles were mined for additional items. In addition, a number of databases for non–peer-reviewed literature and the Web sites of relevant governmental, professional, and advocacy organizations were searched. Two reviewers (Drs Mattke and Martorell) independently reviewed the titles of identified items, to assess whether the articles were likely to provide information on gaps in asthma care and their impact. The 2 reviewers then independently reviewed the abstracts or, if abstracts were unavailable, the full publications, to determine whether the publications contained information relevant to our study. Differences between reviewers were resolved through consensus. Reports were omitted from further consideration if they did not contain any data relevant to our research questions, were not conducted in the United States, were review or opinion articles, or were duplicative publications of the same data.
Estimates of benefit and consistency across studies	
What harms were identified?	
Identify any new studies	

conducted since the SR. Do the new studies change the conclusions from the SR?	

Overarching statement: Even when not specifically indicated, we are interested in how these constructs are impacted by such factors as race, ethnicity, socioeconomic status or its indicators, or the presence of other special health care needs.

Our metric is designed to capture axes related to two distinct conceptual frameworks:

1) Asthma is a model of chronic disease management. In other words, ED visits may arise from acute exacerbations indicating a flare up of disease, and/or suboptimal management of the chronic illness.
2) ED visits for asthma may reflect limitations of primary care beyond the provision of suboptimal treatment, such as insufficient education, limitations of access or availability, breakdowns of communication, or a variety of other factors.

We note that the internal quality of the ED visit to manage the asthma is not the target of this measure. However, communication between the emergency department and the primary care site may prove to be within the scope of this measure, pending the views of our experts and developers.

Construct I: Need to sufficiently specify population for measure

Concept	Implications (Lay Statement)	Lit Review Questions
(Descriptive) **The measure will need to adequately specify the population that we consider to be eligible for an ED with asthma measure.**	The development of measures regarding ED use for children with asthma requires us to understand the strengths and weaknesses for our measure of various approaches to identifying whether or not children have asthma. It further requires us to understand the impact of the availability of various sources of data (such as encounter data, pharmaceutical data, electronic medical record or chart review data) on these strengths and weaknesses. We are aware that the use of the term asthma is variable. We are not interested in diagnoses with the name asthma, but with an operational diagnosis that we will functionally treat as asthma, whether it has been called chronic wheezing, reactive airway disease, chronic infectious bronchitis, etc. We recognize that asthma and its presentation may change over the course of a child's life.	1. When asthma care is evaluated, how is the population of care recipients defined? How is asthma defined? What is the impact of including various types of data (dx 1 or more, drugs, etc) on the sensitivity and specificity of asthma identification? What are practical and valid approaches to identifying asthma? How do the answers to these questions differ between adults and children? 2. Are any groups persistently excluded from studies of asthma care (i.e., are children who have asthma and other comorbid conditions, such as a malignant disease, excluded?). What rationale is provided for the exclusion? 3. Are any non-asthma diagnoses considered to be indicators of asthma or potential asthma (e.g. bronchitis, bronchiolitis, wheezing, atopy) 4. For children up to age 21, how do issues of diagnosis, management, and follow-up differ by age and developmental stage? 5. At what point does literature suggest that reactive airway disease should be managed as asthma? a. What other conditions are managed as asthma? 6. What common current or preexisting comorbid conditions alter the management plan for asthma?

Construct II: Adequacy of management of asthma (as a chronic disease example)

Concept	Implications (Lay Statement)	Lit Review Questions

IIA. **↑Adequacy of asthma management:** **↓ED visits**	Since asthma is a chronic disease characterized by acute exacerbations, the extent to which asthma care is optimized through the use of appropriate medications, the control of the environment, and the preparation of the parent/child dyad to adapt to changes in circumstances (e.g. viral respiratory infection or exposure to cold) should reduce the number of ED visits, irrespective of the number of primary care visits.	1. What are the recommendations of the NHLBI guidelines? a. What does the literature suggest about the usefulness of NHLBI guidelines? b. Are there aspects that it has identified that appear to be missed? 2. What do we know about asthma management, how it's measured, who provides it, patterns of care and how ED visits vary as a consequence? 3. Does identification of PCP improve outcomes of ED visit, including patterns of care, utilization? 4. What do we know about the content of an asthma plan and its relationship to a full program of chronic disease management, and its influence on ED utilization? 5. What evidence is there about the impact on outcomes such as ED use when the child or adolescent is involved in asthma self-management? For example, does it matter if: a. The child has a written asthma plan? b. The child understands their asthma plan? c. The child is given an opportunity to participate in managing care? 6. How is the role of the child in self-management measured? 7. How much are children able to recognize, communicate and act on their asthma? 8. What do we know about the impact of asthma services on asthma management? This includes: a. Treatment from an asthma specialist; b. Social worker; or c. Multidisciplinary personnel 9. To what extent is ED use by children with asthma stimulated by non-asthma related issues? a. How can we identify when that occurs?

		b. What is the evidence that providing other services will reduce the number of ED visits? 10. To what extent do children contribute to their management (including avoiding triggers, recognizing symptoms, medication adherence, etc.)? a. What is the impact and variance by age? 11. What is the evidence regarding adequacy of various medication delivery systems for infants, toddlers, children and adolescents in acute and chronic settings? 12. Is there evidence of prior insult to the lungs such as sequelae of prematurity, etc. that create distinct subpopulations when considering this measure (at risk for ER visit)? 13. What aspects of the health services environment have been identified as contributing to outcomes of asthma management (e.g. school based health care)? 14. Does rate of ED utilization for non-respiratory diagnoses vary between asthmatics and non-asthmatics? 15. What is known about how often children with asthma use the ED over an extended period of time? Does it change over the life course of childhood? How does that vary by child characteristics, including race, SES, urban, suburban vs. rural, and age?

IIB. **↑PCP capacity/knowledge/skill:** **a. ↑Asthma management** **b. ↓Asthma exacerbations** **c. ↑Chronic disease management**	Broadly speaking, patient management of asthma is influenced by the capacity of the PCP practice. This includes the knowledge and skills possessed by the PCP, as well as office support to enhance access and availability of care. PCP includes the ability of the PC office to meet the cultural needs of the patient and their family.	1. What are the diversity of practices or services that may or may not impact ability or capacity of the PCP practice to manage asthma? 2. What do we know about the specific skills and processes that contribute to a primary care practice's capacity? 3. What patterns of visits or medication use or other indicators have been used as markers of well or poorly delivered primary care for asthma in children and/or adults? 4. What is the minimum use of specialists appropriate for children with asthma? How does that vary with history of ED or hospital use? a. When and how does the use of specialists become a marker for higher or lower quality of care? 5. What evidence is there regarding the nature of the PCP practice for children with asthma? For example, the level of continuity with individual clinicians vs. practices, the accessibility of specified clinicians and/or practices during the day and/or after work hours, etc.
IIC. **↑Asthma education:** **a. increases recognition of symptoms >** **b. ↑Management skills**	Enhancing what patients or their families know about asthma may be an important tool to improve care for children with asthma. The likely first effect of such education is to enhance the capacity of a caregiver to identify what symptoms may relate to asthma. This could conceivably increase utilization of both PCP and ED services if this were to increase the caregiver's perceived	1. What are metrics or processes regarding the quality of asthma care? Is it drug ratios (i.e. proportion of prescriptions filled that are for rescue vs control medications), asthma action plan, , capacity of PCP office, relationship to PCP practice, or other specific bundles of care, etc? 2. What constitutes "perfect care"/"best practice" for any specified type of patient? 3. What do we know about the impact of asthma education programs on quality of care, outcomes of care, or utilization of care? Define utilization of care as including: a. PCP utilization, b. ED utilization, c. Referral/specialist utilization,

	need for care for their child's asthma. With a more sophisticated understanding, including having a valid asthma action plan and understanding how to use it, ED care may be reduced and PCP care for asthma may be reduced, as symptoms are less frequent and parents are more competent to manage them when they arise.	d. Non physician care team member utilization, e. Medication usage, f. Hospitalizations, and/or g. Other care utilization areas to consider? Examples may include functional status, quality of life elements, spirometry, role functioning. 4. What is the diversity of asthma education programs and what are the differences in quality of care/outcomes/utilization of care associated with differences? 5. Does referral to an asthma specialist impact quality of care, utilization of care and asthma outcomes? 6. Does referral to a social worker impact utilization of care and asthma outcomes? 7. (Broad) Does involvement of multidisciplinary personnel (beyond allopathic or osteopathic physicians) impact quality of care, utilization of care and asthma outcomes? 8. What are desirable roles and effectiveness of interventions that extend beyond the healthcare system, such as reducing pollution, focusing on environmental justice, housing, dust mites, etc.? 9. How does organization and capacity of the practice setting influence the delivery of asthma management education?
Construct III: **Concept** **IIIA.** **↑Primary care capacity:** **a. ↑ PCP visits (routine, WCC)**	Adequacy of PCP practice site to Implications (Lay Statement) In general, enhanced capacity may affect a patient's access to	handle acute exacerbations of chronic disease and/or acute illnesses Lit Review Questions

b. ↑PCP visits (other acute dx) **c. ↑ PCP visits (asthma)** **d. ↓ED visits (acute dx, asthma)** **IIIA.2** **SUBCONSTRUCT:** **↑Accessibility:** **a. ↑ PCP visits (routine, WCC)** **b. ↑PCP visits (other acute dx)** **c. ↑ PCP visits (asthma)** **d. ↓ED visits (acute dx, asthma)**	care. Capacity can refer to patient services that make it easier for a patient to receive timely care, such as location or hours of offices, to the ability to triage phone calls in a timely and effective way, or may include the materials and services present within an office (e.g. the presence of a treatment room, the capacity to deliver oxygen, nebulizers, etc.) Such capacity may be limited or enhanced by staffing, space, the ability to safely transport someone from the office to a hospital, etc. If PCP office capacity is optimized, ED visits may be reduced as acute and mundane conditions can be managed in a PCP setting. Subsequently, increased capacity of the entire PCP support network will increase number of PCP visits.	1. What do we know about access to the PCP's office as a place to manage asthma, and the subsequent capacity of a PCP and the diversity of practice settings? Additionally, how do we measure capacity and, its impact on QoC, processes of care, asthma outcomes, asthma specific processes and utilization? How do these factors impact ED use or other outcomes? a. In general: i. PCP/specialist ratio in a plan or PCP/child ratio ii. PCP time spent in visit (incl. minutes per sick, well-child, asthma management visit) iii. Nature of training activities iv. How long does it take to schedule a visit (incl. asthma (chronic), acute, follow-up visit) v. Office hours and visit flexibility (incl. after hours coverage, office consult, meet in ED) vi. Phone capabilities: (incl. answering capacity, putting on hold, returning calls, after hours phone service) vii. Level of implementation of patient centered medical home/chronic care model, eg i. Use of registries ii. Standardized tools for measurement iii. Case management iv. Group visits or other education, etc

b. Specifically, ability to manage acute dx in office, which includes:
    i. Do they have a treatment room or capacity to use a room as a treatment room?
    ii. Do they offer rescue treatments (e.g. nebulizers, spacers)?
    iii. Can they measure oxygen saturation?
    iv. Do doctors feel comfortable with acute asthmatic in office?
    v. Can they take time to manage an acute pt in their office?
    vi. Do they have safe and rapid transport to a hospital (how long?)
2. Availability and accessibility of offices (incl. office hours, geographic distribution)
    a. What do we know about linguistic capabilities in the PCP setting influencing use of the ED?
    b. What do we know about proximity of the PCP office to public transit on the utilization of the ED?
3. What do we know about the impact of variations in patterns of care/practice, use of modalities, and/or and receipt of well-child care on asthma management or outcomes (eg ED use)? Does Immunization status reflect on t eh capacity of the PCP, on the state of the child, or on other factors that may relate to asthma outcomes? How about the sufficiency of the number of WCC Visits (eg meets HEDIS standard or AAP standard or does not)? Absolute number of visits to PCP?
4. Are children with more WCC visits less likely to use the ED for acute visits? children who are UTD on their immunizations?
5. What literature is there on the relationship between pediatric ED use

		and other measures of asthma exacerbation/outcomes? 6. What do we know about variability of capacity and management of mundane conditions (e.g. OM, URIs, pharyngitis), office to ED ratios? 7. What do we know about variability of capacity and management of acute conditions requiring interventions (e.g. asthma)? 8. To what extent does ED capacity increase use of ED services? Do hospitals advertise ED services, have fast track for mundane conditions, etc? 9. To what extent does ED have capacity to provide primary care, routine immunizations, etc? How is that built into policies and protocols? 10. At what age does the PCP start meeting alone with child? Time spent in visit? 11. To what extent and at what age do PCP's involve children in self-management and does it vary?

IIIB. **↑Relationship with PCP:** **a. ↑ PCP visits (routine, WCC)** **b. ↑PCP visits (other acute dx)** **c. ↑ PCP visits (asthma)** **d. ↓ED visits (acute dx, asthma)**	Improved relationship with PCP may increase visits to your PCP and decrease ED visits, for both acute and mundane conditions. A good relationship may lead to greater trust and adherence to recommendations (both WCC and asthma care) and drive a preference for seeking care by the PCP over seeking care in another environment. In general, we are referring to relationship of caregiver with PCP and their office staff. We recognize the importance of the relationship of PCP's with patients as well; when the relationship between the PCP and the child rather than caretaker is emphasized in research, we'd like to capture that as well.	1. What exists regarding measuring the quantity and quality of the relationship with PCP? Specifically: a. What's the variation and does it matter? b. How is it measured? c. What do we know about patient experience of care, especially as it relates to relationship with clinicians/PCP d. To what extent is quality of relationship expressed in terms of caregiver vs. child relationships and how does this change with age of child or longevity of connection to a PCP? 2. What evidence is there regarding use of supplemental services outside of regular clinical visits and how do these services impact quality and utilization of care? Define supplemental services as: a. Electronic educational/reminder tools (incl. social media) b. Telephone educational/reminder tools c. Print materials (e.g. educational brochures) d. Disease management, demand management, or other type programs e. Other services to consider? Measure quality, utilization of care should include at least : a. ED visits b. PCP visits 3. How does role of child in self care/management tie into these issues?

Construct IV: The connectedness of care in the primary care and ED setting – before, during, and after of the ED visit

Concept	Implications (Lay Statement)	Lit Review Questions

IV. (Descriptive) **Enhanced integration of ED care of asthma with routine care will have better outcomes**	If primary care is generally pretty good, then the ED visit should be an extraordinary event. In such cases the PCP alerting the ED to current management and the ED assuring appropriate follow up with the PCP is important. In cases where primary care is of lower quality or more variable, the ED visit may enhance the long term management of the child with asthma. And we need to assess this. One of the ways it might do so is to construct an asthma management plan that is then followed by the PCP. Another way is to connect a child without adequate primary care to primary care, especially to someone who is competent to manage the asthma.	1. What evidence supports that ED visits for asthma are most effective when visit is followed by a visit to the PCP? 2. Do utilization patterns in both the ED and primary care setting change following ED visits? 3. Is an effective/more effective use of medications seen following an ED visit? 4. Does the identification of a primary care provider improve outcomes of an ED visit (including patterns of care utilization)? 5. Is pre or intra visit communication with the primary care provider associated with better outcomes? How often does this occur? Are there systematic differences regarding those for whom this does and does not occur? 6. Are ED visits for asthma routinely associated with some form of communication or linkage with PCP? Does that result in better outcomes?

Construct V: Equity is a value in asthma care

Concept	Implications (Lay Statement)	Lit Review Questions
V. (Descriptive) **Equity is a critical construct of quality for children with equity**	Systematic differences in the frequency or nature of ED visits for asthma on the basis of race, ethnicity, family make-up, income/economic status, specifics of insurance status, presence or absence of comorbid special health care needs, etc represents decrements in quality that our measures should identify.	1. Does the literature indicate systematic or predictable differences in the frequency or nature of asthma care for children as it relates to ED visits for asthma that may be interpreted as representing inequitable structures, processes, outcomes, experiences with, or coordination of care? 2. What do we know about how social determinants and diagnosis and management of asthma and its outcomes, specifically as it relates to use of ED? 3. What do we know about the extent to which use of the ED for children with asthma that relates to the external physical and social environment?

Proposed Research Questions

Asthma- We propose to prioritize our Asthma Construct Table, to the following questions:

Acronyms

PCP: Primary Care Provider

ED: Emergency Department

WCC: Well-child care

Baseline Question (for Questions 1, 2 and 3 below):
When asthma care is evaluated, how is the population of asthma care recipients defined at the population level? What are specific implications of how you identify patients with asthma, including various approaches to specifying the denominator of children with asthma? What are practical and valid approaches to identifying asthma at the population level? How do the answers to these questions differ between adults and children?

Question 1 (Construct IIA.2):
For children with asthma, what do we know about asthma management? How is management of asthma described and measured? This includes who (PCP, asthma specialist, ED, etc) primarily manages it as well as who provides it. What are the patterns of care and what do we know about how use of the ED varies as a result of various approaches to management?

- **Question 1a (Construct IIB.3):**

Specifically, have any of these patterns of visits or medication use or other characteristics of care been used as markers of well or poorly delivered primary care for asthma for children and/or adults?

Question 2 (Construct IIB.5):
How has varying asthma care for children been described on the basis of characteristics of the PCP offices or practices? For example, are they characterized by the level of continuity between individual clinicians, the level of conntiuity with any provider in the practice, the accessibility of specified clinicians and/or practices during the day and/or after work hours, etc?

- **Question 2a (Construct IIIA.3):**

What do we know about the impact of variations in patterns of care/practice, use of treatment modalities, and/or receipt of well-child care on asthma management or outcomes (e.g. ED use)? How about the sufficiency of the number of WCC Visits (eg meets HEDIS standard or AAP standard or does not)? Absolute number of visits to PCP?

Question 3 (Construct IIC.7):
(Broad) Does involvement of multidisciplinary personnel (beyond allopathic or osteopathic physicians) impact quality of care, utilization of care and asthma outcomes both within context of a primary care practice or in other clinical settings?

- **Question 3a. (Construct IIIB.2):**

What evidence is there regarding use of supplemental services outside of regular clinical visits and how do these services impact quality and utilization of care?

The following poster describing this measure was submitted for peer review and accepted and presented at the Annual Research Meeting of AcademyHealth in 2014.

New Pediatric Quality Measures Program (PQMP) Measure of Emergency Department (ED) Use for Children with Asthma

LC Kleinman, LE Soloway, CJ Homer, A Vella, A Ting, N Massenburg, B Rabin Fastman, A Vachon, A Balbierz, M Manice, EA Barrow, L Pickering, E Shemesh *Mount Sinai Collaboration for Advancing Pediatric Quality Measures (CAPQuaM); Icahn School of Medicine at Mount Sinai; National Institute for Child Health Quality (NICHQ), Northeast Business Group on Health, and NY State Department of Health*

Icahn School of Medicine at Mount Sinai

OBJECTIVES

To describe a new asthma outcome measure developed for the federal Pediatric Quality Measures Program by the Collaboration for Advancing Pediatric Quality Measures (CAPQuaM), an AHRQ-CMS CHIPRA Center of Excellence.

To describe the approach CAPQuaM is using for measure development

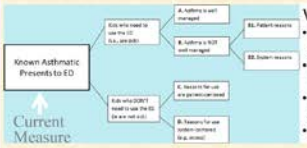
Consortium Partners include: AAP, AAFP, NICHQ, ACOG, CAHMI, NYS Medicaid, NCQA, Institute for Patient- and Family-Centered Care

CHIPRA = Child Health Insurance Program Reauthorization Act

METHODS

1. Qualitative Interviews with Clinicians/Patients
2. Literature Review
3. Criteria and Measure development
   - National multidisciplinary 9-person expert panel
   - 2 Round modified Delphi Process
   - Ratings of >250 clinical scenarios
     - •Inclusion/exclusion/reporting (59)
     - •What establishes ED as appropriate level of care (49)
     - •Establishing sufficiency of prior asthma care (59)
     - •Establishing root source of failures of prior management (34)
     - •Quality of ED Management (62)
   - Stakeholder Review and Input
4. Testing in NY State Medicaid Data

Overarching Conceptual Model



Who is a known asthmatic child?

- Prior asthma that health care plan should be able to identify
- Lower prevalence than asthma (~14-16% on Medicaid in NSCH)
- More prevalent than HEDIS persistent asthma (4.7% in NY State Medicaid)
- CAPQuaM Prevalence: 9.6%

Specifications: Assessing Eligibility and Scanning for Events Month by Month

Any prior hospitalization with asthma as primary or secondary diagnosis

Other Qualifying events after the fifth birthday (age is age at event):

One or more prior ambulatory visits with asthma as the primary diagnosis

(this criterion implies an asthma ED visit in the reporting month), **OR** Two or more ambulatory visits with asthma as a diagnosis, **OR**

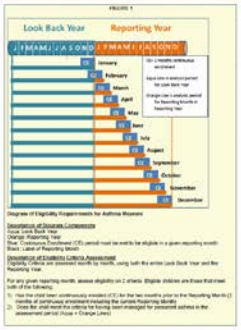
One ambulatory visit with asthma as a diagnosis **AND** at least one asthma related prescription, **OR**

Two or more ambulatory visits with a diagnosis of bronchitis

Other Qualifying events, any age:

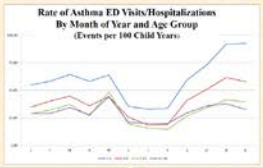
Three or more ambulatory visits with diagnosis of asthma or bronchitis, **OR** Two or more ambulatory visits with a diagnosis of asthma and/or bronchitis **AND** one or more asthma related prescriptions

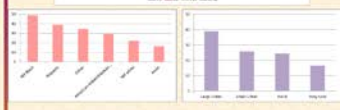
For eligibility purposes, asthma related medicine means long acting beta agonist (alone or in combination) or inhaled corticosteroid (alone or in combination), anti-asthmatic combinations, methylxanthines (alone or in combination), and/or mast cell stabilizers.



Numerator Events include Hospitalizations or ED Visits with Primary or Secondary Diagnosis of Asthma (most Medicaid systems would miss ED visits resulting in hospitalization if only ED visits sought)

Findings: NY Medicaid 2011





CONCLUSIONS

- CAPQuaM Asthma Measure is a true epidemiological rate of undesirable asthma outcomes (ED visits & hospitalizations)
- Measure performs well in NY State Medicaid data
- Captures variations, including racial

IMPLICATIONS

- CAPQuaM measures are now available for quality measurement and QI
- This measure improves upon currently available measures that assess undesirable utilization outcomes for children with asthma and should be adopted widely

FUNDING: AHRQ 1U18HS020518

Appendix D

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 #: 2816
Corresponding Measures:
Measure Title: Appropriateness of Emergency Department Visits for Children and Adolescents with Identifiable Asthma
Measure Steward: University Hospitals Cleveland Medical Center
Brief Description of Measure: This measure estimates the proportion of emergency department (ED) visits that meet criteria for the ED being the appropriate level of care, among all ED visits for identifiable asthma in children and adolescents.
Developer Rationale: Asthma is one of the most common indications for emergency department (ED) visits by children. (1-3) AHRQ's Healthcare Cost and Utilization Project (HCUP) data from the Nationwide Emergency Department Sample (NEDS) found that in 2012, children between 1 and 17 years old had more than 1,895,000 ED visits for asthma with almost 10% resulting in hospitalization.

Evidence suggests that ED visits and hospitalizations in children with asthma vary systematically by how well-equipped that community is to provide primary care, and by the quality of primary care delivered. (4, 5) There is widespread literature illustrating that ED visits and hospitalizations are each undesirable utilization outcomes from poorly managed asthma. There is not a large literature that assesses whether or not pediatric ED visits were appropriate. (6 -10)

A body of literature has explored the value and feasibility of measuring the appropriateness of medical activities using data available in the medical record. (11-14) Early work in adults included assessment of hysterectomy, carotid endarterectomy and cardiac interventions. An independent research project brought the construct of appropriateness to children (15), while Kleinman and colleagues were the first to assess the appropriateness of specific pediatric procedures. (16, 17) A later study demonstrated the feasibility of medical record data for such an assessment. (18) DeAngelis pioneered studies of what constitutes a good reason to use the ED. (6) All of these studies used a definition of appropriateness that compared benefit to likely risk without specific consideration of costs. The need for more studies looking for overuse was recently reviewed. (19) RAND type Delphi panels are accepted around the world as a method for developing criteria to assess appropriateness. (20-22)

Research demonstrates that:

•ED visits are an important issue for child health insurers, including Medicaid, with clinical and financial consequences;

•An overcrowded primary care system contributes to ED use for non-emergent and even non-urgent conditions.

•Pediatric hospitalizations for asthma vary by primary care availability and quality

•ED visits are common for children with asthma, including those in Medicaid

•Assessment of appropriateness using information in the medical record is a well-established and validated method that has been successfully applied to children.

The literature suggests that a measure that assesses whether or not the ED is an appropriate level of care for a child with asthma at the time that they present has intrinsic value. Such a measure would:

•Characterize the process of care in a way that assesses whether a particular ED visit represents overuse

•Allow the outcomes of asthma care to be better characterized in a manner that describes performance and promotes targeted improvement. Inappropriate ED visits represent failures of primary care delivery, availability and/or access. Appropriate visits may represent a failure to control asthma. These have distinct and distinguishable meanings that contribute to the understanding of the quality of asthma care.

•Measuring the quality of asthma care requires assessment of multiple factors. This appropriateness measure helps plans, purchasers, and society to understand the implication of asthma ED visits as outcomes of asthma care. The implications herein is that understanding what is better or worse care requires looking at various factors and not simply a higher or lower appropriateness score. The understanding of this measure is enhanced by considering whether the rate of undesirable outcomes (ED visits and hospitalizations) is high or low and whether other measures of primary care availability and access or asthma quality suggest high levels of performance or not..

An abstract describing the proposed measure was peer-reviewed and subsequently presented to a national audience at AcademyHealth 2014 Annual Research Meeting in San Diego in the “Measuring the Safety, Quality, and Value” section. Feedback was positive regarding the methods, measures, ethics, and importance of this measure.

Research evidence supports the importance and need for our proposed measure that assesses whether the ED represents an appropriate level of care for children with asthma who are seen in the ED.

1.Kharbanda, A.B., et al., Variation in resource utilization across a national sample of pediatric emergency departments. J Pediatr, 2013. 163(1): p. 230-6.
2.Adams, J.G., Emergency department overuse: Perceptions and solutions. JAMA, 2013. 309(11): p. 1173-1174.
3.Institute, N.E.H., A Matter of Urgency: Reducing Emergency Department Overuse. Research Brief, 2010(March).
4.Perrin, J.M., et al., Variations in rates of hospitalization of children in three urban communities. N Engl J Med, 1989. 320(18): p. 1183-7.
5.Perrin, J.M., et al., Primary care involvement among hospitalized children. Arch Pediatr Adolesc Med, 1996. 150(5): p. 479-86.
6.DeAngelis, C., P. Fosarelli, and A.K. Duggan, Use of the emergency department by children enrolled in a primary care clinic. Pediatr Emerg Care, 1985. 1(2): p.61-5.
7.Berns, S.D., et al., Appropriate use of a pediatric emergency department: is the pediatrician called before the visit? Pediatr Emerg Care, 1994. 10(1): p. 13-7.
8.Rudowitz, R., A Look At CBO Projections For Medicaid and CHIP, in The Kaiser Commission on Medicaid and the Uninsured. 2014, The Henry J. Kaiser Family Foundation
9.Taubman, S.L., et al., Medicaid Increases Emergency-Department Use: Evidence from Oregon´s Health Insurance Experiment. Science, 2014. 343(6168): p. 263-268.
10.Smulowitz, P.B., et al., Increased Use of the Emergency Department After Health Care Reform in Massachusetts. Ann Emerg Med, 2014.
11.Brook, R.H., et al., A method for the detailed assessment of the appropriateness of medical technologies. Int J Technol Assess Health Care, 1986. 2(1): p. 53-63.
12.Park, R.E., et al., Physician ratings of appropriate indications for six medical and surgical procedures. Am J Public Health, 1986. 76(7): p. 766-72.
13.Fitch, K., et al., The RAND/UCLA Appropriateness Method User´s Manual. 2001 RAND.
14.Kosecoff, J., et al., The appropriateness of using a medical procedure. Is information in the medical record valid? Med Care, 1987. 25(3): p. 196-201.
15.Kemper, K.J., Medically inappropriate hospital use in a pediatric population. N Engl J Med, 1988. 318(16): p. 1033-7.
16.Kleinman, L.C., et al., The medical appropriateness of tympanostomy tubes proposed for children younger than 16 years in the United States. Jama, 1994. 271(16): p. 1250-5.
17.Kleinman, L.C., E.A. Boyd, and J.C. Heritage, Adherence to prescribed explicit criteria during utilization review. An analysis of communications between attending and reviewing physicians. Jama, 1997. 278(6): p. 497-501.
18.Keyhani, S., et al., Electronic health record components and the quality of care. Med Care, 2008. 46(12): p. 1267-72.
19.Keyhani, S. and A.L. Siu, The underuse of overuse research. Health Serv Res, 2008. 43(6): p. 1923-30.

20.Bernstein, S.J., et al., The appropriateness of hysterectomy. A comparison of care in seven health plans. Health Maintenance Organization Quality of Care Consortium. Jama, 1993. 269(18): p. 2398-402.
21.Taylor, A.J., et al., ACCF/SCCT/ACR/AHA/ASE/ASNC/NASCI/SCAI/SCMR 2010 Appropriate Use Criteria for Cardiac Computed Tomography. A Report of the American College of Cardiology Foundation Appropriate Use Criteria Task Force, the Society of Cardiovascular Computed Tomography, the American College of Radiology, the American Heart Association, the American Society of Echocardiography, the American Society of Nuclear Cardiology, the North American Society for Cardiovascular Imaging, the Society for Cardiovascular Angiography and Interventions, and the Society for Cardiovascular Magnetic Resonance. J Cardiovasc Comput Tomogr, 2010. 4(6): p. 407.e1-33.
22.Basger, B.J., T.F. Chen, and R.J. Moles, Validation of prescribing appropriateness criteria for older Australians using the RAND/UCLA appropriateness method. BMJ Open, 2012. 2(5).

Numerator Statement: The numerator is the number of eligible asthma ED visits in the random sample that also satisfy at least one of the explicit criteria to indicate that the ED is an appropriate level of care. Distinct numerators are reported for children ages 2-5, 6-11, 12-18, and optionally, 19 - 21.
Denominator Statement: The denominator is a random sample of the patients in each age stratum who have visited the emergency department for asthma (as a first or second diagnosis) and meet the specified criteria for having identifiable asthma (defined in s2b).

Separate numerators and denominators are reported for children age 2-5, 6-11, 12-18, and, optionally, 19-21 years. An overall rate across strata is not reported.
Denominator Exclusions: ED visits that are already in the sample OR Children that fall outside of specified age range of 2-21 OR who do not meet time enrollment criteria OR do not meet identifiable asthma prior to the ED visit, OR children with concurrent or pre-existing COPD, Cystic Fibrosis or Emphysema.

At the discretion of the accountability entity, the denominator may be restricted to children 2-18.

Measure Type: Outcome
Data Source: Claims (Only), EHRs Hybrid, Paper Records
Level of Analysis: Facility, Health Plan

New Measure -- Preliminary Analysis

Criteria 1: Importance to Measure and Report

1a. Evidence

1a. Evidence. The evidence requirements for a health outcomes measure include providing rationale that supports the relationship of the health outcome to processes or structures of care. The guidance for evaluating the clinical evidence asks if the relationship between the measured health outcome and at least one clinical action is identified and supported by the stated rationale.

Evidence Summary

This measure was previously reviewed by the Pulmonary Standing Committee (March 2016) as a process measure. Based on feedback from that Committee, the developer has revised and resubmitted the measure as an outcome measure. In the measure's evidence and testing forms, content submitted previously is in black; new information is in blue.

- The developer provides a diagram of the relationship between processes of care and outcomes.
- The developer states "Low levels of appropriateness suggest fewer breakthrough episodes of asthma and hence better quality of asthma care for those who receive it. If the rate of asthma ED visits is high and the rate of appropriateness is low this suggests both high quality care for those receive asthma care and insufficient access/availability of such care. High levels of appropriateness suggest both efficient resource use of the emergency department and that ED visits are a proxy for clinical outcomes since many of the visits represent

breakthrough asthma. High levels of appropriateness combined with a low rate of ED asthma use suggests both efficient use of resources and good asthma outcomes."

- The developer added citations for clinical practice guidelines from the National Heart, Lung, and Blood Institute (NHLBI), National Institutes of Health (2007): "As a general rule, patients with well-controlled asthma should have: ... no emergency department visits; no hospital stays ...". Grade C = Nonrandomized trials and observational studies. Evidence is from outcomes of uncontrolled or nonrandomized trials or from observational studies.
- The developer's criteria for appropriateness are listed as "explicit criteria" in the numerator details:
    - **Disposition of the ED visit was admission to the hospital, OR**
    - **Documented physical findings consistent with respiratory distress, including any of the following:** Labored breathing (including moderate or severe increased work of breathing); OR Retractions, grunting, and/or evidence of accessory muscle use; OR Markedly decreased breath sounds; **OR**
    - **Recorded oxygen saturation below 90%; OR**
    - **An arterial blood gas (ABG) was obtained in the emergency department; OR**
    - **The child had a consultation with a pulmonologist or asthma specialist that was ordered and provided in the ED; OR**
    - **There is clear documentation that prior to arrival in the ED any of the following occurred:** The child was referred to the ED after evaluation by the PCP or other clinician. The evaluation may include an in person visit or auscultation including via telephone OR The child received two or more doses of inhaled rescue medications without sufficient clinical improvement. Documentation of parent report meets the criterion. OR The child was assessed with an objective instrument such as a peak flow meter and was found to be in a pre-defined "red zone" of peak flow measurement as part of an asthma action or similar plan. Documentation requires ALL of the following: a written asthma action plan exists AND defines a "red zone" for which urgent assessment by a clinician is indicated AND an objective assessment was made and its result was in the pre-defined red zone. Documentation of parent report meets the criterion.
- NQF provides specific guidance on evaluating appropriate use measures, as follows:
    - "If there is no empiric evidence, skip Box 10 and go to Box 11. The Committee should agree that the AUC method is a systematic assessment of expert opinion that the benefits of what is being measured outweigh the potential harms (Box 11). If the Committee agrees that it is acceptable (or beneficial) to hold providers accountable for the performance in the absence of empiric evidence (Box 12), then rate as "insufficient evidence with exception."

Questions for the Committee:

- *Does the rationale provided by the developer support a relationship between appropriateness of ED visits and health outcomes?*
- *Is there at least one thing that the provider can do to achieve a change in the measure results?*

Guidance from the Evidence Algorithm: Outcome measure (Box 1) → Relationship between health outcome and provider action (Box 2) → PASS

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

Previous review: Submitted as process measure, did not pass Evidence - H-0; M-2; L-9; I-9

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

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

- The developer states, "asthma is one of the most common indications for emergency department (ED) visits by children". The developer reports on AHRQ data from 2012 indicating "children between 1 and 17 years old had more than 1,895,000 ED visits for asthma with almost 10% resulting in hospitalization." Further, they note that "evidence suggests that ED visits and hospitalizations in children with asthma vary systematically by how well-equipped that community is to provide primary care, and by the quality of primary care delivered. There is widespread literature illustrating that ED visits and hospitalizations are each undesirable utilization outcomes from poorly managed asthma."
- The developer reports the following results (from testing data) identified statistically significant differences between groups at specified levels, e.g., age groups, among racial/ethnic groups, and within age group among racial/ethnic groups:
    - 181 of 335 (54.3%) ED visits were deemed appropriate for children 2 to 5 years
    - 209 of 447 (43.8%) ED visits were appropriate for children 6 to 11 years
    - 165 of 341 (48.4%) visits were appropriate for children 12 to 18 years

Disparities

- The developer states, "Pediatric asthma is more prevalent in minority populations. Lifetime prevalence rates of asthma in Hispanic and African American children are 12.4% and 15.8% respectively."
- Based on its chart audits, the developer reports performance on the measure varies by race/ethnicity and that a Chi-square analysis confirms the differences are statistically significant. For example, Hispanic children had higher rates of questionable use of the ED (55.9% of visits) when compared to non-Hispanic children (47.8%), p=0.002. African American children "showed a trend" toward more questionable use compare to all other children (53.6% vs. 48.7%, p=0.10).
- The developer reports performance on the measure varies by insurance status and that a Chi-square analysis confirms the differences are statistically significant. The appropriate use rates were: Medicaid patients (46.3%); private (59%); uninsured patients (38.6%); other forms of insurance (military and worker's comp) (55.0%), p=0.005.

Question for the Committee:

- *Does the Committee believe there is a gap in care that warrants a national performance measure?*

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

Committee pre-evaluation comments

Criteria 1: Importance to Measure and Report (including 1a, 1b, 1c)

- I am concerned on multiple levels that this measure doesn't really look at appropriateness (appropriate use versus overuse and underuse) but rather focuses on overuse. The denominator is defined by patients who make it to the ER, but what about those that are not appropriately referred? The conceptual model also is based on ER use and appropriateness, but is not captured by the measure as proposed. Moreover, the measure assumes physician referral to the ER is defacto appropriate--I don't see quality evidence supporting this supposition. Other criteria for appropriateness beg quibbling: consultation, ABG (maybe a standing order), decreased breath sounds (reliable?), etc. Furthermore, the use of the second level diagnosis concerns me as it may capture ER visits with nothing to do with asthma. The causal pathway here seems fraught with confounders. And the measure as proposed inadequate to determine appropriate use. I do not support an exception.
- Substantial evidence is provided that demonstrates the improvement in overall health for these chronically ill patients if appropriate asthma visits are reduced as they are usually patients in crisis. As an outcome measure we need to assess whether it is actionable for the unit of measurement. The developers say it can be used by hospitals and insurers/health plans. Further evidence should be provided on how this is actionable at the hospital level as the primary interventions would be associated with PCPs, specialists and urgent care facilities. For a health plan, this data will be sampled from multiple EDs yet this is not part of the sampling algorithm. This can be actionable by the insurer as they control the network.
- I would agree there is a gap in care, and certainly evidence for disparities driving outcomes.
- Most of the rates reported were around the 40-50% range which indicates substantial room for improvement.

Criteria 2: Scientific Acceptability of Measure Properties

2a. Reliability

2a1. Reliability Specifications

2a1. Specifications requires the measure, as specified, to produce consistent (reliable) and credible (valid) results about the quality of care when implemented.

Data source(s): Administrative claims, Electronic Clinical Data: Electronic Health Record, Paper Medical Records

Specifications:

- The level of analysis is facility and health plan; the care setting is ED, hospital.
- Interpretation of score: Better quality = Higher score
- The numerator for this measure is: *The numerator is the number of eligible asthma ED visits in the random sample that also satisfy at least one of the explicit criteria to indicate that the ED is an appropriate level of care. Distinct numerators are reported for children ages 2-5, 6-11, 12-18, and optionally, 19 - 21.*
- ***NQF Note:***
  - The listed "explicit criteria" in the specifications (numerator details) are:
    - **Disposition of the ED visit was admission to the hospital, OR**
    - **Documented physical findings consistent with respiratory distress, including any of the following:** Labored breathing (including moderate or severe increased work of breathing); OR Retractions, grunting, and/or evidence of accessory muscle use; OR Markedly decreased breath sounds; **OR**
    - **Recorded oxygen saturation below 90%; OR**
    - **An arterial blood gas (ABG) was obtained in the emergency department; OR**
    - **The child had a consultation with a pulmonologist or asthma specialist that was ordered and provided in the ED; OR**
    - **There is clear documentation that prior to arrival in the ED any of the following occurred:** The child was referred to the ED after evaluation by the PCP or other clinician. The evaluation may include an in person visit or auscultation including via telephone OR The child received two or more doses of inhaled rescue medications without sufficient clinical improvement. Documentation of parent report meets the criterion. OR The child was assessed with an objective instrument such as a peak flow meter and was found to be in a pre-defined "red zone" of peak flow measurement as part of an asthma action or similar plan. Documentation requires ALL of the following: a written asthma action plan exists AND defines a "red zone" for which urgent assessment by a clinician is indicated AND an objective assessment was made and its result was in the pre-defined red zone. Documentation of parent report meets the criterion.
- The denominator for this measure is: *The denominator is a random sample of the patients in each age stratum who have visited the emergency department for asthma (as a first or second diagnosis) and meet the specified criteria for having identifiable asthma (defined in s2b).*
  Separate numerators and denominators are reported for children age 2-5, 6-11, 12-18, and, optionally, 19-21 years. ***An overall rate across strata is not reported*** [our emphasis].
- The exclusions for the measure are: *ED visits that are already in the sample OR Children that fall outside of specified age range of 2-21 OR who do not meet time enrollment criteria OR do not meet identifiable asthma prior to the ED visit, OR children with concurrent or pre-existing COPD, Cystic Fibrosis or Emphysema.*
  At the discretion of the accountability entity, the denominator may be restricted to children 2-18.
- The ICD-9 and ICD-10 codes are in an appendix.
- The calculation algorithm is stated in S.14.
- Sampling is allowed. At least 500 children per strata should be included in the samples.
- One data source is pharmacy claims, but the developer acknowledges that availability will vary.
- The measure requires stratification by three age groups, as just noted.

Questions for the Committee:

- *Are all the data elements clearly defined? Are the appropriate codes included in the ICD-9 to ICD-10 conversion?*
- *Is the calculation algorithm clear?*
- *Is the potential variability in access to/inclusion of pharmacy data a concern?*
- *Is it a concern that the measure does not report an overall rate?*
- *Is it likely this measure can be consistently implemented?*

2a2. Reliability Testing Testing attachment

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.

SUMMARY OF TESTING

Reliability testing level ☐ **Measure score** ☒ **Data element** ☐ **Both**

Reliability testing performed with the data source and level of analysis indicated for this measure ☒ **Yes** ☐ **No**

Method(s) of reliability testing

- The developer relied on pre-existing data element-level validity testing in the literature to identify children who are being managed for identifiable asthma (denominator), which is permitted by NQF testing guidance.
- For the numerator, the developer did not conduct empirical reliability testing. Instead, it relied on empirical validity testing at the data element level (chart abstraction compare to an authoritative source).
- Per NQF guidance, separate reliability testing is not required if data element-level validity testing is performed.

Results of reliability testing

- Not applicable; see discussion on validity testing at the data element level

Question for the Committee:

o *Do the results demonstrate sufficient reliability so that differences in performance can be identified?*

Guidance from the Reliability Algorithm Not Applicable (rating from data-element validity will apply; highest eligible rating is MODERATE)

Precise specifications (Box 1) → Empirical reliability testing conducted using statistical tests (Box 2) → Empirical validity testing of patient-level data conducted (Box 3) → Insufficient

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

Rationale: See rationale for validity.

2b. Validity

2b1. Validity: Specifications

2b1. Validity Specifications. This section should determine if the measure specifications are consistent with the evidence.

Specifications consistent with evidence in 1a. ☒ **Yes** ☐ **Somewhat** ☐ **No**

Question for the Committee:

o *Are the specifications consistent with the evidence?*

2b2. Validity testing

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.

SUMMARY OF TESTING

Validity testing level ☐ **Measure score** ☒ **Data element testing against a gold standard** ☐ **Both**

Method of validity testing of the measure score:

- ☐ **Face validity only**
- ☐ **Empirical validity testing of the measure score**

Validity testing method:

- The developer reports data element level validity testing. The developer relies, as is permitted by NQF guidance, on other sources for denominator data element validity and conducted empirical testing on the numerator data elements in one facility.

For the denominator:

- The developer relies on literature to support its conclusion of the validity of administrative data elements to identify children who are being managed with identifiable asthma. Per NQF policy:
  - Prior evidence of validity of data elements can be used, including published data, provided it includes the same data elements; uses the same data type; and is conducted on an appropriate sample (i.e., representative, adequate numbers, etc.)
  - The developer attests that the data elements match those assessed in the literature.
- The developer also cites score-level validity testing of two previously-endorsed asthma measures as evidence of data-element level validity. However, this does not meet NQF's requirements for demonstration of score-level validity.
- The developer used NY State Medicaid Managed Care claims data for its analyses.

For the numerator:

- For validity testing at the data element level, three reviewers each looked at 10 charts from one facility, assessing the presence of 6 constructs and an overall visit-level assessment of appropriateness. The developer states it conducted testing at the beginning of data collection and again at the conclusion of data collection.
- The following six numerator appropriateness criteria were tested:
  - Retractions
  - Accessory Muscle Use
  - Markedly diminished BS
  - Hospitalized from ED
  - O2 sat < 90%
  - Referred by PCC

Validity testing results:

Denominator

- For the results of the literature review:
  - The developer attests that the data elements match those assessed in the literature.
  - The sensitivities, specificities, PPVs, kappas, etc., were generally strong, including:
    - Wilchesky et al., asthma diagnosis in in-patient setting: Sp= 96.76 (95%CI 96.5, 97.0)
    - Folwes et al., asthma diagnosis in ambulatory: Sensitivity and specificity was 0.82 and 0.99, respectively. Sensitivity of .82 using claims was higher than sensitivity using self-report at 0.64
    - Wilchesky, et al., asthma diagnosis in clinic/outpatient setting: Asthma claims were highly specific, Sp= 96.76 (95%CI 96.5, 97.0).
- The developer also provides information from various articles related to the use of administrative data for identification of asthma and use of claims data for performance measurement.
  - Age: According to CMS MMIS data requirements, "States are required to submit validated claims data including age or date of birth with a tolerance of 0.1%".

- Asthma diagnosis: In an in-patient/ED setting, "Asthma claims were highly specific, Sp= 96.76 (95%CI 96.5, 97.0)." (Wilchesky, et al)
- In an ambulatory setting, "Sensitivity and specificity was 0.82 and 0.99, respectively. Sensitivity of .82 using claims was higher than sensitivity using self-report at .64" (Folwes, et al)
- In a clinic/outpatient setting, "Asthma claims were highly specific, Sp= 96.76 (95%CI 96.5, 97.0)." (Wilchesky, et al)
- Exclusions: for diagnoses of COPD, cystic fibrosis, emphysema, "Claims had a PPV of 91.9, and a negative predictive value of 92.6, with k of 0.65 (substantial agreement) compared to chart review for chronic pulmonary disease. ICD 10 performed similarly in this study" (Quan et al)

Numerator

- For the six constructs, three reviewers each reviewed 10 charts early in training and at the end of the data collection period. The developer reports this resulted in 180 comparisons with the trainer (6 clinical constructs * 3 * 10 = 180). The table also provides an "all six combined" kappa, but it is unclear why this is lower than the overall assessment and its implications for reliability. NQF staff has requested additional information from the developer.
- For the assessment of overall appropriateness (i.e., a separate assessment and the numerator), the kappas were 0.77, increasing to 0.87 after training.

The developer states, per the Landis and Koch classification, a kappa value of 0.87 indicates almost perfect agreement. **Construct**	**Agreement**	
	Initial Kappa	**Final Kappa**
1. Retractions	0.67	0.87
2. Accessory Muscle Use	0.44	0.89
3. Markedly diminished BS	0.71	0.78
4. Hospitalized from ED	1.0	1.0
5. O2 sat < 90%	0.79	NA*
6. Referred by PCC	1.0	NA*
All six combined	0.76	0.68
Overall: Appropriateness	0.77	0.87
* NA is because there was no variability in the charts reviewed. There was no disagreement in any of the assessments.		

- The explicit criteria for the numerator specifications include several elements for which results are not reported—e.g., "arterial blood gas (ABG) in the emergency department".
  - In an email to NQF staff, the developer stated it tested all elements, but "only reported findings from the criteria that were found to be pertinent within our tested institution. Because these appropriateness criteria are written for implementation nationally, and there is variation among hospitals protocols/procedures done in the ED, we included all the appropriateness criteria."
  - The developer further stated that at the institution where the measure was tested, "ABG ordered/obtained were not reported because ABGs are not ordered in the ED [at this institution], and rely on the PulseOx O2 Saturation level instead. But the criteria included in the specifications are written for national implementation."

Questions for the Committee:

- *Is the test sample size adequate to generalize for widespread implementation?*
- *Should all numerator appropriateness criteria be tested for data-element level validity?*
- *Do the results demonstrate sufficient validity so that conclusions about quality can be made?*
- *Do you agree that the score from this measure as specified is an indicator of quality?*

2b3-2b7. Threats to Validity

2b3. Exclusions:

The developer provides the following:

- There are no numerator exclusions.
- Denominator exclusions include: Children with concurrent or pre-existing Chronic Obstructive Pulmonary Disease (COPD) diagnosis, cystic fibrosis diagnosis, or emphysema diagnosis.
- The developer reports <=2.5% potentially eligible children were excluded by these clinical diagnoses.
- The developer reports that exclusions are clinical and represent construct validity rather than statistical considerations.
- The measure also excludes children who have not been consecutively enrolled in the reporting plan for at least two months prior to the index reporting month, as well as the index reporting month itself, but they note that 20% more children are included than would be if they had a 12-month enrollment requirement.

Questions for the Committee:

- *Are any patients or patient groups inappropriately excluded from the measure?*
- *Are the exclusions/exceptions of sufficient frequency and variation across providers to be needed (and outweigh the data collection burden)?*

2b4. Risk adjustment: **Risk-adjustment method** ☐ **None** ☐ **Statistical model** ☒ **Stratification**

Conceptual rationale for SDS factors included? ☐ **Yes** ☒ **No**

SDS factors included in risk model? ☒ **Yes** ☐ **No**

Risk adjustment summary

- An overall rate is not reported; the measure is reported by age stratifications.
- This measure is adjusted for age group only (ages 2-5, 6-11, 12-18, and optionally, 19-21). The developer has specified stratified analysis for risk-adjustment rather than using a statistical risk modeling approach.
- The developer relies on the NIH NHLBI NAEPP guideline assertion that the goals for asthma severity, control, and responsiveness are identical for all levels of baseline asthma severity as the rationale not to risk-adjust for severity.
- The developer states that additional stratifications are optional, but notes that these were not included as risk adjustment factors due to lack of "*clear biological evidence that ED visits should be more likely in any of the sub categories*".

Conceptual analysis of the need for SDS adjustment:

- Although the developer noted that its funders asked them to consider SDS factors, and it in fact did find patient-level differences for at least some, the developer did not discuss the conceptual rationale of why or how SDS factors (e.g., race/ethnicity, poverty level in the caregivers county of residence, rurality/urbanicity on the caregiver's county of residence, insurance type and plan type) might be associated with appropriateness of ED visits for asthma.

Empirical analysis of SDS factors:

- The developer reported statistically significant differences in appropriateness of ED visits at the patient level by race and ethnicity, as well as for sex and insurance status subgroups. It did not, however, compare rates of appropriateness with and without inclusion of these variables as a risk-adjustment approach, as requested by NQF.

Questions for the Committee:

- *A justification for no risk adjustment for SDS factors is provided. Do you agree with the developer's rationale that there is no conceptual basis for adjusting this measure for SDS factors? Is there evidence that contradicts the developer's rationale?*
- *Do you agree with the developer's decision not to adjust for severity?*

2b5. Meaningful difference *(can statistically significant and clinically/practically meaningful differences in performance measure scores can be identified)*:

- In its testing data, the developer tested the numerator at 1 facility, so cannot demonstrate meaningful differences at the proposed facility level of analysis.
- The developer also analyzed claims data from the New York State Medicaid Managed Care data (including claims from all MCOs that are contracted for Medicaid care), so could not analyze differences among (for example) state Medicaid programs. The developer does not examine differences among MCOs within the data plan.
    - The developer states its analyses found meaningful differences by age groups and statistically significant differences by race/ethnicity and insurance status, which the developer states means "the measure distinguishes signal from noise". The developer posits that this demonstrates the measure detects meaningful differences. NQF's requirement is that testing demonstrate differences among measured entities.

Question for the Committee:

- *Does this measure identify meaningful differences in quality?*

2b6. Comparability of data sources/methods:

n/a

2b7. Missing Data

- The developer does not account for missing data. It cites literature that chart review is an accurate method of identifying the level of appropriateness of a clinical service. Failure to document is a "quality deficit" that the developer does not consider as missing data.
- Use of pharmacy data is on an "if available" basis to identify children with asthma for the denominator; the developer notes any results reported without should be marked as such. The developer reports use of pharmaceutical data expanded the pool by approximately 10,000 children (from 180,000 to 190,000—5.5%). The developer states it "found no evidence this was a threat to validity," but does not provide analyses that the scores with the pharmacy data did not differ from the scores when pharmacy data were excluded. The developer does not have direct access to the data to provide additional analyses at this time.

Question for the Committee:

- *Is the variable use of pharmacy data a threat to validity?*

Guidance from the Validity Algorithm: Specifications consistent with evidence (Box 1) → Potential threats to validity assessed (Box 2)→ Insufficient

The highest possible rating is INSUFFICIENT.

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

RATIONALE: The numerator was only tested at one facility with 10 charts; the measure does not demonstrate meaningful differences. All numerator details/data elements were not tested (e.g., ABG was not available at the institution at which testing was conducted, but its reliability/kappa should be assessed at an institution for which this is policy/practice). Additionally, there is insufficient information for the Committee to discuss whether SDS factors should

or should not be included (i.e., analyses on scoring with and without factors included). Finally, there are insufficient data on the effect on the score of missing pharmacy data.

Committee pre-evaluation comments

Criteria 2: Scientific Acceptability of Measure Properties (including all 2a, 2b, and 2d)

- I am worried that there are substantial challenges in collecting this data consistently, and potential for elements such as pharmacy data systematically influencing outcomes. The outcome measure itself is really not useful as a quality measure without knowing utilization overall.
- This is being presented as an Outcome Measure. Substantial evidence is provided concerning the denominator definition and the codes seem appropriate. The numerator was determined by the RAND/UCLA Delphi method with a panel of experts. If appropriately used this is considered a best approach to applying the Delphi method. Even though its primarily for designing survey questions it is still appropriate for determining the list of indicators for inappropriate use. A clinician should address the appropriateness of the panel measures chosen
- The developers say this will be done like a HEDIS hybrid measure where a stratified sample will be used to generate members of the denominator and the claims and chart review will determine the numerator. Will discuss the validity of the numerator method in that section but it was clear that even in this one hospital pilot, some of the numerator indicators are not collected. The developers say that is okay as the collected indicators are enough. Hard to judge without multiple institutions.
- The algorithm is clear for sampling the denominator (except for exact sample sizes) and it is the “at least one indication” for the numerator.
- The measure is reported stratified by age based on data provided that show age variability on this rate exists. This is appropriate. There is no data to support consistent implementation though the steps to determine the rate are well specified.
- There was no specific reliability testing done. However acceptable validity results on the denominator were provided. Validity testing of the numerator was done at the item level comparing chart to an authoritative source with mixed results.
- I guess according to NQF criteria it is sufficient, although I am very concerned that putting apples, oranges, and pineapples together gets you fruit cocktail...
- There was no specific reliability testing done. However acceptable validity results on the denominator were provided. Validity testing of the numerator was done at the item level comparing chart to an authoritative source with mixed results. Only 30 charts were reviewed at one institution which is insufficient to generalize the results. Some items could not be assessed as they were not recorded in the charts at this institution which makes item level testing insufficient. Kappas were okay for some items but not others. They improved over time but does that imply that the measure should be “practiced” the first year and not used until year 2?The comparison of chart to authoritative source could be assessing: Accuracy of chart data or The level of documentation detail provided by the clinician or The Quality of the data abstractors. No assessment was done comparing the charts to the claims data to assess whether there may be inaccuracies in the claims data. This can be considered a measure of quality but not necessarily at the hospital or health plan level. Rather it is an outcome that can be used for interventions at the provider level. It should be noted that they surveyed nine other institutions to see if they think any of the chart abstraction or item specifications would be problematic and results were consistently positive for successful data collection.
- I am concerned about the pharmacy factor, and also varying availability of data within the specification such as ABG. The testing is not very comforting...
- "Exclusion criteria seems acceptable and should not be burdensome as these can be identified from the claims data. The reasoning for no risk adjustment provided by the developers is sound if this measure is to be used for Quality Improvement as opposed to benchmarking (typically HEDIS measures are for both). QI interventions should be done based on the true population of the hospital or insurer. Interventions might be different in a poorer region vs one that is not or in an area with a large minority population. If benchmarking is also a goal then a risk adjusted rate should be calculated as well given the differences in rates by subgroups that was presented. Meaningful differences across institutions could not be adequately assessed based on the evidence provided. This is critical if the rates will be used for benchmarking but less so if for local quality improvement initiatives. Missing pharmacy data was not adequately addressed as no empirical data showing lack of impact was provided.

Criterion 3. Feasibility

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.

- The developer reports some data elements are in defined fields in electronic sources.
- The developer reports there are no fees.
- The measure requires chart abstraction, and testing found that training improves data element level validity/reliability. The developer notes that chart review is a reliable and accepted method of measuring appropriate use. No information is provided on the minimum number of charts that should be assessed.

Questions for the Committee:

- *Are the required data elements routinely generated and used during care delivery?*
- *Is the data collection strategy ready to be put into operational use?*
- *What training is available, how is it accessed, and what are the costs associated with training a clinician to identify an event? Is there additional burden for that physician?*
- *Can trained nurses or nurse practitioners review records?*

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

Committee pre-evaluation comments

Criteria 3: Feasibility

- I think this is feasible but very burdensome.
- Given the similarity to HEDIS measures, the data collection should be feasible given the needed training. Trained nurse or nurse practitioners should be able to do the chart review and there should not be additional burden for the physician unless the hospital adds fields to the EMR to improve validity of the rate.

Criterion 4: Usability and Use

4. Usability and 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.

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

- The developer plans to assist in the implementation of this measure following NQF endorsement. The developer notes multiple stakeholders are interested in using the measure. It has been approved for inclusion in the National Quality Measures Clearinghouse.

Improvement results

- This measure is not in use and therefore, has no improvement results.

Potential harms

- The developer reports no unintended consequences were observed during testing.
- The developer notes that the measure has a lower risk for gaming than some other measures because both low and high results can demonstrate different areas for improvement.

Vetting of the measure

- N/A

Feedback:

- No feedback provided on QPS. MAP has not reviewed this measure for inclusion in any federal program.

***Questions for the Committee**:*

- *Can the performance results be used to further the goal of high-quality, efficient healthcare?*
- *Do the benefits of the measure outweigh any potential unintended consequences?*

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

Committee pre-evaluation comments

Criteria 4: Usability and Use

- I am not sold on the link of this measure to quality as noted many times above.
- As designed, this measure can be used to assess improvements in the care these asthma patients.

Criterion 5: Related and Competing Measures

Related or competing measures

The developer did not include information on any of the related or competing measures. However, NQF staff identified the following measures that may be related and/or competing.

- 0047: Asthma: Pharmacologic Therapy for Persistent Asthma
- 0728: Asthma Admission Rate (PDI 14)
- 1800: Asthma Medication Ratio
- 2414: Pediatric Lower Respiratory Infection Readmission Measure
- 3189: Rate of Emergency Department Visit Use for Children Managed for Identifiable Asthma: Visits per 100 Child-years (*submitted by the same developer for review in this project)*

Harmonization

No information available.

Endorsement + Designation

The "Endorsement +" designation identifies measures that exceed NQF's endorsement criteria in several key areas. After a Committee recommends a measure for endorsement, it will then consider whether the measure also meets the "Endorsement +" criteria.

This measure is a candidate for the "Endorsement +" designation IF the Committee determines that it: meets evidence for measure focus without an exception; is reliable, as demonstrated by score-level testing; is valid, as demonstrated by score-level testing (not via face validity only); and has been vetted by those being measured or other users.

Eligible for Endorsement + designation: ☐ **Yes** ☒ **No**

RATIONALE IF NOT ELIGIBLE:

The measure has not been vetted.

Pre-meeting public and member comments

- None

NATIONAL QUALITY FORUM—Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:** Click here to enter NQF number
Measure Title: Appropriateness of Emergency Department Visits for Children and Adolescents with Identifiable Asthma
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: 12/9/2016

Instructions

- *Complete 1a.1 and 1a.12 for all measures.*
- *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:

- Health outcome: [3] a rationale supports the relationship of the health outcome to processes or structures of care. Applies to patient-reported outcomes (PRO), including health-related quality of life/functional status, symptom/symptom burden, experience with care, health-related behavior.
- 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.

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 U.S. Preventive Services Task Force (USPSTF) grading definitions and methods, or Grading of Recommendations, Assessment, Development and Evaluation (GRADE) guidelines.

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

☒ Health outcome: Appropriateness of Emergency Department Visits for Children and Adolescents with Identifiable Asthma

☐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

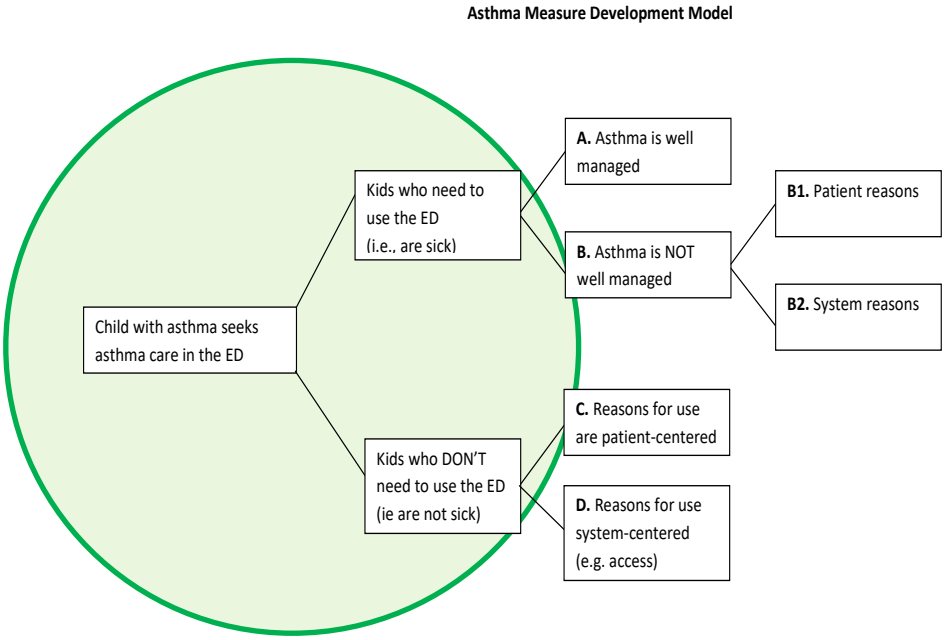
☐ Process: Click here to name what is being measured

☐ 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.12 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.



The green circle highlights the aspects of the conceptual model incorporated into this measure.
Underlying this model is a simple framework:

1. Accessible high quality primary care reduces the need for ED visits by decreasing the number of children who have acute breakthrough episodes requiring the ED. (NHLBI Guideline idenitifies factors with Evidence Levels A, B, and C).
2. Accessible high quality primary care reduces the need for ED visits by decreasing the number of children who come to the ED for asthma care better performed in the office setting.
3. Some children in the ED need to be there. Of those, some episodes were potentially preventable and others were not.
4. Our focus groups highlighted that some parents are comforted by the setting of the ED when they are caring for what they perceive as a vulnerable child. Parent perspectives do not adhere to system perspectives regarding a more strict hierarchy of what care belongs where.

Low levels of appropriateness suggest fewer breakthrough episodes of asthma and hence better quality of asthma care for those who receive it. If the rate of asthma ED visits is high and the rate of appropriateness is low this suggests both high quality care for those receive asthma care and insufficient access/availability of such care.

High levels of appropriateness suggest both efficient resource use of the emergency department and that ED visits are a proxy for clinical outcomes since many of the visits represent breakthrough asthma. High levels of appropriateness combined with a low rate of ED asthma use suggests both efficient use of resources and good asthma outcomes.

A seminal article that began to link primary care services to outcomes at a community or population level and that supports our interpretation appeared in the New England Journal of Medicine in 1989. (N Engl J Med. 1989 May 4;320(18):1183-7. Variations in rates of hospitalization of children in three urban communities.
Perrin JM1, Homer CJ, Berwick DM, Woolf AD, Freeman JL, Wennberg JE.).

Low levels of appropriateness suggest better asthma outcomes with less efficiency of primary care.
High levels of appropriateness suggest more efficient primary care and worse asthma outcomes.

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

1a.2 FOR OUTCOME MEASURES including PATIENT REPORTED OUTCOMES- State the rationale supporting the relationship between the health outcome (or PRO) to at least one healthcare structure, process (e.g., intervention, or service).

1a.3. SYSTEMATIC REVIEW(SR) OF THE EVIDENCE (for INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURES) 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

☐ 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**	• National Heart, Lung, and Blood Institute, National Institutes of Health (NHLBI/NIH) Asthma Guideline 2007 • www.nhlbi.nih.gov/guidelines/asthma (NAEPP Guideline)
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline,	The impact of asthma management and access to care on ED outcomes and clinical control are described well in the NHLBI NAEPP guideline. The Clinical guideline acknowledges ED visits as failures of control.

summarize the conclusions from the SR.	Class C Evidence is suggested when suggesting the relationship between specific periodicity of ambulatory visits with asthma outcomes but higher levls of evidence including Class A with the relationship between treatment and outcomes. Quick Reference: Asthma control focuses on two domains: 1)reducing impairment--- the frequency and intensity of symptoms… and 2) reducing risk – the likelihood of future asthma attacks… [later described as "prevent exacerbations] NHLBI Guideline: *As a general rule, patients with well-controlled asthma should have:* • *Few, if any, asthma symptoms.* • *Few, if any, awakenings during the night caused by asthma symptoms.* • *No need to take time off from school or work due to asthma.* • *Few or no limits on full participation in physical activities.* • *No emergency department visits.* • *No hospital stays.* • *Few or no side effects from asthma medicines.*
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	The National Asthma Education and Prevention Program (NAEPP) guidelines are the prevailing clinical recommendation for children with asthma. The Expert Panel Reports presenting clinical practice duielines for the diagnosis and management of asthma have organized recommendations for asthma care around four components considered essential to effective asthma management: - Measures of assessment and monitoring, obstained by objective tests, physical examination, patient history and patient report, to diagnose and assess the characteristics and severity of asthma and to monitor whether asthma control is achieved and maintained. - Education for partnership in asthma care - Control of environmental factors and comorbid conditions that affect asthma - Pharmacologic therapy This section of the report updates information on each of these four components based on the Expert Panel's review of the scientific literature. The sections that follow present specific clinical recommendations for managing asthma long term and for managing exacerbations that incorporate the four compoenents.

Provide all other grades and definitions from the evidence grading system	Methodology for report: Overall Methods Used To Develop This Report **Background** In June 2004, the Science Base Committee of the NAEPP recommended to the NAEPP CC that its clinical practice guidelines for the diagnosis and management of asthma be updated. In September, under the leadership of Dr. Barbara Alving, M.D. (Chair of the NAEPP CC, and Acting Director of the NHLBI), a panel of experts was selected to update the clinical practice guidelines by using a systematic review of the scientific evidence for the treatment of asthma and consideration of literature on implementing the guidelines. In October 2004, the Expert Panel assembled for its first meeting. Using EPR-2 1997 and EPR-Update 2002 as the framework, the Expert Panel organized the literature searches and subsequent report around the four essential components of asthma care, namely: (1) assessment and monitoring, (2) patient education, (3) control of factors contributing to asthma severity, and (4) pharmacologic treatment. Subtopics were developed for each of these four broad categories. The steps used to develop this report include: (1) completing a comprehensive search of the literature; (2) conducting an indepth review of relevant abstracts and articles; (3) preparing evidence tables to assess the weight of current evidence with respect to past recommendations and new and unresolved issues; (4) conducting thoughtful discussion and interpretation of findings; (5) ranking strength of evidence underlying the current recommendations that are made; (6) updating text, tables, figures, and references of the existing guidelines with new findings from the evidence review; (7) circulating a draft of the updated guidelines through several layers of external review, as well as posting it on the NHLBI website for review and comment by the public and the NAEPP CC, and (8) preparing a final-report based on consideration of comments raised in the review cycle. **Preparation Of Evidence Tables** Evidence tables were prepared for selected topics. It was not feasible to generate evidence tables for every topic in the guidelines. Furthermore, many topics did not have a sufficient body of evidence or a sufficient number of high-quality studies to warrant the preparation of a table. The Panel decided to prepare evidence tables on those topics for which an evidence table would be particularly useful to assess the weight of the evidence-e.g., topics with numerous articles, conflicting evidence, or which addressed questions raised frequently by clinicians. Summary findings on topics without evidence tables, however, also are included in the updated guidelines text.

Evidence tables were prepared with the assistance of a methodologist who served as a consultant to the Expert Panel. Within their respective committees, Expert Panel members selected the topics and articles for evidence tables. The evidence tables included all articles that received a "yes" vote from both the primary and secondary reviewer during the systematic literature review process. The methodologist abstracted the articles to the tables, using a template developed by the Expert Panel. The Expert Panel subsequently reviewed and approved the final evidence tables. A total of 20 tables, comprising 316 articles are included in the current update (see figure 1-1). Evidence tables are posted on the NHLBI Web site.

Ranking The Evidence

The Expert Panel agreed to specify the level of evidence used to justify the recommendations being made. Panel members only included ranking of evidence for recommendations they made based on the scientific literature in the current evidence review. They did not assign evidence rankings to recommendations pulled through from the EPR-2 1997 on topics that are still important to the diagnosis and management of asthma but for which there was little new published literature. These "pull through" recommendations are designated by EPR-2 1997 in parentheses following the first mention of the recommendation. For recommendations that have been either revised or further substantiated on the basis of the evidence review conducted for the EPR-3: Full Report 2007, the level of evidence is indicated in the text in parentheses following first mention of the recommendation. The system used to describe the level of evidence is as follows (Jadad et al. 2000):

- **Evidence Category A:** Randomized controlled trials (RCTs), rich body of data. Evidence is from end points of well-designed RCTs that provide a consistent pattern of findings in the population for which the recommendation is made. Category A requires substantial numbers of studies involving substantial numbers of participants.

- **Evidence Category B:** RCTs, limited body of data. Evidence is from end points of intervention studies that include only a limited number of patients, post hoc or subgroup analysis of RCTs, or meta-analysis of RCTs. In general, category B pertains when few randomized trials exist; they are small in size, they were undertaken in a population that differs from the target population of the recommendation, or the results are somewhat inconsistent.

- **Evidence Category C:** Nonrandomized trials and observational studies. Evidence is from outcomes of uncontrolled or nonrandomized trials or from observational studies.

- **Evidence Category D:** Panel consensus judgment. This category is used only in cases where the provision of some guidance was deemed valuable, but the clinical literature addressing the

subject was insufficient to justify placement in one of the other categories. The Panel consensus is based on clinical experience or knowledge that does not meet the criteria for categories A through C.

In addition to specifying the level of evidence supporting a recommendation, the Expert Panel agreed to indicate the strength of the recommendation. When a certain clinical practice "is recommended," this indicates a strong recommendation by the panel. When a certain clinical practice "should, or may, be considered," this indicates that the recommendation is less strong. This distinction is an effort to address nuances of using evidence ranking systems. For example, a recommendation for which clinical RCT data are not available (e.g., conducting a medical history for symptoms suggestive of asthma) may still be strongly supported by the Panel. Furthermore, the range of evidence that qualifies a definition of "B" or "C" is wide, and the Expert Panel considered this range and the potential implications of a recommendation as they decided how strongly the recommendation should be presented.

Panel Discussion

The first opportunity for discussion of findings occurred within the "topic teams." Teams then presented a summary of their findings during a conference call to all members of their respective committee. A full discussion ensued on each topic, and the committee arrived at a consensus position. Teams then presented their findings and the committee position to the full Expert Panel at an in-person meeting, thereby engaging all Panel members in critical analysis of the evidence and interpretation of the data.

A series of conference calls for each of the 10 committees as well as four in-person Expert Panel meetings (held in October 2004, April 2005, December 2005, and May 2006) were scheduled to facilitate discussion of findings and to dovetail with the three cycles of literature review that occurred over the 18-month period. Potential conflicts of interest were disclosed at the initial meeting.

Report Preparation

Development of the EPR-3: Full Report 2007 was an iterative process of interpreting the evidence, drafting summary statements, and reviewing comments from the various external reviews before completing the final report. In the summer and fall of 2005, the various topic teams, through conference calls and subsequent electronic mail, began drafting their assigned sections of the report. Members of the respective committees reviewed and revised team drafts, also by using conference calls and electronic mail. During the calls, votes were taken to ensure agreement with final conclusions and recommendations. During the December 2005 meeting, Panel members reviewed and discussed all committee drafts.

During the May 2006 meeting, the Panel conducted a thorough review and discussion of the report and reached consensus on the recommendations. For controversial topics, votes were taken to ensure that each individual's opinion was considered. In July, using

	conference calls and electronic mail, the Panel completed a draft of the EPR-3: Full Report 2007 for submission in July/August to a panel of expert consultants for their review and comments. In response to their comments, a revised draft of the EPR-3: Full Report 2007 was developed and circulated in November to the NAEPP Guidelines Implementation Panel (GIP) for their comment. This draft was also posted on the NHLBI Web site for public comment in February 2007. The Expert Panel considered 721 comments from 140 reviewers. Edits were made to the documents, as appropriate, before the full EPR-3: Full Report 2007 was finalized and published. The EPR-3: Full Report 2007 will be used to develop clinical practice guidelines and practice-based tools as well as educational materials for patients and the public. **References** 1. EPR. Expert panel report: guidelines for the diagnosis and management of asthma (EPR 1991). NIH Publication No. 91-3642. Bethesda, MD: U.S. Department of Health and Human Services; National Institutes of Health; National Heart, Lung, and Blood Institute; National Asthma Education and Prevention Program, 1991. 2. EPR-2. Expert panel report 2: guidelines for the diagnosis and management of asthma (EPR-2 1997). NIH Publication No. 97-4051. Bethesda, MD: U.S. Department of Health and Human Services; National Institutes of Health; National Heart, Lung, and Blood Institute; National Asthma Education and Prevention Program, 1997. 3. EPR-Update 2002. Expert panel report: guidelines for the diagnosis and management of asthma. Update on selected topics 2002 (EPR-Update 2002). NIH Publication No. 02 5074. Bethesda, MD: U.S. Department of Health and Human Services; National Institutes of Health; National Heart, Lung, and Blood Institute; National Asthma Education and Prevention Program, June 2003. 4. Jadad AR, Moher M, Browman GP, Booker L, Sigouin C, Fuentes M, Stevens R. Systematic reviews and meta-analyses on treatment of asthma: critical evaluation. BMJ 2000;320(7234):537-40. 5. NHIS. National health interview survey (NHIS 2005). Hyattsville, MD: National Center for Health Statistics (NCHS), Centers for Disease Control and Prevention, 2005. Available at http://www.cdc.gov/nchs/about/major/nhis/reports_2005.htm. Link to the evidence tables themselves: http://www.nhlbi.nih.gov/health-pro/guidelines/current/asthma-guidelines/evidence-tables
Grade assigned to the **recommendation** with definition of the grade	
Provide all other grades and definitions from the recommendation grading system	
Body of evidence:	**Systematic Evidence Review Overview**

• Quantity – how many studies? • Quality – what type of studies?	**Inclusion/Exclusion Criteria** The literature review was conducted in three cycles over an 18-month period (September 2004 to March 2006). Search strategies for the literature review initially were designed to cast a wide net but later were refined by using publication type limits and additional terms to produce results that more closely matched the framework of topics and subtopics selected by the Expert Panel. The searches included human studies with abstracts that were published in English in peer reviewed medical journals in the MEDLINE database. Two timeframes were used for the searches, dependent on topic: January 1, 2001, through March 15, 2006, for pharmacotherapy (medications), peak flow monitoring, and written action plans, because these topics were recently reviewed in the EPR-Update 2002; and January 1, 1997, through March 15, 2006, for all other topics, because these topics were last reviewed in the EPR-2 1997. **Search Strategies** Panel members identified, with input from a librarian, key text words for each of the four components of care. A separate search strategy was developed for each of the four components and various key subtopics when deemed appropriate. The key text words and Medical Subject Headings (MeSH) terms that were used to develop each search string are found in an appendix posted on the NHLBI Web site. **Literature Review Process** The systematic review covered a wide range of topics. Although the overarching framework for the review was based on the four essential components of asthma care, multiple subtopics were associated with each component. To organize a review of such an expanse, the Panel was divided into 10 committees, with about 4-7 reviewers in each (all reviewers were assigned to 2 or more committees). Within each committee, teams of two ("topic teams") were assigned as leads to cover specific topics. A system of independent review and vote by each of the two team reviewers was used at each step of the literature review process to identify studies to include in the guidelines update. The initial step in the literature review process was to screen titles from the searches for relevancy in updating content of the guidelines, followed by reviews of abstracts of the relevant titles to identify those studies meriting full-text review based on relevance to the guidelines and study quality. The combined number of titles screened from cycles 1, 2, and 3 was 15,444. The number of abstracts and articles reviewed for all three cycles was 4,747. Of these, 2,863 were voted to the abstract Keep list following the abstract-review step. A database of these abstracts is posted on the NHLBI Web site. Of these abstracts, 2,122 were advanced for full-text review, which resulted in 1,654 articles serving as a bibliography of references used to update the guidelines, available on the NHLBI Web site. Articles were selected

	from this bibliography for evidence tables and/or citation in the text. In addition, articles reporting new and particularly relevant findings and published after March 2006 were identified by Panel members during the writing period (March 2006-December 2006) and by comments received from the public review in February 2007.
Estimates of benefit and consistency across studies	In summary, the NAEPP "Expert Panel Report 3: Guidelines for the Diagnosis and Management of Asthma-Full Report 2007" represents the NAEPP's ongoing effort to keep recommendations for clinical practice up to date and based upon a systematic review of the best available scientific evidence by a Panel of experts, as well as peer review and critique by the collective expertise of external research/science consultants, the NAEPP CC members, guidelines implementation specialists, and public comment. The relationship between guidelines and clinical research is a dynamic one, and the NAEPP recognizes that the task of keeping guidelines' recommendations up to date is an increasing challenge. In 1991, many recommendations were based on expert opinion because there were only limited randomized clinical trials in adults, and almost none in children, that adequately tested clinical interventions grounded in research findings about the disease process in asthma. The large gaps in the literature defined pressing clinical research questions that have now been vigorously addressed by the scientific community, as the size of the literature reviewed for the current report attests. The NAEPP is grateful to all of the Expert Panel members for meeting the challenge with tremendous dedication and to Dr. William Busse for his outstanding leadership. The NAEPP would particularly like to acknowledge the contributions of Dr. Gail Shapiro, who served on NAEPP Expert Panels from 1991 until her death in August 2006. Dr. Shapiro provided valuable continuity to the Panel's deliberations while simultaneously offering a fresh perspective that was rooted in observations from her clinical practice and was supported and substantiated by her clinical research and indepth understanding of the literature. Dr. Shapiro had a passion for improving asthma care and an unwavering commitment to develop evidence-based recommendations that would also be practical. Dr. Shapiro inspired in others the essence of what NAEPP hopes to offer with this updated Expert Panel Report: a clear vision for clinicians and patients to work together to achieve asthma control.
What harms were 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.

We assert that without a measure of whether or not the reason the child is in the emergency room is sufficient to make it a clinically appropriate visit, it is impossible to interpret whether an ED visit represents a failure of clinical management and control, or a failure of the primary care or other aspects of the health care system to provide care at a more appropriate level of care.

Overarching statement: Even when not specifically indicated, we are interested in how these constructs are impacted by such factors as race, ethnicity, socioeconomic status or its indicators, or the presence of other special health care needs.

Our metric is designed to capture axes related to two distinct conceptual frameworks:

1) Asthma is a model of chronic disease management. In other words, ED visits may arise from acute exacerbations indicating a flare up of disease, and/or suboptimal management of the chronic illness.
2) ED visits for asthma may reflect limitations of primary care beyond the provision of suboptimal treatment, such as insufficient education, limitations of access or availability, breakdowns of communication, or a variety of other factors.

We note that the internal quality of the ED visit to manage the asthma is not the target of this measure. However, communication between the emergency department and the primary care site may prove to be within the scope of this measure, pending the views of our experts and developers.

Construct I: Need to sufficiently specify population for measure

Concept	Implications (Lay Statement)	Lit Review Questions
(Descriptive) **The measure will need to adequately specify the population that we consider to be eligible for an ED with asthma measure.**	The development of measures regarding ED use for children with asthma requires us to understand the strengths and weaknesses for our measure of various approaches to identifying whether or not children have asthma. It further requires us to understand the impact of the availability of various sources of data (such as encounter data, pharmaceutical data, electronic medical record or chart review data) on these strengths and weaknesses. We are aware that the use of the term asthma is variable. We are not interested in diagnoses with the name asthma, but with an operational diagnosis that we will functionally treat as asthma, whether it has been called chronic wheezing, reactive airway disease, chronic infectious bronchitis, etc. We recognize that asthma and its presentation may change over the course of a child's life.	1. When asthma care is evaluated, how is the population of care recipients defined? How is asthma defined? What is the impact of including various types of data (dx 1 or more, drugs, etc) on the sensitivity and specificity of asthma identification? What are practical and valid approaches to identifying asthma? How do the answers to these questions differ between adults and children? 2. Are any groups persistently excluded from studies of asthma care (i.e., are children who have asthma and other comorbid conditions, such as a malignant disease, excluded?). What rationale is provided for the exclusion? 3. Are any non-asthma diagnoses considered to be indicators of asthma or potential asthma (e.g. bronchitis, bronchiolitis, wheezing, atopy) 4. For children up to age 21, how do issues of diagnosis, management, and follow-up differ by age and developmental stage? 5. At what point does literature suggest that reactive airway disease should be managed as asthma? a. What other conditions are managed as asthma? 6. What common current or preexisting comorbid conditions alter the management plan for asthma?

Construct II: Adequacy of management of asthma (as a chronic disease example)

Concept	Implications (Lay Statement)	Lit Review Questions
IIA. **↑Adequacy of asthma management:** **↓ED visits**	Since asthma is a chronic disease characterized by acute exacerbations, the extent to which asthma care is optimized through the use of appropriate medications, the control of the	1. What are the recommendations of the NHLBI guidelines? a. What does the literature suggest about the usefulness of NHLBI guidelines? b. Are there aspects that it has identified that appear to be missed?

	environment, and the preparation of the parent/child dyad to adapt to changes in circumstances (e.g. viral respiratory infection or exposure to cold) should reduce the number of ED visits, irrespective of the number of primary care visits.	2. What do we know about asthma management, how it's measured, who provides it, patterns of care and how ED visits vary as a consequence? 3. Does identification of PCP improve outcomes of ED visit, including patterns of care, utilization? 4. What do we know about the content of an asthma plan and its relationship to a full program of chronic disease management, and its influence on ED utilization? 5. What evidence is there about the impact on outcomes such as ED use when the child or adolescent is involved in asthma self-management? For example, does it matter if: a. The child has a written asthma plan? b. The child understands their asthma plan? c. The child is given an opportunity to participate in managing care? 6. How is the role of the child in self-management measured? 7. How much are children able to recognize, communicate and act on their asthma? 8. What do we know about the impact of asthma services on asthma management? This includes: a. Treatment from an asthma specialist; b. Social worker; or c. Multidisciplinary personnel 9. To what extent is ED use by children with asthma stimulated by non-asthma related issues? a. How can we identify when that occurs? b. What is the evidence that providing other services will reduce the number of ED visits? 10. To what extent do children contribute to their management (including avoiding triggers, recognizing symptoms, medication adherence, etc.)? a. What is the impact and variance by age? 11. What is the evidence regarding adequacy of various medication delivery systems for infants, toddlers, children and adolescents in acute and chronic settings?

	12. Is there evidence of prior insult to the lungs such as sequelae of prematurity, etc. that create distinct subpopulations when considering this measure (at risk for ER visit)? 13. What aspects of the health services environment have been identified as contributing to outcomes of asthma management (e.g. school based health care)? 14. Does rate of ED utilization for non-respiratory diagnoses vary between asthmatics and non-asthmatics? 15. What is known about how often children with asthma use the ED over an extended period of time? Does it change over the life course of childhood? How does that vary by child characteristics, including race, SES, urban, suburban vs. rural, and age?

IIB. **↑PCP capacity/knowledge/skill:** **a. ↑Asthma management** **b. ↓Asthma exacerbations** **c. ↑Chronic disease management**	Broadly speaking, patient management of asthma is influenced by the capacity of the PCP practice. This includes the knowledge and skills possessed by the PCP, as well as office support to enhance access and availability of care. PCP includes the ability of the PC office to meet the cultural needs of the patient and their family.	1. What are the diversity of practices or services that may or may not impact ability or capacity of the PCP practice to manage asthma? 2. What do we know about the specific skills and processes that contribute to a primary care practice's capacity? 3. What patterns of visits or medication use or other indicators have been used as markers of well or poorly delivered primary care for asthma in children and/or adults? 4. What is the minimum use of specialists appropriate for children with asthma? How does that vary with history of ED or hospital use? a. When and how does the use of specialists become a marker for higher or lower quality of care? 5. What evidence is there regarding the nature of the PCP practice for children with asthma? For example, the level of continuity with individual clinicians vs. practices, the accessibility of specified clinicians and/or practices during the day and/or after work hours, etc.

IIC. **↑Asthma education:** **a. increases recognition of symptoms >** **b. ↑Management skills**	Enhancing what patients or their families know about asthma may be an important tool to improve care for children with asthma. The likely first effect of such education is to enhance the capacity of a caregiver to identify what symptoms may relate to asthma. This could conceivably increase utilization of both PCP and ED services if this were to increase the caregiver's perceived need for care for their child's asthma. With a more sophisticated understanding, including having a valid asthma action plan and understanding how to use it, ED care may be reduced and PCP care for asthma may be reduced, as symptoms are less frequent and parents are more competent to manage them when they arise.	1. What are metrics or processes regarding the quality of asthma care? Is it drug ratios (i.e. proportion of prescriptions filled that are for rescue vs control medications), asthma action plan, , capacity of PCP office, relationship to PCP practice, or other specific bundles of care, etc? 2. What constitutes "perfect care"/"best practice" for any specified type of patient? 3. What do we know about the impact of asthma education programs on quality of care, outcomes of care, or utilization of care? Define utilization of care as including: a. PCP utilization, b. ED utilization, c. Referral/specialist utilization, d. Non physician care team member utilization, e. Medication usage, f. Hospitalizations, and/or g. Other care utilization areas to consider? Examples may include functional status, quality of life elements, spirometry, role functioning. 4. What is the diversity of asthma education programs and what are the differences in quality of care/outcomes/utilization of care associated with differences? 5. Does referral to an asthma specialist impact quality of care, utilization of care and asthma outcomes? 6. Does referral to a social worker impact utilization of care and asthma outcomes? 7. (Broad) Does involvement of multidisciplinary personnel (beyond allopathic or osteopathic physicians) impact quality of care, utilization of care and asthma outcomes? 8. What are desirable roles and effectiveness of interventions that extend beyond the healthcare system, such as reducing pollution, focusing on environmental justice, housing, dust mites, etc.? 9. How does organization and capacity of the practice setting influence the delivery of asthma management education?

Construct III: **Adequacy of PCP practice site to handle acute exacerbations of chronic disease and/or acute illnesses**

Concept	Implications (Lay Statement)	Lit Review Questions
IIIA. **↑Primary care capacity:** **a. ↑ PCP visits (routine, WCC)** **b. ↑PCP visits (other acute dx)** **c. ↑ PCP visits (asthma)** **d. ↓ED visits (acute dx, asthma)** **IIIA.2** **SUBCONSTRUCT:** **↑Accessibility:** **a. ↑ PCP visits (routine, WCC)** **b. ↑PCP visits (other acute dx)** **c. ↑ PCP visits (asthma)** **d. ↓ED visits (acute dx, asthma)**	In general, enhanced capacity may affect a patient's access to care. Capacity can refer to patient services that make it easier for a patient to receive timely care, such as location or hours of offices, to the ability to triage phone calls in a timely and effective way, or may include the materials and services present within an office (e.g. the presence of a treatment room, the capacity to deliver oxygen, nebulizers, etc.) Such capacity may be limited or enhanced by staffing, space, the ability to safely transport someone from the office to a hospital, etc. If PCP office capacity is optimized, ED visits may be reduced as acute and mundane conditions can be managed in a PCP setting. Subsequently, increased capacity of the entire PCP support network will increase number of PCP visits.	1. What do we know about access to the PCP's office as a place to manage asthma, and the subsequent capacity of a PCP and the diversity of practice settings? Additionally, how do we measure capacity and, its impact on QoC, processes of care, asthma outcomes, asthma specific processes and utilization? How do these factors impact ED use or other outcomes? a. In general: i. PCP/specialist ratio in a plan or PCP/child ratio ii. PCP time spent in visit (incl. minutes per sick, well-child, asthma management visit) iii. Nature of training activities iv. How long does it take to schedule a visit (incl. asthma (chronic), acute, follow-up visit) v. Office hours and visit flexibility (incl. after hours coverage, office consult, meet in ED) vi. Phone capabilities: (incl. answering capacity, putting on hold, returning calls, after hours phone service) vii. Level of implementation of patient centered medical home/chronic care model, eg i. Use of registries ii. Standardized tools for measurement iii. Case management iv. Group visits or other education, etc b. Specifically, ability to manage acute dx in office, which includes: i. Do they have a treatment room or capacity to use a room as a treatment room? ii. Do they offer rescue treatments (e.g. nebulizers, spacers)? iii. Can they measure oxygen saturation? iv. Do doctors feel comfortable with acute asthmatic in office? v. Can they take time to manage an acute pt in their office?

vi. Do they have safe and rapid transport to a hospital (how long?)

2. Availability and accessibility of offices (incl. office hours, geographic distribution)
   a. What do we know about linguistic capabilities in the PCP setting influencing use of the ED?
   b. What do we know about proximity of the PCP office to public transit on the utilization of the ED?
3. What do we know about the impact of variations in patterns of care/practice, use of modalities, and/or and receipt of well-child care on asthma management or outcomes (eg ED use)? Does Immunization status reflect on t eh capacity of the PCP, on the state of the child, or on other factors that may relate to asthma outcomes? How about the sufficiency of the number of WCC Visits (eg meets HEDIS standard or AAP standard or does not)? Absolute number of visits to PCP?
4. Are children with more WCC visits less likely to use the ED for acute visits? children who are UTD on their immunizations?
5. What literature is there on the relationship between pediatric ED use and other measures of asthma exacerbation/outcomes?
6. What do we know about variability of capacity and management of mundane conditions (e.g. OM, URIs, pharyngitis), office to ED ratios?
7. What do we know about variability of capacity and management of acute conditions requiring interventions (e.g. asthma)?
8. To what extent does ED capacity increase use of ED services? Do hospitals advertise ED services, have fast track for mundane conditions, etc?
9. To what extent does ED have capacity to provide primary care, routine immunizations, etc? How is that built into policies and protocols?
10. At what age does the PCP start meeting alone with child? Time spent in visit?
11. To what extent and at what age do PCP's involve children in self-management and does it vary?

IIIB. **↑Relationship with PCP:** **a. ↑ PCP visits (routine, WCC)** **b. ↑PCP visits (other acute dx)** **c. ↑ PCP visits (asthma)** **d. ↓ED visits (acute dx, asthma)**	Improved relationship with PCP may increase visits to your PCP and decrease ED visits, for both acute and mundane conditions. A good relationship may lead to greater trust and adherence to recommendations (both WCC and asthma care) and drive a preference for seeking care by the PCP over seeking care in another environment. In general, we are referring to relationship of caregiver with PCP and their office staff. We recognize the importance of the relationship of PCP's with patients as well; when the relationship between the PCP and the child rather than caretaker is emphasized in research, we'd like to capture that as well.	1. What exists regarding measuring the quantity and quality of the relationship with PCP? Specifically: a. What's the variation and does it matter? b. How is it measured? c. What do we know about patient experience of care, especially as it relates to relationship with clinicians/PCP d. To what extent is quality of relationship expressed in terms of caregiver vs. child relationships and how does this change with age of child or longevity of connection to a PCP? 2. What evidence is there regarding use of supplemental services outside of regular clinical visits and how do these services impact quality and utilization of care? Define supplemental services as: a. Electronic educational/reminder tools (incl. social media) b. Telephone educational/reminder tools c. Print materials (e.g. educational brochures) d. Disease management, demand management, or other type programs e. Other services to consider? Measure quality, utilization of care should include at least : a. ED visits b. PCP visits 3. How does role of child in self care/management tie into these issues?

Construct IV: The connectedness of care in the primary care and ED setting – before, during, and after of the ED visit

Concept	Implications (Lay Statement)	Lit Review Questions
IV. (Descriptive) Enhanced integration of ED care of asthma with routine care will have better outcomes	If primary care is generally pretty good, then the ED visit should be an extraordinary event. In such cases the PCP alerting the ED to current management and the ED assuring appropriate follow up with the PCP is important. In cases where primary care is of lower quality or more variable, the ED visit may enhance the long term management of the child with asthma. And we need to assess this. One of the ways it might do so is to construct an asthma management plan that is then followed by the PCP. Another way is to connect a child without adequate primary care to primary care, especially to someone who is competent to manage the asthma.	1. What evidence supports that ED visits for asthma are most effective when visit is followed by a visit to the PCP? 2. Do utilization patterns in both the ED and primary care setting change following ED visits? 3. Is an effective/more effective use of medications seen following an ED visit? 4. Does the identification of a primary care provider improve outcomes of an ED visit (including patterns of care utilization)? 5. Is pre or intra visit communication with the primary care provider associated with better outcomes? How often does this occur? Are there systematic differences regarding those for whom this does and does not occur? 6. Are ED visits for asthma routinely associated with some form of communication or linkage with PCP? Does that result in better outcomes?

Construct V: Equity is a value in asthma care

Concept	Implications (Lay Statement)	Lit Review Questions
V. (Descriptive) Equity is a critical construct of quality for children with equity	Systematic differences in the frequency or nature of ED visits for asthma on the basis of race, ethnicity, family make-up, income/economic status, specifics of insurance status, presence or absence of comorbid special health care needs, etc represents decrements in quality that our measures should identify.	1. Does the literature indicate systematic or predictable differences in the frequency or nature of asthma care for children as it relates to ED visits for asthma that may be interpreted as representing inequitable structures, processes, outcomes, experiences with, or coordination of care? 2. What do we know about how social determinants and diagnosis and management of asthma and its outcomes, specifically as it relates to use of ED? 3. What do we know about the extent to which use of the ED for children with asthma that relates to the external physical and social environment?

Proposed Research Questions

Asthma- We propose to prioritize our Asthma Construct Table, to the following questions:

Acronyms

PCP: Primary Care Provider

ED: Emergency Department

WCC: Well-child care

Baseline Question (for Questions 1, 2 and 3 below):

When asthma care is evaluated, how is the population of a t the population level? What are specific implications of how you ident ng various approaches to specifying the denominator of children l and valid approaches to identifying asthma at the population level? How do the answers to these questions differ between adults and children?

Question 1 (Construct IIA.2):

For children with asthma, what do we know about asthma management? How is management of asthma described and measured? This includes who (PCP, asthma specialist, ED, etc) primarily manages it as well as who provides it. What are the patterns of care and what do we know about how use of the ED varies as a result of various approaches to management?

- **Question 1a (Construct IIB.3):**

Specifically, have any of these patterns of visits or medication use or other characteristics of care been used as markers of well or poorly delivered primary care for asthma for children and/or adults?

Question 2 (Construct IIB.5):

How has varying asthma care for children been described on the basis of characteristics of the PCP offices or practices? For example, are they characterized by the level of continuity between individual clinicians, the level of conntiuity with any provider in the practice, the accessibility of specified clinicians and/or practices during the day and/or after work hours, etc?

- **Question 2a (Construct IIIA.3):**

What do we know about the impact of variations in patterns of care/practice, use of treatment modalities, and/or receipt of well-child care on asthma management or outcomes (e.g. ED use)? How about the sufficiency of the number of WCC Visits (eg meets HEDIS standard or AAP standard or does not)? Absolute number of visits to PCP?

Question 3 (Construct IIC.7):

(Broad) Does involvement of multidisciplinary personnel (beyond allopathic or osteopathic physicians) impact quality of care, utilization of care and asthma outcomes both within context of a primary care practice or in other clinical settings?

- **Question 3a. (Construct IIIB.2):**

What evidence is there regarding use of supplemental services outside of regular clinical visits and how do these services impact quality and utilization of care?

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

We conducted a scoping review as follows:
We identified key constructs of asthma ED use measures for consideration. We created a table of these constructs in technical and lay language, and listed research questions for the review to answer. Our contractor (a national accrediting body experienced in measure development), prepared for us a literature review in 2 stages and we supplemented this with targeted reviews as needed to answer specific questions that arose during the measure development process.

The above construct table was used to guide the review and was the basis for the first round of review. Following the table, we include a list of questions for focused review that guided round 2 of the review, which resulted in a detailed summary of 91 articles from the peer-reviewed literature. In addition to this review, the CAPQuaM scientific team conducted an ad hoc series of reviews to answer specific questions such as the reliability of administrative data to identify asthma, and the value of expert panels and the RAND/UCLA appropriateness method. The CAPQuaM degree 360 method starts with a topic area and the measures emerge during the process, in this case necessitating the specified ad hoc reviews.

We searched peer reviewed and gray literature from 1985-2014 over the course of these reviews. Literature was summarized for our expert panel, which met in late 2013.

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

Our approach to developing this measure stems from several vibrant and scientifically sound traditions. We first discuss research involving the soundness of our data sources, which include both administrative data to identify cases (and a fraction of numerator qualifications) and chart review (medical record audit) to confirm some denominator inclusions and to identify most numerator inclusion. This is a generally accepted and standard approach with acceptable reliability.

Brook and Davies [1] trace the early history of quality measurement and remind us of the importance of medical chart audit as an approach to quality measurement. Lohr and Brook at RAND and Roos in Manitoba, Canada pioneered the use of electronically-available administrative data (generated by routine health care operations, such as billings) as proxies for health care processes. Administrative data carefully used reduces burden of quality measurement. [2-6]

As the National Committee for Quality Assurance (NCQA) developed the Healthcare Employee Data Information Set (HEDIS) as the de facto measurement system for managed care, attention turned to the use of administrative data for routine performance measurement. Research demonstrated that administrative data could have a role in producing quality measures, with augmentation by chart review often necessary. Administrative data are not typically sufficient for detailed clinical assessment. [7-11] HEDIS developed a hybrid approach, using administrative data and chart review, which this measure borrows heavily from. [12, 13]

We have used rigorous and transparent methods [14] to assemble a national expert panel that included pediatricians, family physicians, pediatric and general emergency room specialists, a pediatric pulmonologist and a pediatric allergist from practices and medical schools around the country. This work was conducted in collaboration with national clinical societies (AAP, AAFP) and CAPQuaM's diverse other partner organizations, including NY State DoH/Medicaid. NCQA is an important technical consultant and partner. The specific criteria that we operationalize in this

measure were all rated by the expert panel with a median score of 8 or 9 on a 9 point scale (9 high) as circumstances for which the ED is an appropriate level of care. The use of Expert Panels has been demonstrated to be useful in measure development and health care evaluation, including for children.

Select references documenting other aspects of performance gap, and supporting our process and data sources are also noted (15-35).

1. Brook, R.H. and A. Davies-Avery, *Mechanisms for Assuring Quality of US Medical Care Services: Past, Present, and Future*. 1977: Santa Monica, CA.
2. Brook, R.H. and K.N. Lohr, *Efficacy, effectiveness, variations, and quality. Boundary-crossing research.* Med Care, 1985. **23**(5): p. 710-22.
3. Lohr, K.N., *Use of insurance claims data in measuring quality of care.* Int J Technol Assess Health Care, 1990. **6**(2): p. 263-71.
4. Roos, L.L., et al., *Using Administrative Data Banks for Research and Evaluation: A Case Study.* Evaluative Research 1979. **236**(3).
5. Roos, L.L., Jr., J.P. Nicol, and S.M. Cageorge, *Using administrative data for longitudinal research: comparisons with primary data collection.* J Chronic Dis,1987. **40**(1): p. 41-9.
6. Roos, L.L., Jr., et al., *How good are the data? Reliability of one health care data bank.* Med Care, 1982. **20**(3): p. 266-76.
7. Dresser, M.V., et al., *Clinical quality measurement. Comparing chart review and automated methodologies.* Med Care, 1997. **35**(6): p. 539-52.
8. Newton, K.M., et al., *The use of automated data to identify complications and comorbidities of diabetes: a validation study.* J Clin Epidemiol, 1999. **52**(3): p. 199-207.
9. Thompson, B.L., et al., *Measuring clinical performance: comparison and validity of telephone survey and administrative data.* Health Serv Res, 2001. **36**(4): p. 813-25.
10. Angier, H., et al., *Variation in outcomes of quality measurement by data source.* Pediatrics, 2014. **133**(6): p. e1676-82.
11. Weiskopf, N.G. and C. Weng, *Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research.* Journal of the American Medical Informatics Association, 2013. **20**(1): p. 144-151.
12. Pawlson, L.G., S.H. Scholle, and A. Powers, *Comparison of administrative-only versus administrative plus chart review data for reporting HEDIS hybrid measures.* Am J Manag Care, 2007. **13**(10): p. 553-8.
13. NCQA. *National Committee for Quality Assurance*. [cited 2014 7/30/14]; Available from: http://www.ncqa.org/.
14. Brook, R.H., et al., *A method for the detailed assessment of the appropriateness of medical technologies.* International journal of technology assessment in health care, 1986. **2**(01): p. 53-63.
15. DeAngelis, C., P. Fosarelli, and A.K. Duggan, *Use of the emergency department by children enrolled in a primary care clinic.* Pediatr Emerg Care, 1985. **1**(2): p.61-5.
16. Talreja, N., et al., *Modifiable factors associated with severe asthma exacerbations in urban patients.* Ann Allergy Asthma Immunol, 2012. **109**(2): p.128-32.
17. Auger, K.A., et al., *Medical home quality and readmission risk for children hospitalized with asthma exacerbations.* Pediatrics, 2013. **131**(1): p. 64-70.
18. Ducharme, F.M., et al., *Written action plan in pediatric emergency room improves asthma prescribing, adherence, and control.* Am J Respir Crit Care Med, 2011. **183**(2): p. 195-203.

19. Farber, H.J., *Optimizing maintenance therapy in pediatric asthma.* Curr Opin Pulm Med, 2010. **16**(1): p. 25-30.
20. Smith, S.R., D.B. Wakefield, and M.M. Cloutier, *Relationship between pediatric primary provider visits and acute asthma ED visits.* Pediatr Pulmonol, 2007. **42**(11): p. 1041-7.
21. Parchman, M.L. and S. Culler, *Primary care physicians and avoidable hospitalizations.* J Fam Pract, 1994. **39**(2): p. 123-8.
22. Prevention, C.f.D.C.a. *Home-based Multi-trigger, Multi-component interventions*. 2013 [cited 2013 May 20]; Available from: http://www.cdc.gov/asthma/interventions.htm.
23. Price, J.H., et al., *Racial/ethnic disparities in chronic diseases of youths and access to health care in the United States.* Biomed Res Int, 2013. **2013**: p. 1-12; Available in open access http://dx.doi.org/10.1155/2013/787616
24. Homer, C.J., et al., *Does quality of care affect rates of hospitalization for childhood asthma?* Pediatrics, 1996. **98**(1): p. 18-23.
25. Finkelstein, J.A., et al., *Quality of care for preschool children with asthma: the role of social factors and practice setting.* Pediatrics, 1995. **95**(3): p. 389-94.
26. Flores, G., et al., *Keeping children out of hospitals: parents' and physicians' perspectives on how pediatric hospitalizations for ambulatory care-sensitive conditions can be avoided.* Pediatrics, 2003. **112**(5): p. 1021-30.
27. Knudson, A., et al., *Disparities in pediatric asthma hospitalizations.* J Public Health Manag Pract, 2009. **15**(3): p. 232-7.
28. Bindman, A.B., et al., *Preventable hospitalizations and access to health care.* Jama, 1995. **274**(4): p. 305-11.
29. National Heart, Lung and Blood Institute. *Asthma Guidelines*. 2011 February 2011 [cited 2014 7/30/2014].
30. *A Matter of Urgency: Reducing Emergency Department Overuse* in *A NEHI Research Brief* 2010, New England Healthcare Institute.
31. Martin, B.C., *Emergency medicine versus primary care: a case study of three prevalent, costly, and non-emergent diagnoses at a community teaching hospital.* J Health Care Finance, 2000. **27**(2): p. 51-65.
32. Finkelstein, J.A., et al., *Comparing asthma care for Medicaid and non-Medicaid children in a health maintenance organization.* Arch Pediatr Adolesc Med, 2000. **154**(6): p. 563-8.
33. Owens, P.L., et al., *Care of children and adolescents in U.S. hospitals*, in *HCUP Fact Book No. 4*. Agency for Healthcare Research and Quality: Rockville: MD.
34. Pearson, W.S., et al., *State-based Medicaid costs for pediatric asthma emergency department visits.* Prev Chronic Dis, 2014. **11**: p. E108.
35. Taubman, S.L., et al., *Medicaid Increases Emergency-Department Use: Evidence from Oregon's Health Insurance Experiment.* Science, 2014. **343**(6168): p. 263-268.

The appropriate use criteria were derived from a set developed by an expert panel who synthesized the literature and their expert opinion into explicit criteria using the RAND/UCLA appropriateness method. Criteria that were rated 8 or 9 by the panel were included for this measure. The criteria set includes:

1) Hospitalization directly from the ED;
2) Documented physical findings consistent with respiratory distress, including:
    a) Labored breathing with retractions and/or evidence of accessory muscle use;
    b) Markedly decreased breath sounds;
3) O_2 saturation level less than 90 percent on percutaneous assessment;
4) An ABG obtained (or ordered);

5) Consultation ordered and obtained with a pulmonologist asthma specialist, an order of an arterial blood gas (ABG), or a consult with a pulmonary or asthma specialist.
6) Parent/caregiver referred to the ED after evaluation from the PCP or other office/clinic;
7) Parent/caregiver report of administering two or more doses of inhaled rescue medications without meaningful clinical improvement;
8) Parent/caregiver report that the child was in a pre-defined "red zone" of peak flow measurement as part of an asthma action or similar plan; or,
9) Parent/caregiver report of a rapid and life-threatening deterioration after a similar prior episode. This criterion is not included in the specifications for this measure.

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

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

1a. Evidence to Support the Measure Focus – See attached Evidence Submission Form
nqf_evidence_attachment_12_11_16.docx
1a.1 For Maintenance of Endorsement: Is there new evidence about the measure since the last update/submission?
Please update any changes in the evidence attachment in red. Do not remove any existing information. If there have been any changes to evidence, the Committee will consider the new evidence. If there is no new evidence, no updating of the evidence information is needed.

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 PRO-PM (e.g. HRQoL/functional status, symptom/burden, experience with care, health-related behaviors), provide evidence that the target population values the measured PRO and finds it meaningful. (Describe how and from whom their input was obtained.)
IF a COMPOSITE (e.g., combination of component measure scores, all-or-none, any-or-none), SKIP this question and provide rationale for composite in question 1c.3 on the composite tab.
Asthma is one of the most common indications for emergency department (ED) visits by children. (1-3) AHRQ's Healthcare Cost and Utilization Project (HCUP) data from the Nationwide Emergency Department Sample (NEDS) found that in 2012, children between 1 and 17 years old had more than 1,895,000 ED visits for asthma with almost 10% resulting in hospitalization.

Evidence suggests that ED visits and hospitalizations in children with asthma vary systematically by how well-equipped that community is to provide primary care, and by the quality of primary care delivered. (4, 5) There is widespread literature illustrating that ED visits and hospitalizations are each undesirable utilization outcomes from poorly managed asthma. There is not a large literature that assesses whether or not pediatric ED visits were appropriate. (6 -10)

A body of literature has explored the value and feasibility of measuring the appropriateness of medical activities using data available in the medical record. (11-14) Early work in adults included assessment of hysterectomy, carotid endarterectomy and cardiac interventions. An independent research project brought the construct of appropriateness to children (15), while Kleinman and colleagues were the first to assess the appropriateness of specific pediatric procedures. (16, 17) A later study demonstrated the feasibility of medical record data for such an assessment. (18) DeAngelis pioneered studies of what constitutes a good reason to use the ED. (6) All of these studies used a definition of appropriateness that compared benefit to likely risk without specific consideration of costs. The need for more studies looking for overuse was recently reviewed. (19) RAND type Delphi panels are accepted around the world as a method for developing criteria to assess appropriateness. (20-22)

Research demonstrates that:

•ED visits are an important issue for child health insurers, including Medicaid, with clinical and financial consequences;

•An overcrowded primary care system contributes to ED use for non-emergent and even non-urgent conditions.

•Pediatric hospitalizations for asthma vary by primary care availability and quality

•ED visits are common for children with asthma, including those in Medicaid

•Assessment of appropriateness using information in the medical record is a well-established and validated method that has been successfully applied to children.

The literature suggests that a measure that assesses whether or not the ED is an appropriate level of care for a child with asthma at the time that they present has intrinsic value. Such a measure would:

•Characterize the process of care in a way that assesses whether a particular ED visit represents overuse

•Allow the outcomes of asthma care to be better characterized in a manner that describes performance and promotes targeted improvement. Inappropriate ED visits represent failures of primary care delivery, availability and/or access. Appropriate visits may represent a failure to control asthma. These have distinct and distinguishable meanings that contribute to the understanding of the quality of asthma care.

•Measuring the quality of asthma care requires assessment of multiple factors. This appropriateness measure helps plans, purchasers, and society to understand the implication of asthma ED visits as outcomes of asthma care. The implications herein is that understanding what is better or worse care requires looking at various factors and not simply a higher or lower appropriateness score. The understanding of this measure is enhanced by considering whether the rate of undesirable outcomes (ED visits and hospitalizations) is high or low and whether other measures of primary care availability and access or asthma quality suggest high levels of performance or not..

An abstract describing the proposed measure was peer-reviewed and subsequently presented to a national audience at AcademyHealth 2014 Annual Research Meeting in San Diego in the “Measuring the Safety, Quality, and Value” section. Feedback was positive regarding the methods, measures, ethics, and importance of this measure.

Research evidence supports the importance and need for our proposed measure that assesses whether the ED represents an appropriate level of care for children with asthma who are seen in the ED.

1.Kharbanda, A.B., et al., Variation in resource utilization across a national sample of pediatric emergency departments. J Pediatr, 2013. 163(1): p. 230-6.
2.Adams, J.G., Emergency department overuse: Perceptions and solutions. JAMA, 2013. 309(11): p. 1173-1174.
3.Institute, N.E.H., A Matter of Urgency: Reducing Emergency Department Overuse. Research Brief, 2010(March).
4.Perrin, J.M., et al., Variations in rates of hospitalization of children in three urban communities. N Engl J Med, 1989. 320(18): p. 1183-7.
5.Perrin, J.M., et al., Primary care involvement among hospitalized children. Arch Pediatr Adolesc Med, 1996. 150(5): p. 479-86.
6.DeAngelis, C., P. Fosarelli, and A.K. Duggan, Use of the emergency department by children enrolled in a primary care clinic. Pediatr Emerg Care, 1985. 1(2): p.61-5.

7.Berns, S.D., et al., Appropriate use of a pediatric emergency department: is the pediatrician called before the visit? Pediatr Emerg Care, 1994. 10(1): p. 13-7.
8.Rudowitz, R., A Look At CBO Projections For Medicaid and CHIP, in The Kaiser Commission on Medicaid and the Uninsured. 2014, The Henry J. Kaiser Family Foundation
9.Taubman, S.L., et al., Medicaid Increases Emergency-Department Use: Evidence from Oregon´s Health Insurance Experiment. Science, 2014. 343(6168): p. 263-268.
10.Smulowitz, P.B., et al., Increased Use of the Emergency Department After Health Care Reform in Massachusetts. Ann Emerg Med, 2014.
11.Brook, R.H., et al., A method for the detailed assessment of the appropriateness of medical technologies. Int J Technol Assess Health Care, 1986. 2(1): p. 53-63.
12.Park, R.E., et al., Physician ratings of appropriate indications for six medical and surgical procedures. Am J Public Health, 1986. 76(7): p. 766-72.
13.Fitch, K., et al., The RAND/UCLA Appropriateness Method User´s Manual. 2001 RAND.
14.Kosecoff, J., et al., The appropriateness of using a medical procedure. Is information in the medical record valid? Med Care, 1987. 25(3): p. 196-201.
15.Kemper, K.J., Medically inappropriate hospital use in a pediatric population. N Engl J Med, 1988. 318(16): p. 1033-7.
16.Kleinman, L.C., et al., The medical appropriateness of tympanostomy tubes proposed for children younger than 16 years in the United States. Jama, 1994. 271(16): p. 1250-5.
17.Kleinman, L.C., E.A. Boyd, and J.C. Heritage, Adherence to prescribed explicit criteria during utilization review. An analysis of communications between attending and reviewing physicians. Jama, 1997. 278(6): p. 497-501.
18.Keyhani, S., et al., Electronic health record components and the quality of care. Med Care, 2008. 46(12): p. 1267-72.
19.Keyhani, S. and A.L. Siu, The underuse of overuse research. Health Serv Res, 2008. 43(6): p. 1923-30.
20.Bernstein, S.J., et al., The appropriateness of hysterectomy. A comparison of care in seven health plans. Health Maintenance Organization Quality of Care Consortium. Jama, 1993. 269(18): p. 2398-402.
21.Taylor, A.J., et al., ACCF/SCCT/ACR/AHA/ASE/ASNC/NASCI/SCAI/SCMR 2010 Appropriate Use Criteria for Cardiac Computed Tomography. A Report of the American College of Cardiology Foundation Appropriate Use Criteria Task Force, the Society of Cardiovascular Computed Tomography, the American College of Radiology, the American Heart Association, the American Society of Echocardiography, the American Society of Nuclear Cardiology, the North American Society for Cardiovascular Imaging, the Society for Cardiovascular Angiography and Interventions, and the Society for Cardiovascular Magnetic Resonance. J Cardiovasc Comput Tomogr, 2010. 4(6): p. 407.e1-33.
22.Basger, B.J., T.F. Chen, and R.J. Moles, Validation of prescribing appropriateness criteria for older Australians using the RAND/UCLA appropriateness method. BMJ Open, 2012. 2(5).

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 (4b) under Usability and Use.*

In testing we found that for children age 2-5, 181 of 335 (54.3%) were deemed appropriate, with the breakdown of reasons for appropriateness presented in the Testing Form 2b.2.2. Other age groups found that children 6-11, 209 of 447 (43.8%) ED visits were appropriate, while for adolescents aged 12-18, 165 of 341 (48.4%) visits were appropriate. These numbers were sufficient to identify statistically significant differences in the proportion that were appropriate between age groups, among racial/ethnic groups, and within age group among racial/ethnic groups. These data demonstrate that the specified sample size is sufficient to find meaningful differences between groups at the various specified levels. In our work, validating and testing the measure for the rate of appropriateness, we have demonstrated the capacity to identify the included events (ED visits and hospitalizations) using administrative data and our specifications for identifiable asthma. That aspect of testing was conducted using state wide data from the NY State Medicaid Managed Care Program.

We also incorporate by reference work done by our partner NCQA that demonstrate the capacity of administrative data to identify a population with asthma.

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.**

In her seminal article nearly three decades ago, DeAngelis included an asthma attack as an appropriate indication for use of the ED.[1] As a common chronic illness characterized by remissions and potentially preventable exacerbations undesirable utilization outcomes for asthma have been a frequent target for measurement for three decades. Reducing the relative number of ED visits during the care for asthmatic children remains a high priority on the national agenda. The universal delivery of optimal asthma care has the potential to lower costs and improve quality of life. Understanding which ED visits represent failures of clinical prevention and which instead represent a mismatch of service level to clinical need can help to move these goals forward. The submitted measure is a step in this direction.

ED visits for asthma can be reduced through both enhanced access to care and through better quality of care. The NIH's National Asthma Education and Prevention Program Guideline has been shown to reduce the frequency of breakthrough asthma and of ED visits and hospitalizations when implemented. The literature points to two general characteristics of asthma care delivery systems that correlate with ED utilization. One is the effective use of preventive and routine care measures, such as multidisciplinary practice or a medical home model, the presence of an asthma action plan, the use of controller medications supplemented by judicious use of rescue medications. [2-6] The other is the availability of primary care or urgent care visits as a step before ED use in the context of either a general pediatric or an asthma specialty practice. [6, 7] Conversely, a lack of comprehensive asthma care, which includes primary and secondary prevention schemas, and a lack of available urgent care services are both commonly cited as reasons for preventable ED visits. It has been demonstrated that the children who used the ED underutilized primary care services [6] and it has also been demonstrated that interventions that attempt to provide comprehensive, multidisciplinary care are able to decrease ED utilization for asthma care.[8] We acknowledge that environmental management and control is a nonclinical opportunity to improve the quality of life for children with asthma and to reduce health care utilization, but do not focus on these issues in this submission.

High rates of asthma visits to the ED suggest widespread deficiencies in asthma care. The literature shows that lack of proper asthma care is disparate with minority children bearing undue burden. [9-11]

The literature also presents different perspectives on appropriate use of the ED for pediatric asthma. Pediatric asthma is one of the leading conditions when it comes to potentially avoidable ED visits. [12] Asthma has been classified both as an avoidable hospitalization condition (AHC) and as an ambulatory care sensitive condition. This describes that a meaning proportion of ED visits or hospital admissions could have been avoided with proper outpatient care. [12, 13] Poor outpatient care can be an outcome of a number of variables. As noted, the availability of primary care can reduce such inappropriate and costly visits. [7, 12, 14 -17]

Assessing the extent to which ED use for asthma is appropriate can inform health policy, manpower planning, and clinical quality improvement activities. It can help to answer the question of how much of ED use potentially may be prevented by better management of the underlying asthma, versus how much requires other, process or structural improvements to reduce use of the ED when a lower level of care would meet the clinical needs of the child. Refractory asthma or those with unavoidable environmental exposures leading to an acute exacerbation requiring medical care are likely to be identified as appropriate, reminding us that NOT all asthma ED visits are preventable even with optimal care.

With a better understanding of ED use, health care organizations and policy makers could develop better informed approaches to optimizing services for children with asthma. And hopefully children and their families may increasingly be spared the inconvenience, risk, and costs of ED visits for asthma.

References:

1. DeAngelis, C., P. Fosarelli, and A.K. Duggan, Use of the emergency department by children enrolled in a primary care clinic. Pediatr Emerg Care, 1985. 1(2): p.61-5.
2. Talreja, N., et al., Modifiable factors associated with severe asthma exacerbations in urban patients. Ann Allergy Asthma Immunol, 2012. 109(2): p.128-32.
3. Auger, K.A., et al., Medical home quality and readmission risk for children hospitalized with asthma exacerbations. Pediatrics, 2013. 131(1): p. 64-70.
4. Ducharme, F.M., et al., Written action plan in pediatric emergency room improves asthma prescribing, adherence, and control. Am J Respir Crit Care Med, 2011. 183(2): p. 195-203.
5. Farber, H.J., Optimizing maintenance therapy in pediatric asthma. Curr Opin Pulm Med, 2010. 16(1): p. 25-30.
6. Smith, S.R., D.B. Wakefield, and M.M. Cloutier, Relationship between pediatric primary provider visits and acute asthma ED visits. Pediatr Pulmonol, 2007. 42(11): p. 1041-7.
7. Parchman, M.L. and S. Culler, Primary care physicians and avoidable hospitalizations. J Fam Pract, 1994. 39(2): p. 123-8.
8. Prevention, C.f.D.C.a. Home-based Multi-trigger, Multi-component interventions. 2013 [cited 2013 May 20]; Available from: http://www.cdc.gov/asthma/interventions.htm.
9. Price, J.H., et al., Racial/ethnic disparities in chronic diseases of youths and access to health care in the United States. Biomed Res Int, 2013. 2013: p. 1-12; Available in open access http://dx.doi.org/10.1155/2013/787616
10. Homer, C.J., et al., Does quality of care affect rates of hospitalization for childhood asthma? Pediatrics, 1996. 98(1): p. 18-23.
11. Finkelstein, J.A., et al., Quality of care for preschool children with asthma: the role of social factors and practice setting. Pediatrics, 1995. 95(3): p. 389-94.
12. Flores, G., et al., Keeping children out of hospitals: parents´ and physicians´ perspectives on how pediatric hospitalizations for ambulatory care-sensitive conditions can be avoided. Pediatrics, 2003. 112(5): p. 1021-30.
13. Knudson, A., et al., Disparities in pediatric asthma hospitalizations. J Public Health Manag Pract, 2009. 15(3): p. 232-7.
14. Bindman, A.B., et al., Preventable hospitalizations and access to health care. Jama, 1995. 274(4): p. 305-11.
15. National Heart, Lung and Blood Institute. Asthma Guidelines. 2011 February 2011 [cited 2014 7/30/2014].
16. A Matter of Urgency: Reducing Emergency Department Overuse in A NEHI Research Brief 2010, New England Healthcare Institute.
17. Martin, B.C., Emergency medicine versus primary care: a case study of three prevalent, costly, and non-emergent diagnoses at a community teaching hospital. J Health Care Finance, 2000. 27(2): p. 51-65.

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 (4b) under Usability and Use.*

Race/Ethnicity

Our medical chart audit found that the measure varies by race/ethnicity. Hispanic children had higher rates of questionable use of the ED (55.9% of visits) when compared to non-Hispanic children (46.8%), p=.002. Black children showed a trend toward more questionable use compared to all other children (53.6% questionable vs 48.7%, p=.10). Overall, Blacks had an appropriate use rate of 51.3%, Whites 56.5%, Hispanics 44.1% and other races 45.2%. For ages 2-5, Blacks had an appropriate use rate of 57.1%, Whites 63.6%, Hispanics 50.9%, and other races 51.9%. For ages 6-11, Blacks had an appropriate use rate of 49.3%, Whites 50.0%, Hispanics 46.3% and other races 39.8%. For ages 12-18, Blacks had an appropriate use rate of 49.3%, Whites 66.7%, Hispanics 46.3% and other races 46.7%. Chi-square analysis confirms that these differences are statistically significant.

Insurance Status
Overall, the appropriate use rate for Medicaid patients was 46.3%, Private insurance 59.0%, Uninsured patients 38.6% and other forms of insurance (military and Worker's comp) 55.0% (p=.005). Within the age strata, for ages 2-5 the appropriate use rate for Medicaid patients was 53.9%, Private 67.4%, Uninsured 40.9% and other was 20%. For ages 6-11, the appropriate use rate for Medicaid patients was 41.5%, Private 57.7%, Uninsured 35.7% and other was 52.6%. For ages 12-18, the appropriate use rate for Medicaid patients was 46.1%, Private 54.5%, Uninsured 42.1% and other was 68.8%. Chi-square analysis demonstrates the presence of statistically significant differences.

Socioeconomic Status
The measure is specified to be stratified in 2 ways to assess aspects related to socioeconomic status: Public versus Commercial Insurance, and by 5 strata defined by the percent of the population in poverty in their county of residence.

Rurality/Urbanicity
These measures are specified to be reported by Urban Influence Codes (UIC), which have been developed by the USDA based on a number of criteria to describe the levels of urbanicity and rurality. This is intended not only to report within plan differences but to allow for aggregation as appropriate. While each UIC has its own meaningful definition, some researchers choose to aggregate various codes. We recommend consideration of the aggregation schema of Bennett and colleagues at the South Carolina Rural Research Center. (2) Their aggregation scheme brings together Codes 1 & 2 as Urban; 3, 5, & 8 as micropolitan rural; 4, 6, & 7 as rural adjacent to a metro area; and 9, 10, 11, & 12 as remote rural. We observe that UIC 5 might as well be aggregated with 4, 6, & 7 as an adjacent rural area. Further, while this approach to rurality does not map exactly to the population density based definition of frontier (< 6 persons per square mile) as articulated in the Affordable Care Act, use of such categories is consistent with the ACA's intent that the Secretary ask that data that are collected for racial and ethnic disparities also look at underserved frontier counties. Frontier health care may be approximated by analysis of the remote rural categories. (3)

This judgment was confirmed after CAPQuaM consulted with Gary Hart, Director of the Center for Rural Health at the University of North Dakota School of Medicine & Health Sciences, who is heading a HRSA-funded project to develop new methods to analyze frontier health. We clarified that his work suggests that UIC 9-12 is the best overall approach to using county level data to study frontier health. Inclusion of UIC 8 would make the analysis more sensitive to including frontier areas but at a meaningful cost in sensitivity.

Those interested in care specific to large cities may wish to aggregate the rural area and analyze UIC 1 and 2 separately. Frontier health care may be approximated by analysis of the remote rural categories. (3) The New York State Medicaid data were sensitive to urbanicity with higher rates of ED utilization in the most urban areas and lowest in the most rural areas and other areas intermediate between the two.

For aggregation and as an imperfect approximation one can also group as urban (1 and 2), suburban (3-6) and rural (7-9). This is what we have used for our NY Medicaid analysis to demonstrate that variations are observed for this measure using UIC codes.

1.Kawachi I, B.L., Neighborhoods and Health. 2003, New York, NY: Oxford University Press.
2.Bennett, K.J., Olatosi B. & Probst, J.C., Health Disparities: A rural-urban chartbook. 2008, Columbia, South Carolina: South Carolina Rural Health Research Center.
3.Hart, G., Frontier/Remote, Island, and Rural Literature Review. 2012.

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
Performance data provided in 1b.4

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):*
Respiratory : Asthma

De.6. **Cross Cutting Areas** *(check all the areas that apply):*
«crosscutting_area»

De.7. **Target Population Category** *(Check all the populations for which the measure is specified and tested if any):*
Children, 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.)*
We will create a webpage as soon as possible, likely early in the new year.

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:** Asthma_5_12_06_16.xlsx

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

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

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).
The numerator is the number of eligible asthma ED visits in the random sample that also satisfy at least one of the explicit criteria to indicate that the ED is an appropriate level of care. Distinct numerators are reported for children ages 2-5, 6-11, 12-18, and optionally, 19 - 21.

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).
Children and adolescents who have a qualifying ED visit associated with asthma as the first or second diagnosis and are in the random sample;

AND have at least one of the following:

•Disposition of the ED visit was admission to the hospital,
OR
•Documented physical findings consistent with respiratory distress, including any of the following:
o Labored breathing (including moderate or severe increased work of breathing); OR
o Retractions, grunting, and/or evidence of accessory muscle use; OR
o Markedly decreased breath sounds;
OR
•Recorded oxygen saturation below 90%;
OR
•An arterial blood gas (ABG) was obtained in the emergency department;
OR
•The child had a consultation with a pulmonologist or asthma specialist that was ordered and provided in the ED;
OR
•There is clear documentation that prior to arrival in the ED any of the following occurred:
o The child was referred to the ED after evaluation by the PCP or other clinician. The evaluation may include an in person visit or auscultation including via telephone. OR
o The child received two or more doses of inhaled rescue medications without sufficient clinical improvement. Documentation of parent report meets the criterion. OR
o The child was assessed with an objective instrument such as a peak flow meter and was found to be in a pre-defined "red zone" of peak flow measurement as part of an asthma action or similar plan. Documentation requires ALL of the following: a written asthma action plan exists AND defines a "red zone" for which urgent assessment by a clinician is indicated AND an objective assessment was made and its result was in the pre-defined red zone. Documentation of parent report meets the criterion.

S.6. **Denominator Statement** *(Brief, narrative description of the target population being measured)*
The denominator is a random sample of the patients in each age stratum who have visited the emergency department for asthma (as a first or second diagnosis) and meet the specified criteria for having identifiable asthma (defined in s2b).

Separate numerators and denominators are reported for children age 2-5, 6-11, 12-18, and, optionally, 19-21 years. An overall rate across strata is not reported.

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).

Denominator Elements:

The presence of identifiable asthma (see table 1) is established each month from administrative data using the specified algorithm.

Descriptive definitions for being managed for identifiable asthma are as follows. Specifications follow the descriptive definitions. Identifiable asthma is present in any child who has:

- Any prior hospitalization with asthma as primary or secondary diagnosis;

OR

- Other qualifying events, all ages:

o Three or more ambulatory visits with diagnosis of asthma or bronchitis, OR
o Two or more ambulatory visits with a diagnosis of asthma and/or bronchitis AND one or more asthma-related prescriptions

OR

- Other qualifying events, occurring after the fifth birthday:

o One or more prior ambulatory visits with asthma as the primary diagnosis AND a subsequent ED visit in the Reporting Month,
OR
o Two or more ambulatory visits with asthma as a diagnosis,
OR
o One ambulatory visit with asthma as a diagnosis AND at least one asthma-related prescription,
OR
o Two or more ambulatory visits with a diagnosis of bronchitis

For eligibility purposes, asthma-related medicine means long-acting beta-agonist (alone or in combination) or inhaled corticosteroid (alone or in combination), anti- asthmatic combinations, methylxanthines (alone or in combination), and/or mast cell stabilizers. See below further regarding this specification. Note that leukotriene modifiers and short term beta agonists are excluded for the purpose of establishing identifiable asthma. Data from the year prior to the reporting year are used, as well as all months prior to the reporting month in the reporting year (see Appendix Figure 1). Detailed specifications for asthma related medicine can begin with the NCQA NDC list (ASM-C_DASM-C_final_2012, found by clicking through at (http://www.ncqa.org/HEDISQualityMe asurement/HEDISMeasures/HEDIS20 12/HEDIS2012FinalNDCLists.aspx) Eliminate medications in the following 2 categories: leukotriene modifiers, short-acting inhaled beta-2 agonists). May use equivalent updated lists when provided by NCQA. Even if included in NCQA list, we further exclude indacaterol, a recently approved long acting beta agonist that is indicated in the US only for the treatment of COPD.

The analysis should be conducted on a month by month basis as described herein: Within the group of children who meet the criteria for identifiable asthma, identify and maintain a unique patient identifier, age, and all stratification variables. We call the time frame during which eligibility is established to be the Assessment Period.

For each month of the Reporting Year, determine eligibility for each patient, as of the last day of the month prior to the reporting month. This illustration assumes that the Reporting Year is 2011. When assessing January 2011, consider all of Calendar Year 2010 as the Assessment Period for assessing the presence or absences of identifiable asthma. For February, 2011 the Assessment Period includes all of calendar year 2010 AND January 2011. Repeat this progression monthly so that for December, 2011 identifiable asthma one would identify children with identifiable asthma using an Assessment Period from January 2010 through November 2011. For each month, assess whether the continuous enrollment criterion is met prior to including the month in the denominator. For example, for January 2011, the child must have been enrolled in November and

December, 2010 (plus January 2011). Another example, for December 2011, to be eligible the child must have been enrolled in October 2011 and November 2011, as well as December. See Figure 1 in Appendix.

Develop Denominator sample according to Appendix Figure 2 and consistent with the instructions in sections S.18 and S.20.

Codes used for definitions are specified in s2b and include specifications of Hospitalization, Emergency Department Visits, Ambulatory/Office Visits, Asthma diagnosis,

Please note Figures 1 and 2 in the Appendix and the specifications in s2b are considered INTEGRAL to these specifications and are not optional.

S.8. **Denominator Exclusions** *(Brief narrative description of exclusions from the target population)*
ED visits that are already in the sample OR Children that fall outside of specified age range of 2-21 OR who do not meet time enrollment criteria OR do not meet identifiable asthma prior to the ED visit, OR children with concurrent or pre-existing COPD, Cystic Fibrosis or Emphysema.

At the discretion of the accountability entity, the denominator may be restricted to children 2-18.

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.)*
Denominator Exclusions
1) Children with concurrent or pre-existing:
a. Chronic Obstructive Pulmonary Disease (COPD) diagnosis;
OR
b. Cystic Fibrosis diagnosis;
OR
c. Emphysema diagnosis.

OR
2) Children without identifiable asthma as specified
OR
3) Outside of specified age range
OR
4) Events occurring in patients who have not been enrolled in the reporting plan for at least two consecutive months before the index reporting month (a total of 3 consecutive months, including the reporting month).

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.)*
This measure requires stratification by age group. Several additional stratifications are optional but may be required by the accountability entity or provided by the reporting entity. These variables include race/ethnicity, rurality/urbanicity and county level of poverty.

Stratify by age group (reporting entity should specify whether to use age at month of qualifying event or age on first day of reporting year):
•Age 2-5 years (second birthday to the day before the 6th birthday);
•Age 6-11 years (sixth birthday to the day before the 12th birthday);

•Age 12-18 years (twelfth birthday to the day before the 18th birthday); and
•Age 19-21 years (nineteenth birthday to the day before the 21st birthday).

Age strata are to be reported distinctly and not combined.

Optional stratifications require data elements such as:
•Race/Ethnicity
•Insurance type (Public, Commercial, Uninsured)
•Benefit type (if insured): HMO, PPO, Medicaid Primary Care Case Management (PCCM) Plan, Fee for Service (FFS), other
•Zip code, state and county or equivalent area of parent/caregiver's residence. Record FIPS if available

Stratification variables details
•Race/Ethnicity: Hispanic, Non-Hispanic Black, Non-Hispanic White; Non- Hispanic Asian/Pacific Islander, other Non-Hispanic
•Public vs Commercial (Private Insurance).
•HMO vs PPO vs FFS vs PCCM vs other; Within Medicaid, States may ask for reporting of FFS vs Managed Care or other relevant enrollment categories (e.g., TANF, SSI).
•Urban Influence Code. Identify the Urban Influence Code or UIC. (2013 urban influence codes available at: http://www.ers.usda.gov/data-products/urban- influence-codes.aspx#.UZUvG2cVoj8). Use parent or primary caregiver's place of residence to determine UIC. State and county names can be linked or looked up directly or zip codes can be linked to county indirectly, using the Missouri Census Data Center (http://mcdc.missouri.edu/). These data will link to county or county equivalents as used in various states.

Urban Influence Codes (UIC) have been developed by the USDA to describe levels of urbanicity and rurality. While each UIC has its own meaningful definition, some researchers choose to aggregate various codes. Well regarded schemas for aggregation of codes include Bennett and colleagues at the South Carolina Rural Research Center. Their aggregation scheme brings together Codes 1 & 2 as Urban; 3,5, & 8 as micropolitan rural; 4,6, & 7 as rural adjacent to a metro area; and 9, 10, 11, & 12 as remote rural. We acknowledge that UIC 5 (adjacent rural area) may appropriately be aggregated with 4,6,&7 as rural. Frontier health care may be approximated by analysis of the remote rural categories (UIC 9, 11 and 12). Alternatively, Gary Hart, Director of the Center for Rural Health at the University of North Dakota School of Medicine & Health Science suggests that UIC 9-12 is the best overall approach to using county level data to study frontier health. Inclusion of UIC 8 would make the analysis more sensitive to including frontier areas but at a meaningful cost in specificity.

Those interested in care specific to large cities may wish to aggregate the rural area and analyze UIC 1 and 2 separately.

When stratifying by urbanicity or UIC, the reporting and accountability entities should specify clearly what if any aggregating schema was used.

•Identify the Level of Poverty in the parent or primary caregiver's county of residence. The percent of all residents in poverty by county or county equivalent are available from the US Department of Agriculture at http://www.ers.usda.gov/data-products/county-level-data-sets/download- data.aspx. Our stratification standards are based on 2011 US population data that we have analyzed with SAS 9.3. Using parent or primary caregiver's state and county of residence (or equivalent) or FIPS code, use the variable PCTPOVALL_2011 to categorize into one of 5 Strata:
o Lowest Quartile of Poverty if percent in poverty is <=12.5%
o Second Quartile of Poverty if percent in poverty is >12.5% and <=16.5%
o Third Quartile of poverty if percent in poverty is >16.5% and <=20.7%
o First Upper Quartile (75th-90th) if percent in poverty is >20.7% and <=25.7%
o Second Upper Quartile (>90th percentile)

These classification standards may be updated by the accountability entity suing more recent data if desired.

Note: if needed, the Missouri Census Data Center may be used to link zip codes to county equivalents. http://mcdc.missouri.edu/

S.11. **Risk Adjustment Type** (Select type. Provide specifications for risk stratification in measure testing attachment)
Stratification by risk category/subgroup
If other:

S.12. **Type of score:**
Rate/proportion
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 = Higher score

S.14. **Calculation Algorithm/Measure Logic** (*Diagram or describe the calculation of the measure score as an ordered sequence of steps including identifying the target population; exclusions; cases meeting the target process, condition, event, or outcome; time period for data, aggregating data; risk adjustment; etc.)*
Step 1: Select starting cohort
Identify the upper age limit to be used, either 18 or 21. The measure is specified from 2 to 21 years, with 19-21 year olds considered optional at the discretion of the accountability entity.

Appendix Figures 1 and 2 and Data Codes in S2b provide an overview and guide for eligibility and sample selection.

Step 2: Conduct analysis of administrative data using the specifications described in denominator description to identify children within the specified age range with identifiable asthma. The analysis should be conducted on a month by month basis as described herein:

Determine eligibility for each patient, as of the last day of the month prior to the reporting month. For example, if the goal is to report for January 2011, first identify children with identifiable asthma (above), and analyze all of calendar year 2010 when doing so. Continuous enrollment criterion requires that the child was enrolled in November and December of 2010. Next, for February analyze all of calendar year 2010 AND January 2011. Continuous enrollment criterion requires that the child was enrolled in December 2010 and January 2011. Repeat this progression monthly so that for December, one would identify children with identifiable asthma and analyze all of calendar year 2010 AND January through November 2011 when doing so. Continuous enrollment criterion requires that for December the child was also enrolled in October 2011 and November 2011. Appendix Figure 1 describes and illustrates the month by month analysis.

Step 3: Identify ED Visits and hospitalizations for asthma in eligible children.
Considering only the children who were identified as eligible in the given month
according to Step 2, perform a month-by-month analysis to identify and log all ED visits with asthma as a primary or secondary diagnosis and all hospitalizations with asthma as a primary or secondary diagnosis for each reporting month, using specifications described in denominator and the codes described above and in S2b. Maintain stratification data elements, age, and unique identifiers.

Step 4: Stratify by age and develop random samples.
Stratify by age group (use age at month of qualifying event):

- Age 2-5 years (second birthday to the day before the 6th birthday);
- Age 6-11 years (sixth birthday to the day before the 12th birthday);
- Age 12-18 years (twelfth birthday to the day before the 18th birthday); and
- Age 19-21 years (nineteenth birthday to the day before the 21st birthday).

For each age group develop a random sample of 500 events as described in the sampling section below and illustrated in Appendix Figure 2.

Appendix Figure 2 is necessary to guide sample development. Several key remarks may help Figure 2 to be more understandable:

Before sample selection can be randomized, eligibility needs to be determined based on 3 key factors:
- Identifiable asthma diagnosis AND
- Month by month time analysis AND
- Asthma emergency department (ED) visit OR Asthma hospitalization

After eligibility is determined, the randomized sample can fall into one of three groups only:
A. Asthma ED visit only OR
B. Asthma hospitalization on same day as ED visit OR
C. Asthma hospitalization only

A. Asthma ED visit only qualifies for (at least) denominator inclusion
B. Asthma hospitalization on same day as ED visit qualifies for denominator AND numerator inclusion
C. Asthma hospitalization only needs further investigation to determine denominator inclusion
. • Do NOT include in denominator
-- if sample was not hospitalized (admitted) from an asthma ED Visit, OR
-- if ED visit was already in the sample under any criteria (avoid duplication)
ELSE,
. • Do include in both Denominator AND Numerator

Step 5: Collect stratification data elements from administrative data.
Collect the following data elements for all eligible children in each randomized sample. These data elements are used for reporting stratified results. Entities that are interested in assuring large samples for specific stratified analyses may choose to incorporate a further stratified sampling scheme and oversample to assure that there is a sample size of 100-500 per stratification category (e.g. race or ethnicity of interest). Such a sampling scheme must employ an appropriate weighting system (using the reciprocal of the likelihood for selection as a weight, c.f. Rao, P., 2000. Sampling Methodologies with Applications. New York: Chapman & Hall) to estimate overall performance. Alternatively, the stratified samples may be used only for reporting stratum specific performance comparison and not for estimating the overall performance. Approximate 95% confidence interval widths (assuming a rate of 50% appropriateness) are shown in the sampling specifications. We specify to oversample by 25% to account for potential loss in our event identifications.

Stratification data elements include:
- Race
- Ethnicity
- Insurance type (Public, Commercial, Uninsured)
- Benefit type (if insured): HMO, PPO, Medicaid Primary Care Case Management (PCCM) Plan, Fee for Service (FFS), other
- Zip code, state and county or equivalent area of parent/caregiver's residence. Record FIPS if available

Step 6: Categorize stratification variables as described in the stratification section S.12.

Step 7. Conduct Chart Audit (Medical Record Review) of GROUP A ED Visits.
Group A ED visits that have been selected for inclusion in the sample require a chart audit to assess eligibility for the numerator based on the explicit appropriateness criteria. They have already qualified for inclusion in the denominator. Eligibility for the numerator is established based on documentation of any of the following items. Review may be terminated once any qualification for the numerator is identified.

- Disposition of the child from the ED was admission to an inpatient hospital, OR
- Documented physical findings consistent with respiratory distress, including:
  - o Labored breathing with retractions and/or grunting; or
  - o Labored breathing with evidence of accessory muscle use; or,
  - o Markedly decreased breath sounds; OR
- O2 saturation level below 90% documented in the ED; OR
- An ABG obtained and reported in the ED; OR
- The child had a consultation with a pulmonologist or asthma specialist that was ordered and provided in the ED; OR
- Specific documentation that:
  - o The child was referred to the ED after evaluation by the PCP or other licensed clinician practitioner; OR
  - o The child received two or more doses of inhaled rescue medications without sufficient clinical improvement; OR
  - o The child was assessed with an objective instrument such as a peak flow meter and was found to be in a pre-defined “red zone” of peak flow measurement as part of a pre-specified asthma action or similar plan.

There is no specified order for review. Some institutions may prefer to record all reasons for numerator qualification to support ongoing or planned improvement activities.

Note 1: Evidence for hospitalization above requires that the child was admitted to any hospital as an inpatient. This includes admission directly to a medical or pediatric ICU or inpatient floor or transfer directly to an inpatient facility. If a child is transferred to another hospital, confirmation that the child actually was admitted directly (i.e., was not first admitted to another ED prior to admission) is necessary prior to qualifying for the numerator. Such confirmation may include evidence from the administrative data review in Step 2. Other potential sources for this information include ED discharge summary, disposition on a flow, admit, or discharge form, or documentation by doctors, nurses, nurse practitioners or physician assistants.

Note 2: Evidence that the child was referred to the ED requires documentation of both of two requirements. The requirements are:

- The child/adolescent was referred by a clinician to come to the ED; and
- The child/adolescent was evaluated by the clinician prior to referral. Generally such evaluations will be in person. Assessment of respiratory distress by listening or speaking to the child/adolescent over the telephone is sufficient if such an examination is clearly documented. Report of this requirement being met by the child/adolescent or parent/caregiver is sufficient to meet this criterion. Report of contact from the referring physician can also fulfill this criterion. Nursing notes, triage notes and clinician notes, particularly history of present illness (HPI) are common sources for this data.

Note 3: Evidence of a parent or caregiver report that the child received two or more doses of an inhaled rescue medication with insufficient clinical improvement typically will be found in triage, nursing, clinician, or respiratory therapy notes. It may also be documented as a part of medication reconciliation during intake. It requires documentation:

- That multiple treatments of medication were provided by inhalation or injection prior to arrival in the ED;
- That the medication(s) provided were specifically rescue medications and are not a part of the of the child/adolescent’s preventive or maintenance regimen; and,

• That the child continued to be in distress following the treatments (alternately that the child did not improve substantially).

Note 4: Parent / caregiver report that their child was in a pre-defined "red zone" of peak flow measurement includes documentation:

• That a pre-specified asthma plan (action plan) exists and defines a "red zone" based upon an objective respiratory measurement, such as a peak flow rate; and

• That the objective assessment was made prior to coming to the ED and that the results were in the pre-specified "red zone."

Note 5: Reports of the physical exam typically may be found on triage, nursing, physician, nurse practitioner, physician assistant, or respiratory therapist notes. Diverse language may be used to describe similar findings, for example:

• The term pulling may be used to describe retractions. Retractions may be described as nasal flaring (particularly in infants), or by location (see below);

• Increased work of breathing may be indicated or it may be described by physical findings such as the use of accessory muscles, such as sub or intercostal muscles, supraclavicular or suprasternal. "Mildly" increased work of breathing or "minimal" retractions do not meet these criteria.

• Labored breathing, significant increased work of breathing, respiratory distress (moderate or greater), difficulty breathing, poor air entry (or air exchange or air movement) may all describe findings that meet this criterion. Grunting indicates that the child or adolescent is generating clearly audible sounds with each breath concomitant with apparent increased work of breathing. These may be found in the general description or respiratory section of the physical exam.

• Markedly (or severely) reduced breath sounds and descriptions of poor air movement are typically a part of an auscultation during the pulmonary exam.

Note 6: Documented evidence of the percent oxygen (O2) saturation from a transcutaneous assessment can be located in a flow sheet, nursing, respiratory therapy, or physician/nurse practitioner/physical assistant note or may be recorded as part of the physical exam. The O2 saturation may be obtained initially at triage and is often assessed periodically during the visit. Any O2 saturation less than 90 satisfies the criteria.

Note 7: An ABG requires drawing of a blood specimen from an artery and is distinguished from a venous blood gas, which would not fulfill this criterion. This typically would be found in a laboratory results section of the record or commented as a finding in a clinician's note, such as a respiratory therapist, doctor, PA, NP, or RN. An ABG is typically comprised of at least a pO2, pCO2, and pH.

Note 8: Consultation with a pulmonary specialist or other asthma specialist requires both an order for such a physician consultation and evidence that the consultation occurred, including a note from the consultant specialist. Typically a consultation from a pulmonologist, pediatric pulmonologist, allergist, or pediatric allergist would fill this criterion.

Identify which ED visits meet at least one criterion for the Numerator.
Maintain stratification variables.

Step 8: Conduct Chart Audit (Medical Record Review) to Assess Eligibility of GROUP C Hospitalizations for Inclusion in Denominator.
Within each stratification group (as determined above), identify the asthma hospitalizations for which there were not associated ED visits (Group C) found in the administrative data. An asthma ED visit and asthma hospitalization are said to be associated on the basis of the administrative data review only if they occur on the same service data and at the same institutions and if the hospital discharge date is after the ED service date. Such hospitalizations should have been included in Group B. Other hospitalizations require a review of the

medical record to determine if they were admitted or transferred directly from an ED visit that was not otherwise in the sample (i.e., was not identified via the administrative data analysis).

The chart audit/medical record review seeks evidence that the child was admitted to the hospital directly from the ED or transferred directly from another hospital's ED. Evidence may include an ED note (physician, nurse, physician assistant, nurse practitioner), flow, or face sheet that indicates the disposition of the ED visit was hospital admission.

It may also include a note from within the hospitalization (including the admission note or any physician, nurse, physician assistant, nurse practitioner note), flow sheet, face sheet, or discharge summary that indicates that the hospitalization came directly from (was admitted from or transferred directly from) an ED. In either case, the ED visit is only eligible for inclusion if the chart review specifies the date and institution of the ED visit sufficiently to assure that it can be uniquely identified and all duplication avoided. Others are excluded.

For example if an ED visit was identified in Group A and the resulting hospitalization appeared in Group C (either because of a different service date or different institution), the Group A ED visit would be included and the Group C hospitalization excluded as a duplicate (even though there was a preceding ED visit). If the child is uniquely included in the sample for that month and there is clear evidence that the admission came directly from an ED (e.g., was not transferred from another hospital after having been admitted from the ED) this measure can be satisfied. De-duplication requires the elimination of any duplications that remain in the sample, considering the unit of analysis to be the ED visit. In other words, all ED visits must be included only once. Further, an ED visit identified via the hospitalization that also was a transfer from another ED visit already in the sample should have been removed as a duplicate. Similarly all hospitalizations lacking sufficient document that the child was admitted or transferred directly from an ED visit or lacking sufficient detail to allow confirmation that the ED visit referred to in the notes is not already in the sample elsewhere (e.g., from Group A) should have been removed.

Those Group C hospitalizations that can be identified as resulting from a unique (unduplicated) ED visit are included in BOTH the numerator and the denominator.

Step 9: Calculate and report the measure.
a) For each age stratum, count the number of events in the sample that qualify for the denominator (ND).
b) For each age stratum, count the number of events in the sample and in the denominator that qualify for the numerator (NN).
c) For each stratum, calculate the percent of appropriate ED visits as Percent Appropriate = 100 * (NN / ND). Report to one decimal place.

Step 10: Report each stratification category listed below, that have an N of at least 50.
a) Race and ethnicity
b) Insurance type (Public/Medicaid, Private/Commercial, None, other)
c) Benefit type: HMO vs PPO vs FFS vs PCCM vs other
d) Urban Influence Code or UIC.
e) Level of poverty in the county of residence.

Step 11. Calculate and report 95% confidence intervals (using binomial distribution for each stratum) for each age specific stratum and for all of the Step 9 stratifications.

a) Calculate the standard error as the square root of each proportion by [1-the same proportion] divided by the number in the denominator.
b) Multiply the standard error by 1.96.
c) Subtract that value from the measured proportion. Report the greater of 0 and that number as the lower bound of the 95% confidence interval.

d) Add the product from b to the measured proportion. Use the lesser of that sum or 1 as the upper bound of the 95% confidence interval.
e) To report as percent, multiply by 100.

S.15. **Sampling** *(If measure is based on a sample, provide instructions for obtaining the sample and guidance on minimum sample size.)*
IF a PRO-PM, identify whether (and how) proxy responses are allowed.
Children and adolescents who have a qualifying ED visit associated with asthma as the first or second diagnosis and are in the random sample;

AND have at least one of the following:

•Disposition of the ED visit was admission to the hospital,
OR
•Documented physical findings consistent with respiratory distress, including any of the following:
o Labored breathing (including moderate or severe increased work of breathing); OR
o Retractions, grunting, and/or evidence of accessory muscle use; OR
o Markedly decreased breath sounds;
OR
•Recorded oxygen saturation below 90%;
OR
•An arterial blood gas (ABG) was obtained in the emergency department;
OR
•The child had a consultation with a pulmonologist or asthma specialist that was ordered and provided in the ED;
OR
•There is clear documentation that prior to arrival in the ED any of the following occurred:
o The child was referred to the ED after evaluation by the PCP or other clinician. The evaluation may include an in person visit or auscultation including via telephone. OR
o The child received two or more doses of inhaled rescue medications without sufficient clinical improvement. Documentation of parent report meets the criterion. OR
o The child was assessed with an objective instrument such as a peak flow meter and was found to be in a pre-defined "red zone" of peak flow measurement as part of an asthma action or similar plan. Documentation requires ALL of the following: a written asthma action plan exists AND defines a "red zone" for which urgent assessment by a clinician is indicated AND an objective assessment was made and its result was in the pre-defined red zone. Documentation of parent report meets the criterion.

The denominator is a random sample of the patients in each age stratum who have visited the emergency department for asthma (as a first or second diagnosis) and meet the specified criteria for having identifiable asthma (defined in s2b).

Separate numerators and denominators are reported for children age 2-5, 6-11, 12-18, and, optionally, 19-21 years. An overall rate across strata is not reported.

Denominator Elements:

The presence of identifiable asthma (see table 1) is established each month from administrative data using the specified algorithm.

Descriptive definitions for being managed for identifiable asthma are as follows. Specifications follow the descriptive definitions. Identifiable asthma is present in any child who has:

- Any prior hospitalization with asthma as primary or secondary diagnosis;

OR

- Other qualifying events, all ages:

o Three or more ambulatory visits with diagnosis of asthma or bronchitis, OR
o Two or more ambulatory visits with a diagnosis of asthma and/or bronchitis AND one or more asthma-related prescriptions

OR

- Other qualifying events, occurring after the fifth birthday:

o One or more prior ambulatory visits with asthma as the primary diagnosis AND a subsequent ED visit in the Reporting Month,
OR
o Two or more ambulatory visits with asthma as a diagnosis,
OR
o One ambulatory visit with asthma as a diagnosis AND at least one asthma-related prescription,
OR
o Two or more ambulatory visits with a diagnosis of bronchitis

For eligibility purposes, asthma-related medicine means long-acting beta-agonist (alone or in combination) or inhaled corticosteroid (alone or in combination), anti- asthmatic combinations, methylxanthines (alone or in combination), and/or mast cell stabilizers. See below further regarding this specification. Note that leukotriene modifiers and short term beta agonists are excluded for the purpose of establishing identifiable asthma. Data from the year prior to the reporting year are used, as well as all months prior to the reporting month in the reporting year (see Appendix Figure 1). Detailed specifications for asthma related medicine can begin with the NCQA NDC list (ASM-C_DASM-C_final_2012, found by clicking through at (http://www.ncqa.org/HEDISQualityMe asurement/HEDISMeasures/HEDIS20 12/HEDIS2012FinalNDCLists.aspx) Eliminate medications in the following 2 categories: leukotriene modifiers, short-acting inhaled beta-2 agonists). May use equivalent updated lists when provided by NCQA. Even if included in NCQA list, we further exclude indacaterol, a recently approved long acting beta agonist that is indicated in the US only for the treatment of COPD.

The analysis should be conducted on a month by month basis as described herein: Within the group of children who meet the criteria for identifiable asthma, identify and maintain a unique patient identifier, age, and all stratification variables. We call the time frame during which eligibility is established to be the Assessment Period.

For each month of the Reporting Year, determine eligibility for each patient, as of the last day of the month prior to the reporting month. This illustration assumes that the Reporting Year is 2011. When assessing January 2011, consider all of Calendar Year 2010 as the Assessment Period for assessing the presence or absences of identifiable asthma. For February, 2011 the Assessment Period includes all of calendar year 2010 AND January 2011. Repeat this progression monthly so that for December, 2011 identifiable asthma one would identify children with identifiable asthma using an Assessment Period from January 2010 through November 2011. For each month, assess whether the continuous enrollment criterion is met prior to including the month in the denominator. For example, for January 2011, the child must have been enrolled in November and December, 2010 (plus January 2011). Another example, for December 2011, to be eligible the child must have been enrolled in October 2011 and November 2011, as well as December. See Figure 1 in Appendix.

Develop Denominator sample according to Appendix Figure 2 and consistent with the instructions in sections S.18 and S.20.

Codes used for definitions are specified in s2b and include specifications of Hospitalization, Emergency Department Visits, Ambulatory/Office Visits, Asthma diagnosis,

Please note Figures 1 and 2 in the Appendix and the specifications in s2b are considered INTEGRAL to these specifications and are not optional.

ED visits that are already in the sample OR Children that fall outside of specified age range of 2-21 OR who do not meet time enrollment criteria OR do not meet identifiable asthma prior to the ED visit, OR children with concurrent or pre-existing COPD, Cystic Fibrosis or Emphysema.
At the discretion of the accountability entity, the denominator may be restricted to children 2-18.

Denominator Exclusions
1) Children with concurrent or pre-existing:
a. Chronic Obstructive Pulmonary Disease (COPD) diagnosis;
OR
b. Cystic Fibrosis diagnosis;
OR
c. Emphysema diagnosis.

OR
2) Children without identifiable asthma as specified
OR
3) Outside of specified age range
OR
4) Events occurring in patients who have not been enrolled in the reporting plan for at least two consecutive months before the index reporting month (a total of 3 consecutive months, including the reporting month).

This measure requires stratification by age group. Several additional stratifications are optional but may be required by the accountability entity or provided by the reporting entity. These variables include race/ethnicity, rurality/urbanicity and county level of poverty.

Stratify by age group (reporting entity should specify whether to use age at month of qualifying event or age on first day of reporting year):
•Age 2-5 years (second birthday to the day before the 6th birthday);
•Age 6-11 years (sixth birthday to the day before the 12th birthday);
•Age 12-18 years (twelfth birthday to the day before the 18th birthday); and
•Age 19-21 years (nineteenth birthday to the day before the 21st birthday).

Age strata are to be reported distinctly and not combined.

Optional stratifications require data elements such as:
•Race/Ethnicity
•Insurance type (Public, Commercial, Uninsured)
•Benefit type (if insured): HMO, PPO, Medicaid Primary Care Case Management (PCCM) Plan, Fee for Service (FFS), other
•Zip code, state and county or equivalent area of parent/caregiver's residence. Record FIPS if available

Stratification variables details
•Race/Ethnicity: Hispanic, Non-Hispanic Black, Non-Hispanic White; Non- Hispanic Asian/Pacific Islander, other Non-Hispanic
•Public vs Commercial (Private Insurance).
•HMO vs PPO vs FFS vs PCCM vs other; Within Medicaid, States may ask for reporting of FFS vs Managed Care or other relevant enrollment categories (e.g., TANF, SSI).

•Urban Influence Code. Identify the Urban Influence Code or UIC. (2013 urban influence codes available at: http://www.ers.usda.gov/data-products/urban- influence-codes.aspx#.UZUvG2cVoj8). Use parent or primary caregiver's place of residence to determine UIC. State and county names can be linked or looked up directly or zip codes can be linked to county indirectly, using the Missouri Census Data Center (http://mcdc.missouri.edu/). These data will link to county or county equivalents as used in various states.

Urban Influence Codes (UIC) have been developed by the USDA to describe levels of urbanicity and rurality. While each UIC has its own meaningful definition, some researchers choose to aggregate various codes. Well regarded schemas for aggregation of codes include Bennett and colleagues at the South Carolina Rural Research Center. Their aggregation scheme brings together Codes 1 & 2 as Urban; 3,5, & 8 as micropolitan rural; 4,6, & 7 as rural adjacent to a metro area; and 9, 10, 11, & 12 as remote rural. We acknowledge that UIC 5 (adjacent rural area) may appropriately be aggregated with 4,6,&7 as rural. Frontier health care may be approximated by analysis of the remote rural categories (UIC 9, 11 and 12). Alternatively, Gary Hart, Director of the Center for Rural Health at the University of North Dakota School of Medicine & Health Science suggests that UIC 9-12 is the best overall approach to using county level data to study frontier health. Inclusion of UIC 8 would make the analysis more sensitive to including frontier areas but at a meaningful cost in specificity.

Those interested in care specific to large cities may wish to aggregate the rural area and analyze UIC 1 and 2 separately.

When stratifying by urbanicity or UIC, the reporting and accountability entities should specify clearly what if any aggregating schema was used.

•Identify the Level of Poverty in the parent or primary caregiver's county of residence. The percent of all residents in poverty by county or county equivalent are available from the US Department of Agriculture at http://www.ers.usda.gov/data-products/county-level-data-sets/download- data.aspx. Our stratification standards are based on 2011 US population data that we have analyzed with SAS 9.3. Using parent or primary caregiver's state and county of residence (or equivalent) or FIPS code, use the variable PCTPOVALL_2011 to categorize into one of 5 Strata:
o Lowest Quartile of Poverty if percent in poverty is <=12.5%
o Second Quartile of Poverty if percent in poverty is >12.5% and <=16.5%
o Third Quartile of poverty if percent in poverty is >16.5% and <=20.7%
o First Upper Quartile (75th-90th) if percent in poverty is >20.7% and <=25.7%
o Second Upper Quartile (>90th percentile)

These classification standards may be updated by the accountability entity suing more recent data if desired. Note: if needed, the Missouri Census Data Center may be used to link zip codes to county equivalents. http://mcdc.missouri.edu/

Within each age group, randomly select 500 ED visits among those identified in Step 4. Analyze each age strata's random sample distinctly:

Sort into three groups according to Appendix Figure A.1.b.

• Group A: Those with asthma ED visits ONLY and no associated asthma hospitalization to the same hospital on the same date. These ED visits are INCLUDED in the Denominator and receive Medical Record Review to assess eligibility for the Numerator;

• Group B: Those with both asthma ED Visits and asthma Hospitalizations at the same facility on the same date and for whom the hospital discharge date is after the ED date of service. These ED visits are INCLUDED in both the Denominator and in the Numerator. No further review is necessary to establish appropriateness;

• Group C: Those with asthma Hospitalizations ONLY and no associated asthma ED Visit to the same hospital on the same date. Please note that children admitted to the ED one date and admitted to the hospital the next day (from the same ED visit) will be identified in this group. Group C Hospitalizations are subject to Medical Record Review to assess eligibility for the Denominator. If they are eligible for the denominator they will be included in BOTH the Numerator and Denominator.

Please note that the terms medical chart and medical record are used interchangeably, as are the terms audit and review in this context.
Figure 2 in the Appendix is integral to sample selection and no optional.
Notes:
•Determining eligibility for sample selection precedes determining eligibility for measure.
•On the basis of the Administrative Data Analysis, children who are potentially eligible for the measure will be identified and segregated into Groups A, B, and C (the blue boxes In Figure 2 of Appendix).
•Children are eligible for Group B if three things are found in the administrative data: ED Visit; Hospitalization on same day and same institution; and Hospital discharge is after date of ED visit.
•National and NY State data suggest that approximately ¾ of childhood asthma hospitalizations are admitted from ED, that about 1 in 9 childhood asthma ED visits result in hospitalization and that children admitted from the ED may not have their ED visit coded in administrative data.
•Medical record review determines eligibility for numerator among the Group A children, all of whom have already qualified to be included in the denominator.
•Group B children are eligible for both the numerator and the denominator on the basis of administrative data analysis alone and do not require chart review.
•Medical record review determines eligibility for inclusion in the measure (denominator!) for Group C children. If they are eligible for the denominator (i.e. that have been admitted directly from an unduplicated ED visit) then they are also qualified for the numerator.

The impact of sample size on the width of the confidence interval is illustrated by assuming 50% appropriateness and a variety of sample size to calculate the width of the confidence intervals around the estimate obtained above. Variations from 50% will bring down the size of the confidence interval.

N=	50,	+ / - 13%
N=	75,	+ / - 11%
N=	100,	+ / - 10%
N=	150,	+ / - 8%
N=	200,	+ / - 7%
N=	250,	+ / - 6%
N=	400,	+ / - 5%

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.)*
IF a PRO-PM, specify calculation of response rates to be reported with performance measure results.
N/A

S.17. **Data Source** *(Check ONLY the sources for which the measure is SPECIFIED AND TESTED).*
If other, please describe in S.18.
Claims (Only), EHRs Hybrid, Paper Records

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 is collected.)*
IF a PRO-PM, identify the specific PROM(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)*
No data collection instrument provided

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

S.21. **Care Setting** *(Check ONLY the settings for which the measure is SPECIFIED AND TESTED)*
Emergency Department, 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.)*
N/A

2. Validity – See attached Measure Testing Submission Form
nqf_testing_attachment_12_20_16_LKV.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. (Do not remove prior testing information – include date of new information in red.)

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. (Do not remove prior testing information – include date of new information in red.)

2.3 For maintenance of endorsement
Risk adjustment: For outcome, resource use, cost, and some process measures, risk-adjustment that includes SDS factors is no longer prohibited during the SDS Trial Period (2015-2016). Please update sections 1.8, 2a2, 2b2, 2b4, and 2b6 in the Testing attachment and S.14 and S.15 in the online submission form in accordance with the requirements for the SDS Trial Period. NOTE: These sections must be updated even if SDS factors are not included in the risk-adjustment strategy. If yes, and your testing attachment does not have the additional questions for the SDS Trial please add these questions to your testing attachment:

What were the patient-level sociodemographic (SDS) variables that were available and analyzed in the data or sample used? For example, patient-reported data (e.g., income, education, language), proxy variables when SDS data are not collected from each patient (e.g. census tract), or patient community characteristics (e.g. percent vacant housing, crime rate).

Describe the conceptual/clinical and statistical methods and criteria used to select patient factors (clinical factors or sociodemographic 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)

What were the statistical results of the analyses used to select risk factors?

Describe the analyses and interpretation resulting in the decision to select SDS 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)
No - This measure is not risk-adjusted

NATIONAL QUALITY FORUM—Measure Testing (subcriteria 2a2, 2b2-2b7)

Measure Number (*if previously endorsed*)**:** Click here to enter NQF number

Measure Title: Appropriateness of Emergency Department Visits for Children and Adolescents with Identifiable Asthma

Date of Submission: 12/20/2016

Type of Measure:

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

Instructions

- 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.
- **For all measures, sections 1, 2a2, 2b2, 2b3, and 2b5 must be completed.**
- **For outcome and resource use measures**, section **2b4** also must be completed.
- If specified for **multiple data sources/sets of specificaitons** (e.g., claims and EHRs), section **2b6** also must be completed.
- 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 (2b2-2b6) must be in this form. An appendix for *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.
- Maximum of 20 pages (*incuding questions/instructions;* minimum font size 11 pt; do not change margins). ***Contact NQF staff if more pages are needed.***
- Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.
- For information on the most updated guidance on how to address sociodemographic variables and testing in this form refer to the release notes for version 6.6 of the Measure Testing Attachment.

Note: The information provided in this form is intended to aid the Steering 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 **PRO-PMs and composite performance measures**, reliability should be demonstrated for the computed performance score.

2b2. 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 **PRO-PMs and composite performance measures**, validity should be demonstrated for the computed performance score.

2b3. Exclusions are supported by the clinical evidence; otherwise, they are supported by evidence of sufficient frequency of occurrence so that results are distorted without the exclusion; [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]

2b4. 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 sociodemographic 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.

2b5. 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.

2b6. If multiple data sources/methods are specified, there is demonstration they produce comparable results.

2b7. For **eMeasures, composites, and PRO-PMs** (or other measures susceptible to missing data), 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.
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.23*)	**Measure Tested with Data From:**
☒ abstracted from paper record	☒ abstracted from paper record
☒ administrative claims	☒ administrative claims
☐ clinical database/registry	☐ clinical database/registry
☒ abstracted from electronic health record	☒ abstracted from electronic health record
☐ eMeasure (HQMF) implemented in EHRs	☐ eMeasure (HQMF) implemented in EHRs
☐ other: Click here to describe	☐ other: Click here to describe

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*).

New York State Medicaid claims data 2010 – 2012.

Our work builds off of work performed by our CAPQuaM partner and steering committee member, NCQA. For specific data reliability and signal to noise analyses, we incorporate by reference (and will present more selectively) NCQA data relevant to their submission for NQF – endorsed asthma related measures 0036, 1799, and 1800.

Their analyses demonstrate the capacity to use administrative data to identify the applicable denominator population. There is nearly complete overlap of the denominator codes and there is overlap of the denominator elements. Where codes differ it is specific to decisions made by the CAPQuaM expert panel which was aware of the NCQA measures.

Newly abstracted data was also used for this measure.

1.3. What are the dates of the data used in testing? 10/2009 – 11/2013

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.26*)	**Measure Tested at Level of:**
☐ individual clinician	☐ individual clinician
☐ group/practice	☐ group/practice
☒ hospital/facility/agency	☒ hospital/facility/agency
☒ health plan	☒ health plan
☐ other: Integrated delivery system, population, state, region, county	☐ other: Integrated delivery system, population, state, region, county

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*)

We surveyed 9 hospitals around the country for data availability, data source, and ease of data abstraction related to this measure. Most of these hospitals used electronic medical records, but we also surveyed several that at the time did not. Respondents included: VP Quality & Patient Safety; Performance Improvement Coordinator; Director of Medical Affairs, QI RN, Quality and Performance Improvement Project Management Coordinator; Clinical Manager Emergency Services & the Resource Pool; Case Manager Women and Children's Service Line; and the Director of Quality Management. 8 of the 9 specifically indicated that they were answering with an institutional perspective and 5 of the nine specifically reviewed charts to assist in answering the survey. At least 19 charts were reviewed.

All 9 respondents indicated that race/ethnicity were not difficult to collect and was present in the medical records. The answer set regarding difficulty of abstraction included: Not Available; Not Difficult to Collect; Difficult to Collect; and Very Difficult to Collect. Eight of nine indicated that date of birth was not difficult to collect and all agreed that it was in the medical record, to allow for simple assessment of age. All 9 agreed that payment source was likewise not difficult to collect from the patient's record. All indicated that the presence of prior asthma was typically in the record and not difficult to assess via chart audit. Eight of nine indicated that abstraction of clinical data to assess the appropriateness of the ED was not difficult to accomplish via chart audit. In contrast 2 of 9 indicated that a forma asthma severity score assessment was not available in the ED chart, and three of nine that no formal interpretation of such a score would be available in the chart.

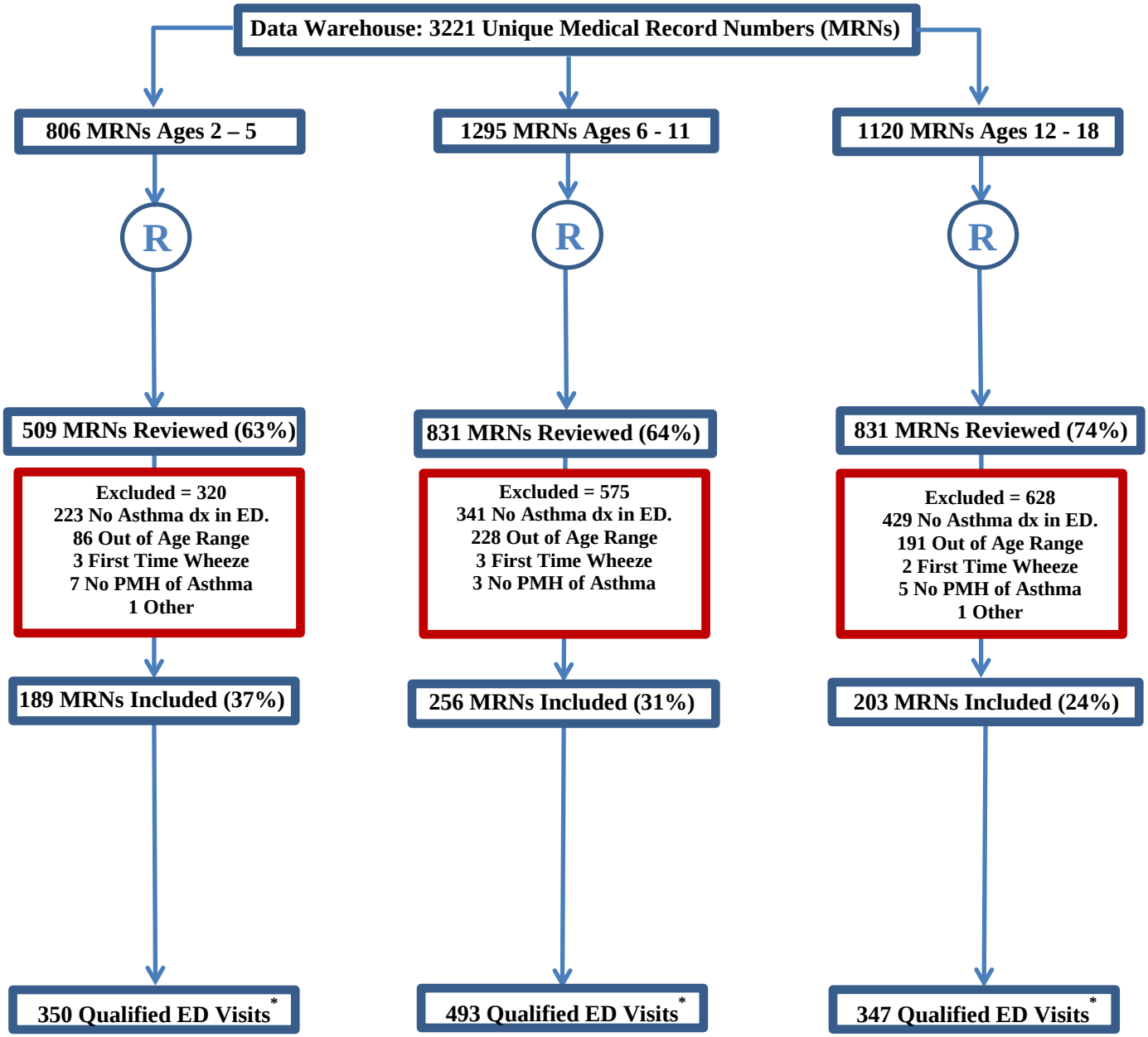
Nine of nine indicated that the following were not difficult to collect: collection of oxygen saturation; identifying the lowest recorded oxygen saturation; identifying whether or not an arterial blood gas (ABG) had been collected was not difficult to collect; identifying the level of respiratory distress (mild, moderate, severe dyspnea); presence of retractions; and admission to the hospital from the ED. More varied was information regarding whether or not the child had been referred into the ED by the PCP (4 = "Not Difficult", 2= "Difficult"; 2= "Very Difficult", 1="Very Difficult", and 2="Not Available"). This last finding was among those aspects of testing that led our specifications to name indicate appropriateness based on evidence of presence of at least one of the criteria, and not to consider the absence to be meaningful in and of itself. We further extend this principle to recognize that certain

aspects of care that are based upon parental response are not likely to fully captured in the medical record, even if such documentation ought to be part of a standard of high quality care. Thus we speak of “Appropriate” and “Questionable” ED visits, rather than “inappropriate” visits.

Foundational analyses for this measure included:

Analysis of NY State Medicaid Managed Care claims data, including claims from all MCO’s that are contracted for Medicaid care by our partner, the NY State Dept of Health. We identified eligible populations and events from both RY 2011 and 2012 and include children from counties in nine urban influence codes and in counties poverty level 1-3. NY State does not have any counties in the lowest 25% of poverty or with UIC of 10-12. New York has more than 60 counties and numerous health plan vendors. Analysis in year 2011 provided very similar data to 2012

Figure 1: Chart Selection for Review of ED Appropriate Use for Pediatric Asthma



*** Some patients had multiple ED visits for asthma**
R = Randomized

.

For the NCQA analysis, nine health plans covering a variety of geographic areas within the United States were asked to provide a complete administrative data file consisting of any member in their commercial and Medicaid product lines for anyone that had a diagnosis code for asthma during the calendar years of 2009-2010. The complete member-level administrative file used for analysis included a total of more than 82,000 health plan members with asthma.

The specific measure demonstration and testing was done at one site, a New York City Academic Medical Center. In this testing, **sample selection can be summarized in the diagram above**.

The eligible observation period was October 2009 to November 2013. Please note, because of the limitations of the data systems available for testing randomization happened at the level of the patient. For patients with 1-3 visits for asthma in the included time frame, we included all visits. For patients with more than 3 visits, the first three visits were included. Hence the average number of visits per child was 350/189= 1.9 for the younger children, 493/256=1.9 for the school age children, and 347/203=1.7 for adolescents. In NY State Medicaid in 2011, the median number of visits per child was 1, the 75th percentile was 2 and the 90th percentile was 3 (N=26,169 children). Hence this finding is plausible and consistent, given the 4 year time frame that we sampled.

The eligible observation period was October 2009 to November 2013. Please note, because of the limitations of the data systems available for testing randomization happened at the level of the patient. For patients with 1-3 visits for asthma in the included time frame, we included all visits. For patients with more than 3 visits, the first three visits were included. Hence the average number of visits per child was 350/189= 1.9 for the younger children, 493/256=1.9 for the school age children, and 347/203=1.7 for adolescents. In NY State Medicaid in 2011, 40,855 children experienced 61,327 eligible emergency department visits. The median number of visits was 1.5. 10% of children had 3 or more eligible visits and the mode number of visits per child was 1. Hence the findings in our testing is plausible and consistent, given the 4 year time frame that we sampled.

1.6. How many and which <u>patients</u> 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*)

Refer to Figure 1 in 1.5. Using the institutional data warehouse, we randomly identified medical record numbers of children who had both ED visits and asthma diagnoses in the specified time frame. Because we were not using claims data to select them, charts had to be reviewed for evidence of prior asthma and to assure that ED visit and asthma diagnosis were concurrent and in the selection time frame. ED visits were excluded if there was not evidence that they were known to be asthmatic, if the ED visit did not have asthma as the first or second diagnosis, or if the ED visit was not in the specified time frame. We included up to 3 visits per selected child, using the first 3 visits when more than three were present. Inclusion criteria included an ED visit with previously established asthma as a primary or secondary diagnosis as documented in the electronic medical record. We developed 3 samples stratified by age: 2-5 years, 6-11 years, and 12-18 years. For ages 2 – 5, we included 350 visits; ages 6 – 11, 493 visits; ages 12 – 18, 347 visits. So included in the measure testing was a total of 1200 ED visits were included in the chart review testing.

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.

Assessment of the capacity to identify the eligible population and qualifying events was performed in NY State Medicaid data in both 2011 and 2012 reporting years.

Our construct for the CAPQuaM measure was defined by the multidisciplinary national expert panel using a RAND type modified Delphi process. The panel initially used the term persistent asthma to

describe asthma that was pre-existing and should have been recognized as asthma by the health care system prior to the timing of the ED visit. This construct was renamed by our stakeholder group to be identifiable asthma to avoid confusion with other uses of the term persistent asthma. The construct was intended to be more inclusive than HEDIS' persistent asthma diagnosis, while still removing from consideration those whose asthma was unlikely to have been actively managed at the time.

Holding steady the continuous enrollment criterion at 12 months, HEDIS criteria identified a rate of persistent asthma of 3.1% with the CAPQuaM criteria identifying identifiable asthma at a rate of 8.6%. This ratio is 2.8, which is between 2-3, which is what we had predicted (based on the team's reading of the literature) and was the goals we were hoping to achieve with our criteria and was interpreted to suggest construct validity for our measure. Using data form the National Survey of Children's Health, we estimated the expected rate of asthma in the NY State Medicaid child population to be between 15 - 16%, indicating that our criteria did provide a meaningful filter as we had intended.

We found that by reducing the continuous enrollment period down to three months as was suggested by members of our steering committee that we could increases the number of children eligible for the measure by several tens of thousands while still restricting the measure to those who had received sufficient care for asthma to be identified, and requiring continuous enrollment for attribution to the extent felt important by our multi-stakeholder group.

Assessment of data elements for identifying a population with asthma was performed by NCQA in nine geographically diverse managed care plans.

Assessment of appropriateness was performed in 1200 pediatric ED visits from a single medical center.

1.8 What were the patient-level sociodemographic (SDS) variables that were available and analyzed in the data or sample used? For example, patient-reported data (e.g., income, education, language), proxy variables when SDS data are not collected from each patient (e.g. census tract), or patient community characteristics (e.g. percent vacant housing, crime rate).

Insurance status and race/ethnicity for the single site analysis.

Race, ethnicity, zip code, level of poverty in the zip code of caregiver residence, and urban influence in the county of caregiver residence for the NY State 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*)

Validity testing was performed at the data element level for both the numerator and the

denominator.

See section 2b2 for validity testing of data elements

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 noted below, reliability for finding the presence or absence of appropriateness criteria could be trained and kappas suggest high reliability.

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?*)
Excellent reliability.

2b2. VALIDITY TESTING

2b2.1. 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*)

2b2.2. 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)*

Please see descriptions of testing above as well.

As described below, the literature also supports the use of claims data to identify the presence of asthma. The table summarizes these findings.

Data element	**Reference** (e.g., Quam, et al., 1993)	**Data source** (e.g., Medicare FFS outpatient data)	**Statistical results** (e.g., kappa, sensitivity, specificity, etc.)
Numerator: Validation of Numerator Data Elements was performed by the CAPQuaM development team and the results are summarized in this section following the table.			
Denominator			
Age	NYSDOH CAPQuaM Analysis – internal testing	NY State Medicaid Data	Meaningful variation by age groups as predicted, with peaks in younger children and older adolescents.
	CMS MMIS data requirements Exemplar specifications at https://www.cms.gov/Research-Statistics-Data-and-Systems/Computer-Data-and-Systems/MSIS/downloads/msisdd2010.pdf	State Medicaid MMIS systems	States are required to submit validated claims data including age or date of birth with a tolerance of 0.1%
Asthma diagnosis in inpatient/ED setting	Wilchesky, M., Tamblyna, R. M., & Huang, A. (2004). Validation of diagnostic codes within medical services claims. Journal of Clinical Epidemiology, 57, 131-141.	Drug utilization review, the Charlson comorbidity index and the Johns Hopkins Adjusted Care Group Case-Mix profile (ADGs).	Asthma claims were highly specific, Sp= 96.76 (95%CI 96.5, 97.0).

Asthma diagnosis in ambulatory setting	Fowles, J. B., Fowler, E. J., & Craft, C. (1998). Validation of claims diagnoses and self-reported conditions compared with medical records for selected chronic diseases. Journal of Ambulatory Care Management, 21(1), 24-34.	Multispecialty group practice in Minneapolis, Minnesota	Sensitivity and specificity was 0.82 and 0.99, respectively. Sensitivity of .82 using claims was higher than sensitivity using self-report at .64
Asthma diagnosis in clinic/outpatient setting	Wilchesky, M., Tamblyna, R. M., & Huang, A. (2004). Validation of diagnostic codes within medical services claims. Journal of Clinical Epidemiology, 57, 131-141.	Drug utilization review, the Charlson comorbidity index and the Johns Hopkins Adjusted Care Group Case-Mix profile (ADGs).	Asthma claims were highly specific, Sp= 96.76 (95%CI 96.5, 97.0).
Bronchitis diagnosis in ambulatory setting	Improving Healthcare for the Common Good (IPRO). Avoidance of Antibiotic Treatment in Adults with Acute Bronchitis. May 2011. http://www.health.ny.gov/health_care/managed_care/reports/docs/adults_antibiotic.pdf	New York Medicaid managed care members	An IPRO analysis of ambulatory claims data in NY State Medicaid found that of 651 individuals with an administrative claim for bronchitis, 629 (96.6%) were confirmed by chart review.
Fill of short acting beta agonist Fill of asthma controller medication	Samnaliev, M., Baxter, J. D., & Clark, R. E. (2009). Comparative evaluation of two asthma care quality measures among Medicaid	Using complete claims and pharmaceutical data for 19,076 patients with persistent asthma (based on Health Effectiveness and Data Information Set criteria) in five Medicaid populations	Sensitivity and specificity were combined into one statistic, the area under the ROC curve. For controller medications, the area under ROC curve is 0.705, which represents good agreement.

• anti- asthmatic combination • antibody inhibitor • inhaled steroid combinations • inhaled corticosteroids (alone or in combination) • leukotriene modifiers • methylxanthines (alone or in combination) • mast cell stabilizers	beneficiaries. Chest, 135(5), 1193-1196. Mudd KE, Bollinger ME, Hus VD, et al. Concordance of Mediaciad and pharmacy record data in inner-city children with asthma. Contemporary Clinical Trials 29(2008) 13-20 Grymonpre R, Xheang M, Fraser M, et al. cvalidity of Precritpion Claims Database to Estimate Medication Adherence in Older Persons e.g. Samnaliev M, Baxter JD, and Clark RE. Comparative Evaluation of Two Asthma Care Quality Measure Among Medicaid Beneficiaries. Berger WE, Legorreta AP, Blaiss MS, et al. The Utility of	(Colorado, Georgia, Indiana, New Jersey, Washington) using ICD-9-CM code 493.x to measure filling prescriptions of asthma control medication. Comparison of pharmacy records and Medicaid clams Manitoba prescription claims and pill count for medication adherence A number of studies found that asthma drug data using the similar HEDIS data elements that we propose were valid for predicting things like emergency department use in asthma patients. As indicated in this article: “HEDIS has become an important industry standard…adopted by	For inner city children on Medicaid, Medicaid claims was sensitive compared to pharmacy records, identifying 91.3% of pharmacy claims for ICS, 94.7% for SABA and 90.4% for leukotriene modifiers (Table 2) Using a much stronger standard of actual compliance, this study found for multiple condition for two conditions in adults that there was strong concordance (79% and 88% respectively) between pill counts and administrative claims data. Not specific for asthma meds Controller medication use was associated with fewer ED visits across 5 states, with OR ranging from 0.30 to 0.47, all significant, overall 0.34 (0.32-0.36). Used actual HEDIS pharmacy code set as do we. Low Controller use had an adjusted odds ratio of 1.72 (1.42-2.08) of ED visit or

	the HEDIS Asthma Measure to predict asthma related outcomes. Annals of Allergy, Asthma, and Immunology. 93:538-545. 2004.	regulators, consumers, and public purchasers of health care" Commercial claims	hospitalization. Those with moderate and higher adherence had graded reductions in undesirable outcomes in the predicted fashion (OR, .84 and 0.72 respectively)
Exclusions			
Diagnosis of COPD	Rawson NS, Malcolm E., validity of the recording of ischaemic heart disease and chronic obstructive pulmonary disease in the Saskatchewan health care datafiles. State Med. 1995. Dec 30: 14 (24):2627-43.	Administrative health care datafiles of the Canadian province of Saskatchewan	Comparisons between hospital data and medical charts for chronic airways obstruction patients showed excellent diagnostic agreement at 94%. In other words, the charted discharge diagnosis from the patient's medical record showed exact agreement for 94.2% of these patients.
	Ginde AA, Tsai CL, Blanc PG, Camargo CA Jr. Positive predictive value of ICD-9-CM codes to detect acute exacerbation of COPD in the emergency department. Jt Comm J Qual Patient Saf.2008;34(11):678–680.	Two academic emergency departments.	The overall positive predictive value for the presence of any of the specified codes, including COPD, was 97%. The positive predictive value for a code of 496 alone was 60% (95% CI 32-84%).
	Gershon AS, Wang C, Guan J, Vasilevska-Ristovska J, Cicutto L, To T. Identifying individuals with physician diagnosed	Claims in Ontario, Canada	The combination of one or more outpatient ICD-9 codes (491.xx, 492.xx, 496.xx) and ICD-10 inpatient ICD-10 codes (J41, J43,

	COPD in health administrative databases. Copd. 2009;6(5):388–394. doi: 10.1080/15412550903140865.		J44) had a sensitivity of 85% and specificity of 78.4% among 113 patients with COPD and 329 patients without COPD.
Diagnosis of COPD Diagnosis of cystic fibrosis Diagnosis of emphysema (Exclusions identified anywhere are excluded. The measure is written to over exclude if need be, but our data suggest that exclusions are uncommon.)	Quan, H., Li, B., Saunders, L. D., Parsons, G. A., Nilsson, C. I., Alibhai, A., et al. (2008). Assessing validity of icd-9-cm and icd-10 administrative data in recording clinical conditions in a unique dually coded database. HSR: Health Services Research, 43(4), 1424.	Four teaching hospitals in Alberta, Canada	Claims had a PPV of 91.9, and a negative predictive value of 92.6, with *k* of 0.65 (substantial agreement[i]) compared to chart review for chronic pulmonary disease. ICD 10 performed similarly in this study
	NCQA: http://www.qualityforum.org/QPS/QPSTool.aspx?m=367&e=1	The presence of diagnostic exclusions was extensively tested on the entire field test population (>82,000 members) to determine the effect on eligible population and the measure results experienced as a result of the application of clinical exclusions.	This measure was deemed valid by the expert panel and approved by NCQA's Committee on Performance Measurement (CPM) for continued inclusion in HEDIS[ii]

Data element	**Reference** (e.g., Quam, et al., 1993)	**Data source** (e.g., Medicare FFS outpatient data)	**Statistical results** (e.g., kappa, sensitivity, specificity, etc.)
Race/ Ethnicity	Kressin, NR, Chang, BH, Hendricks, A, Kazis, LE. Agreement Between Administrative Data and Patients' Self-Reports of Race/Ethnicity. American Journal of Public Health. Oct. 2003. 93 (10): 1734-1739.	Federal administrative data	Among patients with known race/ethnicity, there was a 97.9%, 92.0%, and 83.4% agreement between self-report race/ethnicity and administrative data for white, African American, and Hispanic, respectively. (Table 2, p. 1736)
	Blustein, J. The Reliability of Racial Classifications in Hospital Discharge Abstract Data. American Journal Public Health. 1994; 84:1018-1021.	Statewide Planning and Research Cooperative System, a hospital discharge abstract database maintained by the New York State Department of Health.	Percentage of concordance and kappa of reported racial classifications: - Black: 99%; 089 (95% CI: 0.82, 0.96) - White: 95%; 0.72 (95% CI: 0.64, 0.80) (Table 3, page 1020)
	Klinger, EV, Carlini, SV, Gonzalez, I, et al., Accuracy of Race, Ethnicity, and Language Preference in an Electronic Health Record. 2014. J Gen Intern Med. 30(6):719-23.	Thirteen primary care clinics' electronic health records.	When comparing electronic health record to self-report the sensitivity, specificity and ppv for Black, Hispanic and white are as follows (Table 2, page 721): - Black: Se: 70.9, Sp: 98.8, PPV: 95.5 - Hispanic: Se: 83.8, Sp: 99.8; PPV: 98.9 - White: Se: 93.8; Sp: 97.0; PPV: 98.3
	Escarce, JJ and McGuire, TG., Methods for Using Medicare Data to Compare Procedure Rates among Asians, Blacks, Hispanics, Native Americans, and Whites. Health Services Research. Oct. 2003. 38(5): 1303-1318.	Physician claims data	When comparing enrollment database and survey, probability for White, Black, and Hispanic are 0.954, 0.943, 0.977, respectively. (Table 2, page 1309)

We develop our measure using scientifically sound principles. We first discuss research involving the soundness of our data sources, which include both administrative data to identify cases (and a fraction of numerator qualifications) and chart review (medical record audit) to confirm some denominator inclusions and to identify most numerator inclusion. This is a generally accepted and standard approach with acceptable reliability.

We use administrative data to identify the age of the child, various stratification variables and the presence of asthma, as well as the presence of an asthma ED visit or hospitalization. These are routinely used to support billing by CMS, Medicaid, and private insurers and are routinely used in quality measurement. Administrative data are not typically sufficient for detailed clinical assessment.[1-5] HEDIS developed a hybrid approach, using administrative data and chart review that this measure borrows heavily from. [6, 7]

There is moderate agreement (kappa = 0.45 – 0.50) when comparing administrative data regarding the presence of constructs such as recent asthma attacks, use of asthma medications, attack or medication, attack and medication, using 1 year of administrative claims data. The agreement improves from 0.55 to 0.60 when using two years of data. (8). We expect that these kappas would be significantly higher were the analyses restricted to children with disease that met our construct criteria for identifiable asthma.

The explicit criteria that we use were developed using a slightly modified version of the RAND/UCLA Appropriateness Method that maintained the key aspects of that approach, including a detailed literature review, a multidisciplinary and geographically diverse expert panel comprised of both clinicians and researchers, and the two Round modified Delphi Process. The general reliability of this approach is well established. [9, 10] It has been applied successfully to pediatric services previously. [11-13] We have used as criteria for this measure those specifications whose median rating is 8 or 9, the two highest ratings.

In our testing of the criteria during chart audit used a simple paper data collection instrument that was largely a checklist of yes/no for the various items. After a brief training by the physician who organized the testing three non-clinical research assistants (one MPH, 2 Bachelors) conducted chart audits. Kappa is presented in the next section (**2b2.3)**.

1. Dresser, M.V., et al., *Clinical quality measurement. Comparing chart review and automated methodologies.* Med Care, 1997. **35**(6): p. 539-52.
2. Newton, K.M., et al., *The use of automated data to identify complications and comorbidities of diabetes: a validation study.* J Clin Epidemiol, 1999. **52**(3): p. 199-207.
3. Thompson, B.L., et al., *Measuring clinical performance: comparison and validity of telephone survey and administrative data.* Health Serv Res, 2001. **36**(4): p. 813-25.
4. Angier, H., et al., *Variation in outcomes of quality measurement by data source.*
a. Pediatrics, 2014. **133**(6): p. e1676-82.
5. Weiskopf, N.G. and C. Weng, *Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research.* Journal of the American Medical Informatics Association, 2013. **20**(1): p. 144-151.
6. Pawlson, L.G., S.H. Scholle, and A. Powers, *Comparison of administrative-only versus administrative plus chart review data for reporting HEDIS hybrid measures.* Am J Manag Care, 2007. **13**(10): p. 553-8.
7. NCQA. *National Committee for Quality Assurance*. [cited 2014 7/30/14]; Available from: http://www.ncqa.org/
8. Huzel, L, et al. Diagnosing Asthma: The fit between survey and administrative database. Can Respir J. 2002 Nov-Dec;9(6):407-12.
9. Fitch, K., et al., *The RAND/UCLA Appropriateness Method User's Manual*. 2001 RAND.
10. Kosecoff, J., et al., *The appropriateness of using a medical procedure. Is information in the medical record valid?* Med Care, 1987. **25**(3): p. 196-201.
11. Kleinman, L.C., et al., *The medical appropriateness of tympanostomy tubes proposed for children younger than 16 years in the United States.* Jama, 1994. **271**(16): p. 1250-5.
12. Kleinman, L.C., E.A. Boyd, and J.C. Heritage, *Adherence to prescribed explicit criteria during utilization review. An analysis of communications between attending and reviewing physicians.* Jama, 1997. **278**(6): p. 497-501.
13. Keyhani, S., et al., *Overuse of tympanostomy tubes in New York metropolitan area: evidence from five hospital cohort.* Bmj, 2008. **337**: p. a1607.

Table 1. 360 Degree Pediatric Quality Measure Development: Overview

Stage	Phase	Innovation	Product(s)
1. Clinical Criteria Development	a. Input Development	1. Focus groups of caregivers of children with asthma who have used the ED 2. Interviews with front line clinicians: primary care, asthma docs, and ED docs	1. Literature review 2. Summary of consumer perspectives, values and understanding relevant to clinical issue of interest 3. Summary of findings form clinician interviews
	b. RAND/UCLA 2 Round Modified Delphi Process	1. Inclusion of consumer perspectives as a key input; 2. Use of this method to identify appropriateness criteria in national performance measure development;	1. Explicit criteria that rank a comprehensive and mutually exclusive set of clinically detailed scenarios;
2. Boundary Guideline Development	Criteria Enhancement	1. Iterative process to enhance reliability and internal consistency of the explicit criteria set with a goal of outlining three boundary spaces	1. Internally consistent set of explicit criteria that are stable in their representation of the expert panel perspective. "Enhanced criteria"
	Guideline Articulation	1. Stakeholder (including experts, users, clinicians, consumers and others) informed review of the enhanced criteria. 2. Definition of zones of potential overuse, potential underuse, and professional interaction and decision-making based upon the explicit criteria 3. Stakeholder valuations of potential deviations from guideline 4. Boundary Guideline	1. Boundary Guideline 2. Prioritization list
3. Creation of Measure	Specification	1. Translation of guideline into specification of necessary data 2. Iterative process to define optimally efficient sources of data to allow for measurement and stratification	1. Initial specification of measure
	Review	1. Constructive peer review of specifications by stakeholders in Steering Committee and SAB	1. Final specifications of measure including variables for stratification as needed
	Fielding and testing of measure	1. Measure testing	1. Functional experience and practical understanding of measure, its scoring, variability, and interpretation

This measure was developed and assessed using a pre-specified process and consistent with CAPQuaM's peer reviewed 360 degree method outlined in the table above.
Explicit criteria were developed using a variation of the two-round modified Delphi process RAND/UCLA Appropriateness Method with a multidisciplinary and geographically diverse expert panel comprised of both clinicians and researchers. Identifiable asthma was based on panel findings and appropriateness criteria included for this measure were those that were both available in the chart and highly rated.

Development included a series of alpha tests to refine specifications by conducting iterative analyses in New York State Medicaid data. Conclusions from alpha tests include:

1) The reporting period and the assessment period could not overlap completely, leading to use of 2 years of data as shown in the specifications' diagram. The optimal approach was to divide the reporting year into 12 reporting months. ED events in that month are eligible for the numerator if persistent asthma criteria have been satisfied (combining the look-back year and all prior months in the reporting year) and the child has been continuously

enrolled for the two months immediately prior to the reporting month. The optimal unit for the denominator is in child-months;

2) Using both revenue codes and CPT codes increased our sensitivity meaningfully, a choice validated by consultation with coding and billing experts;
3) NY State Medicaid data and national survey data (HCUP) converged to demonstrate the importance of including hospitalizations as numerator events even when the underlying construct is ED visits. This is consistent with policies of many payers to request providers not to submit both ED and hospital claims for the same day. Error would be far less by considering both ED visits and hospitalizations as numerator events, than by not including hospitalizations.
4) The expert panel only wanted numerator events for which the children were already known to the accountable entity as having asthma and established definitions for such "identifiable asthma". Identifiable asthma was intended to be more restrictive than the 15-16% identified by our analysis of the 2011 NSCH as having ever been told they had asthma and much less restrictive than the HEDIS definition of persistent asthma. Alpha testing in NY State Medicaid demonstrated the expected results:
    a. Holding steady the continuous enrollment criterion at 12 months, HEDIS criteria identified a rate of persistent asthma of 3.1%, the CAPQuaM criteria identifying identifiable asthma at a rate of 8.6%. This ratio is 2.8 (our predicted and target result was between 2-3 based the literature achieve and our intended construct).
    b. Relaxing the continuous enrollment period to 3 months was suggested by members of our stakeholder steering committee. Doing so increased the eligible number by several tens of thousands while still restricting the measure to those who had received sufficient care for asthma to be identified, and requiring continuous enrollment for attribution to the extent felt important by our multi-stakeholder group.

The use of Expert Panels has been demonstrated to be useful in measure development and health care evaluation, including for children. [1] Use of the medical record as a valid source of information to judge appropriateness is well accepted. [2] Chart audits are used frequently to generate research in Emergency Medicine. [3, 4]

Key panel ratings are shown. Constructs rated 7 or higher are endorsed, 8 or higher strongly endorsed, and 2 or lower strongly rejected.

Scenario	**MED**
Wheezing on presentation to the ED establishes that the ED was an appropriate level of care for that child.	5
Retractions or labored breathing during the ED visit establishes that the ED was an appropriate level of care for that child.	**9**
Decreased breath sounds establish that the ED was an appropriate level of care.	6
Markedly decreased breath sounds establish that the ED was an appropriate level of care.	**7**
Obtaining an ABG in the ED establishes the ED as an appropriate level of care for that child.	**9**
Oxygen saturation less than 90% establishes that the ED was an appropriate level of care for that child.	**9**
Hospitalization following the ED visit establishes that the ED was an appropriate level of care for that child.	**9**
An ED visit less than 72 hours following a previous ED visit in a child with asthma establishes that the ED was an appropriate level of care for that child.	4
Prescription of an oral steroid burst establishes that the ED was an appropriate level of care for that child.	4
An ED visit less than one week following a hospital discharge in a child with asthma establishes that the ED was an appropriate level of care for that child.	4
An ED visit less than 72 hours following a hospital discharge in a child with asthma establishes that the ED was an appropriate level of care for that child.	3
A specialty consultation in the ED establishes that the ED was an appropriate level of care for that child.	**8**
Homelessness establishes that the ED was an appropriate level of care for that child.	3
Parent report that the PCP is generally unavailable for urgent asthma care establishes that the ED was an appropriate level of care for that child.	5
Parent report of inability to reach the PCP during the current event establishes that the ED was an appropriate level of care for that child.	6

Parent report that they were referred into the ED by phone contact a clinician establishes that the ED was an appropriate level of care for that child.	**8**
Parent report that they were referred to the ED after being seen by a clinician establishes that the ED was an appropriate level of care for that child.	**9**
Parent report that the child did not respond to a dose of a rescue medication establishes that the ED was an appropriate level of care for that child.	6
Parent report that they are unable to afford needed asthma medications establishes that the ED was an appropriate level of care for that child.	3
Parent report that they are unable to obtain needed care because of financial barriers establishes that the ED was an appropriate level of care for that child.	3

The proportion of visits found to be appropriate varied by age and there are biological reasons that make plausible such differences not only being related to health services. Therefore we have specified this measure to be reported as stratified by age. Our data showed that within the 2 – 5 year age group, 54.3% were appropriate, within the 6 – 11 year age group, 44.3% were appropriate and within the 12 – 18 year age group, 48.3% were appropriate, p =.019. The breakdown is as follows:

- For children 2-5: 181 of 335 audits (54.3%) were deemed appropriate.
- For children 6-11: 209 of 477 audits (43.8%) were deemed appropriate.
- Adolescents aged 12-18: 165 of 341 audits (48.4%) were deemed appropriate based upon information in the chart audit.

Criteria for appropriateness that were met were recorded and did vary by age.

1. Brook, R.H., et al., *A method for the detailed assessment of the appropriateness of medical technologies.* International journal of technology assessment in health care, 1986. **2**(01): p. 53-63.
2. Kosecoff, J., et al., *The appropriateness of using a medical procedure: is information in the medical record valid?* Med Care, 1987: p. 196-201.
3. Gilbert, E.H., et al., *Chart reviews in emergency medicine research: where are the methods?* Ann Emerg Med, 1996. **27**(3): p. 305-308.
4. Worster, A., et al., *Reassessing the methods of medical record review studies in emergency medicine research.* Ann Emerg Med, 2005. **45**(4): p. 448-51.

The development team's goal was to develop an ICD10 code set that was fully consistent with the intent of the original measure. Our process began by performing general equivalency mapping using the forward mapping from www.icd9data.com. We then did a de novo review of the CMS ICD 10 CM set to seek to identify codes that might be appropriate for asthma. We reviewed potential codes identified by both sources and developed a new list of codes appropriate for inclusion criteria and a new list of codes appropriate for exclusion criteria. Drs. Kleinman and Sharma reviewed the lists independently and then achieved consensus in a conference call review and discussion. Key team members for this work were Suzanne Lo, MPH who staffed and coordinated this work, Sandeep Sharma, MD, Dr.PH and Lawrence Kleinman, MD, MPH. Dr. Sharma was a lead developer for one of CAPQuaM's 2 asthma measures and Dr. Kleinman is both CAPQuaM PI and was a lead developer for both measures. The guidance for the intended constructs for both ICD9 and ICD10 coding were the findings from a RAND style modified Delphi panel that incorporated 9 national experts over the course of the measure development process.

2b2.3. What were the statistical results from validity testing? (*e.g., correlation; t-test*)
NUMERATOR DATA ELEMENT ASSESSMENT:

We assessed the reliability of data abstraction. Three reviewers each reviewed 10 charts early in training and after further practice. This results in 180 comparisons with the trainer (6 clinical constructs * 3 * 10 = 180). Another 30 comparisons may be made based upon the global scoring chart review as appropriate if any criterion was met. The 6 constructs were findings from the chart review : retractions; accessory muscles being used; markedly reduced breath sounds; ED visit resulted in hospitalization, oxygen saturation was documented in the ED below 90%; PCP referred the patient into the ED. In our pretesting we found that accessory muscle use did not enhance the sensitivity of the appropriateness construct, so we integrated the two with grunting into a category of labored breathing.

	Agreement	
Construct	**Initial Kappa**	**Final Kappa**

7. Retractions	0.67	0.87
8. Accessory Muscle Use	0.44	0.89
9. Markedly diminished BS	0.71	0.78
10. Hospitalized from ED	1.0	1.0
11. O2 sat < 90%	0.79	NA*
12. Referred by PCC	1.0	NA*
All six combined	0.76	0.68
Overall: Appropriateness	0.77	0.87

* NA is because there was no variability in the charts reviewed. There was no disagreement in any of the assessments.

The key assessment (since it is the bottom line) is agreement regarding appropriateness, which is the highlighted row near the bottom of the table. We found that after training and with practice kappa moved from an already strong 0.77 to an excellent 0.87, confirming excellent reliability at the level of the numerator. We confirm that non-medically trained research assistants can be trained to do sufficient quality data abstractions to assess appropriateness of ED visits using these specifications.

IDENTIFYING A POPULATION WITH ASTHMA:

For the foundational NCQA work, NCQA's field test retested a number of previously validated criteria for identifying an eligible population with persistent asthma using administrative claims data. Using the dataset provided, NCQA examined several different scenarios to determine the effects of different specification criteria on this particular population. This information was combined with multiple years of HEDIS data collection of this measure to examine the reliability of collecting this measure through administrative claims.

Score level reliability of the HEDIS 2011 submissions (2010 data) was assessed using the beta-binomial model. Beta-binomial is a better fit when estimating the reliability of simple pass/fail rate measures as is the case with most HEDIS® health plan measures. The beta-binomial model assumes the plan score is a binomial random variable conditional on the plan´s true value that comes from the beta distribution. The beta distribution is usually defined by two parameters, alpha and beta. Alpha and beta can be thought of as intermediate calculations to get to the needed variance estimates. The beta distribution can be symmetric, skewed or even U-shaped.

Reliability used here is the ratio of signal to noise. The signal in this case is the proportion of the variability in measured performance that can be explained by real differences in performance. A reliability of zero implies that all the variability in a measure is attributable to measurement error. A reliability of one implies that all the variability is attributable to real differences in performance. The higher the reliability score, the greater is the confidence with which one can distinguish the performance of one plan from another. A reliability score greater than or equal to 0.7 is considered very good.

We cite these not as specific evidence of score level performance of the submitted measure, but as evidence that the HEDIS measures that rely on the same administrative data elements for their denominator have the capacity to distinguish signal to noise at a very high level. If the population assessment were inadequate, then these other measures which use the same data elements to establish their denominators could not achieve such high reliability scores. This is because failure to distinguish signal from noise at the level of the HEDIS denominators would lead to non-differential misclassification error which is a major bias towards the null, in other towards noise and away from signal. Hence these provide strong indirect evidence of the validity of our approach to capturing the measure's denominator.

While there is moderate agreement (kappa = 0.45 - 0.50) when comparing administrative data regarding the presence of constructs such as recent asthma attacks, use of asthma medications, attack or medication, attack and medication, using

1 year of administrative claims data to parent report, the agreement improves from 0.55 to 0.60 when using two years of data. (Huzel, L. et al. Diagnosing Asthma: The fit between survey and administrative database. Canada Resp. Journal 2002.) We expect that these kappas would be significantly higher were the analyses restricted to children with disease that met our construct criteria for identifiable asthma.

Further, we identify asthma visits and medications using the same data that an insurance company or Medicaid would use for payment, including ICD9 codes, CPT codes, and revenue codes. We have had conversations with expert coders and New York State Department of Health Office of Health Insurance Programs to confirm our choices.

The literature also supports our work. ICD-09 and ICD-10 codes for asthma on patients' medical charts typically match claims data. ICD-9-CM administrative data have been validated using various methodologies for various purposes (5-17). Studies have shown high sensitivity and specificity for diagnoses obtained from administrative data among children with high-risk conditions including asthma, (18), and high predictive value among adolescents and adults with asthma. (19) (20) HEDIS criteria using administrative data support peer reviewed research, for example in patients with persistent asthma based on HEDIS criteria in five Medicaid programs (Colorado, Georgia, Indiana, New Jersey, Washington) using ICD-9-CM code 493.x (21). Fowles and colleagues report sensitivity and specificity of claims compared with ambulatory medical records to identify asthma was 0.82 and 0.99, respectively. (22) Wilchesky compared chart abstraction to diagnoses obtained from administrative database: asthma claims were highly specific, Sp= 96.76 (95%CI 96.5, 97.0). (23) Bronstein et al found that 88.3% of diagnoses asthma on claims agreed with medical record, with a negative predictive value of 0.85 and a positive predictive value of 0.88.They conclude that claims are generally an accurate indicator of the content of a patient encounter. (24) Steinwachs et al. compared billed claims to medical records based on date of visit and diagnosis, they found for asthma there was 90.9 percent of billed visits in record on same date and 82.8 percent of billed visits with same diagnosis in record on same date. (25) Quan et al documented the validity of ICD-9-CM and ICD-10 coding systems in coding clinical information and found that ICD-10 data was generally comparable with that of ICD-9-CM data in recording clinical information. (26)

From a public health perspective, asthma surveillance systems in several states, including Maine, North Carolina, Connecticut and Michigan, have shown the feasibility of using administrative data to identify children having asthma, based on primary and secondary diagnosis codes reported on inpatient and outpatient claims. (27-30) Researchers also classified children with evidence of persistent asthma using HEDIS criteria, (31). Another study showed the usefulness of ICD9 493.x to identify asthma for a quality measure using Maryland data. Like our measure, those researchers excluded children with a diagnosis of cystic fibrosis (ICD9 277). (32) regarding our capacity to identify exclusions, Quan et al found that claims had a PPV of 91.9, and a negative predictive value of 92.6, with *k* of 0.65 (substantial agreement[1]) compared to chart review. FICD 10 performed similarly in this study.

REFERENCES

(1) National Committee for Quality Assurance. (1 994). Health care quality improvement studies in managed care set1ing.s: A guide for state medicaid agencies. Washington, DC: Author.

(2) Leatherman, S., Peterson, E., Heinen, L., & Quam, L. (1991). Quality screening and management using claims data in a managed care setting. Quality Review Bulletin, 17, 349-359.

(3) Parente, S.T., Weiner,J.P., Garnick, D.W., Richards, T.M., Fowles,J., Lawthen, A.G., Chandler, P., & Palmer, R.H. (1995). Developing a quality improvement database using health insurance data: A guided tour with applications to Medicare's National Claims History File. American Journal of Medical Quality, 10(4), 162-176.

(4) Weiner, J.P., Powe,. R., Steinwachs, D.M., & Dent, G. (1990). Applying insurance claims data to assess quality of care: A compilation of potential indicators. Quality Review Bulletin, 16, 424-438.

(5) Best, W.R., S.F. Khuri, M. Phelan, K. Hur, W.G. Henderson, J.G. Demakis, and J. Daley. 2002. ''Identifying Patient Preoperative Risk Factors and Postoperative Adverse Events in Administrative Databases: Results from the Department of Veterans Affairs National Surgical Quality Improvement Program.'' Journal of American College of Surgeons 194: 257–66.

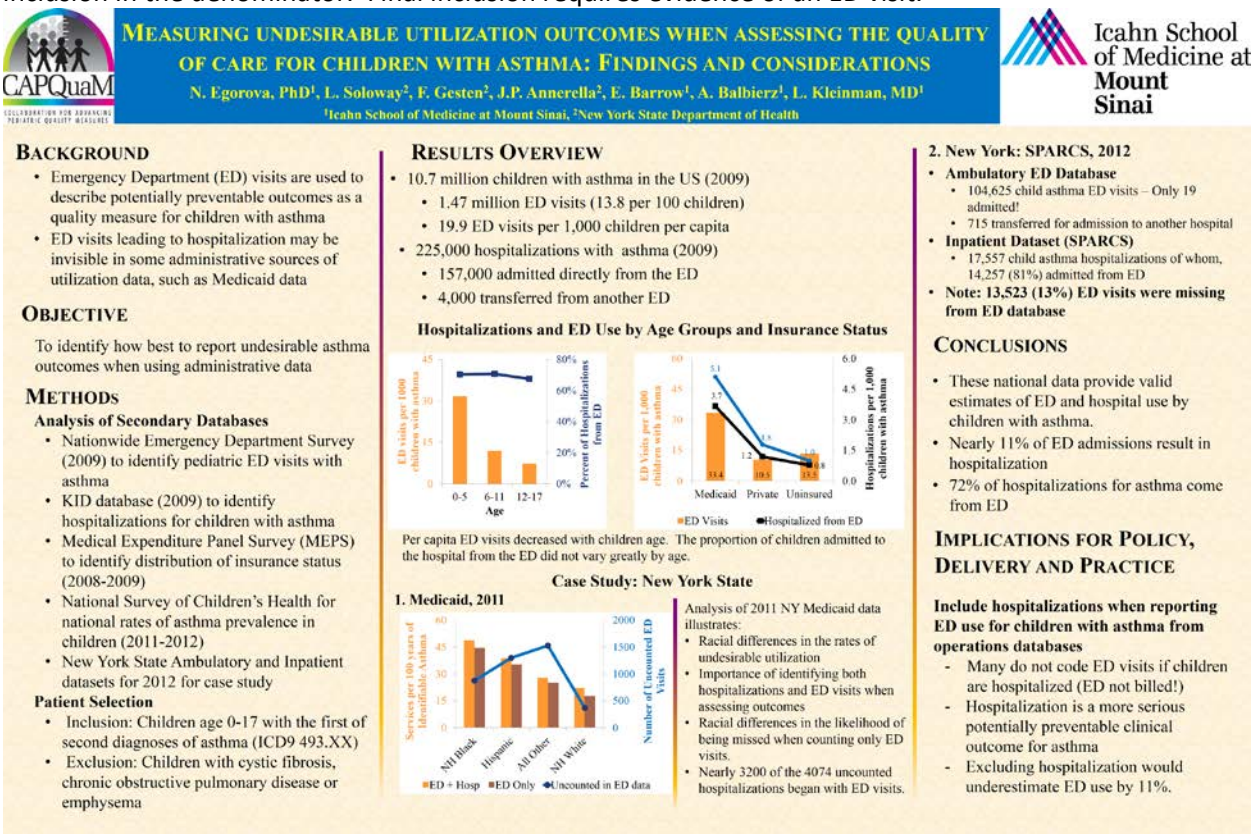
(6) Geraci, J.M., C.M. Ashton, D.H. Kuykendall, M.L. Johnson, and L. Wu. 1997. ''International Classification of Diseases, 9th Revision, Clinical Modification Codes in Discharge Abstracts are Poor Measures of Complication Occurrence in Medical Inpatients.'' Medical Care 35: 589–602.

(7) Hsia, D.C., C.A. Ahern, B.P. Ritchie, L.M. Moscoe, and W.M. Krushat. 1992. ''Medicare Reimbursement Accuracy under the Prospective Payment System, 1985 to 1988.'' Journal of the American Medical Association 268: 896–9.

(8) Iezzoni, L.I., S. Burnside, L. Sickles, M.A.Moskowitz, E. Sawitz, and P.A. Levine. 1988. ''Coding of Acute Myocardial Infarction. Clinical and Policy Implications.'' Annals of Internal Medicine 109: 745–51.
(9) Jollis, J.G., M. Ancukiewicz, E.R. DeLong, D.B. Pryor, L.H. Muhlbaier, and D.B. Mark. 1993. ''Discordance of Databases Designed for Claims Payment versus Clinical Information Systems. Implications for Outcomes Research.'' Annals of Internal Medicine 119: 844–50.
(10) Lee, D.S., L. Donovan, P.C. Austin, Y. Gong, P.P. Liu, J.L. Rouleau, and J.V. Tu. 2005. ''Comparison of Coding of Heart Failure and Comorbidities in Administrative and Clinical Data for Use in Outcomes Research.'' Medical Care 43: 182–8.
(11) Muhajarine, N., C. Mustard, L.L. Roos, T.K. Young, and D.E. Gelskey. 1997. ''Comparison of Survey and Physician Claims Data for Detecting Hypertension.'' Journal of Clinical Epidemiology 50: 711–8.
(12) Quan, H., G. Parson, and W. Ghali. 2002. ''Validity of Information on Comorbidity Derived from ICD-9-CM Administrative Data.'' Medical Care 40: 675–85.
(13) Romano, P.S., B.K. Chan, M.E. Schembri, and J.A. Rainwater. 2002. ''Can Administrative Data Be Used to Compare Postoperative Complication Rates across Hospitals?'' Medical Care 40: 856–67.
(14) Romano, P.S., and D.H. Mark. 1994. ''Bias in the Coding of Hospital Discharge Data and Its Implications for Quality Assessment.'' Medical Care 32: 81–90.
(15) Weingart, S.N., L.I. Iezzoni, R.B. Davis, R.H. Palmer, M. Cahalane, M.B. Hamel, K. Mukamal, R.S. Phillips, D.T. Davies Jr., and N.J. Banks. 2000. ''Use of Administrative Data to Find Substandard Care: Validation of the Complications Screening Program.'' Medical Care 38: 796–806.
(16) Yasmeen, S., P.S. Romano, M.E. Schembri, J.M. Keyzer, and W.M. Gilbert. 2006. ''Accuracy of Obstetric Diagnoses and Procedures in Hospital Discharge Data.'' American Journal of Obstetrics and Gynecology 194: 992–1001.
(17) Quam, L., Ellis, L. B. M., Venus, P., Clouse, J., Taylor, C. G., & Leatherman, S. (1993). using claims data for epidemiologic research. the concordance of claims-based criteria with the medical recordand patient survey for identifying a hypertensive population. Medical Care, 31(6), 498-507.
(18) Daley MF, Barrow J, Pearson K, Crane LA, Gao D, Stevenson JM, et al. Identification and recall of children with chronic medical conditions for influenza vaccination. Pediatrics 2004;113(1 Pt 1):e26-33.
(19) Vollmer WM, O'Connor EA, Heumann M, Frazier EA, Breen V, Villnave J, Buist AS. Searching multiple clinical information systems for longer time periods found more prevalent cases of asthma. J Clin Epidemiol 2004;57:392-7.
(20) Twiggs, J. E., Fifield, J., Apter, A. J., Jackson, E. A., & Cushman, R. A. (2002). Stratifying medical and pharmaceutical administrative claims as a method to identify pediatric asthma patients in a medicaid managed care organization. Journal of Clinical Epidemiology, 55(9), 938-944.
(21) Samnaliev, M., Baxter, J. D., & Clark, R. E. (2009). Comparative evaluation of two asthma care quality measures among medicaid beneficiaries. Chest, 135(5), 1193-1196.
(22) Fowles, J. B., Fowler, E. J., & Craft, C. (1998). validation of claims diagnoses and self-reported conditions compared with medical records for selected chronic diseases. Journal of Ambulatory Care Management, 21(1), 24-34.
(23) Wilchesky, M., Tamblyna, R. M., & Huang, A. (2004). Validation of diagnostic codes within medical services claims. Journal of Clinical Epidemiology, 57, 131-141.
(24) Bronstein, J. M., Santer, L., & Johnson, V. (2000). The use of medicaid claims as a supplementary soure of information on quality of asthma care. Journal for Healthcare Quality, 22(6), 13-18.
(25) Steinwachs, D. M., Stuart, M. E., Scholle, S., Starfield, B., Fox, M. H., & Weiner, J. P. (1998). A comparison of ambulatory medicaid claims to medical records: A reliability assessment. American Journal of Medical Quality, 13(2), 63-69.
(26) Quan, H., Li, B., Saunders, L. D., Parsons, G. A., Nilsson, C. I., Alibhai, A., et al. (2008). Assessing validity of icd-9-cm and icd-10 administrative data in recording clinical conditions in a unique duallycoded database. HSR: Health Services Research, 43(4), 1424.
(27) Deprez RD, Asdigian NL, Oliver LC, Anderson N, Caldwell E, Baggott LA. Development of a prototype system for statewide asthma surveillance. Am J Public Health 2002;92:1946-51.
(28) Buescher PA, Jones-Vessey K. Using Medicaid data to estimate state- and county-level prevalence of asthma among low-income children. Matern Child Health J 1999;3:211-6.
(29) Children's Health Council. Asthma affects nearly one in ten children enrolled in HUSKY Part A in 1999–2000. Hartford (CT): Children's Health Council; 2001.
(30) Dombkowski, K. J., Wasilevich, E. A., & Lyon-Callo, S. K. (2005). Pediatric asthma surveillance: Using medicaid claims. Public Health Reports, 120, 515.
(31) National Committee for Quality Assurance. Use of appropriate medications for people with asthma. HEDIS 2003, volume 2: technical specifications. Washington: NCQA; 2003.
(32) Stuart, B., Singhal, P. K., Magder, L. S., & Zuckerman, I. H. (2003). How robust are health plan quality indicators to data loss? A monte carlo simulation study of pediatric asthma treatment. HSR: Health Services Research, 38(6), 1547.
(33) Schneeweiss, S., & Avorn, J. (2005). A review of uses of health care utilization databases for epidemiologic research on therapeutics. Journal of Clinical Epidemiology, 58(4), 323-337.

Our own research looking at NY State Medicaid and national all payer data (see poster presented at peer-reviewed AcademyHealth national meeting) is consistent with expert and other recommendations that to identify all ED visits, one also needs to include hospitalizations for asthma as potential indicators of an otherwise unrecognized ED visit, which we have done and incorporated into the specifications.

This is the poster presenting our original research regarding the inclusion of hospitalizations when considering potential inclusion in the denominator. Final inclusion requires evidence of an ED visit.



MEASURING UNDESIRABLE UTILIZATION OUTCOMES WHEN ASSESSING THE QUALITY OF CARE FOR CHILDREN WITH ASTHMA: FINDINGS AND CONSIDERATIONS

N. Egorova, PhD[1], L. Soloway[2], F. Gesten[2], J.P. Annerella[2], E. Barrow[1], A. Balbierz[1], L. Kleinman, MD[1]

[1]Icahn School of Medicine at Mount Sinai, [2]New York State Department of Health

Icahn School of Medicine at Mount Sinai

BACKGROUND

- Emergency Department (ED) visits are used to describe potentially preventable outcomes as a quality measure for children with asthma
- ED visits leading to hospitalization may be invisible in some administrative sources of utilization data, such as Medicaid data

OBJECTIVE

To identify how best to report undesirable asthma outcomes when using administrative data

METHODS

Analysis of Secondary Databases

- Nationwide Emergency Department Survey (2009) to identify pediatric ED visits with asthma
- KID database (2009) to identify hospitalizations for children with asthma
- Medical Expenditure Panel Survey (MEPS) to identify distribution of insurance status (2008-2009)
- National Survey of Children's Health for national rates of asthma prevalence in children (2011-2012)
- New York State Ambulatory and Inpatient datasets for 2012 for case study

Patient Selection

- Inclusion: Children age 0-17 with the first of second diagnoses of asthma (ICD9 493.XX)
- Exclusion: Children with cystic fibrosis, chronic obstructive pulmonary disease or emphysema

RESULTS OVERVIEW

- 10.7 million children with asthma in the US (2009)
  - 1.47 million ED visits (13.8 per 100 children)
  - 19.9 ED visits per 1,000 children per capita
- 225,000 hospitalizations with asthma (2009)
  - 157,000 admitted directly from the ED
  - 4,000 transferred from another ED

Hospitalizations and ED Use by Age Groups and Insurance Status

Per capita ED visits decreased with children age. The proportion of children admitted to the hospital from the ED did not vary greatly by age.

Case Study: New York State

1. Medicaid, 2011

Analysis of 2011 NY Medicaid data illustrates:

- Racial differences in the rates of undesirable utilization
- Importance of identifying both hospitalizations and ED visits when assessing outcomes
- Racial differences in the likelihood of being missed when counting only ED visits.
- Nearly 3200 of the 4074 uncounted hospitalizations began with ED visits.

2. New York: SPARCS, 2012

- **Ambulatory ED Database**
  - 104,625 child asthma ED visits – Only 19 admitted!
  - 715 transferred for admission to another hospital
- **Inpatient Dataset (SPARCS)**
  - 17,557 child asthma hospitalizations of whom, 14,257 (81%) admitted from ED
- **Note: 13,523 (13%) ED visits were missing from ED database**

CONCLUSIONS

- These national data provide valid estimates of ED and hospital use by children with asthma.
- Nearly 11% of ED admissions result in hospitalization
- 72% of hospitalizations for asthma come from ED

IMPLICATIONS FOR POLICY, DELIVERY AND PRACTICE

Include hospitalizations when reporting ED use for children with asthma from operations databases

- Many do not code ED visits if children are hospitalized (ED not billed!)
- Hospitalization is a more serious potentially preventable clinical outcome for asthma
- Excluding hospitalization would underestimate ED use by 11%.

2b2.4. 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?*)

Our interpretation is that administrative data are reliable and valid for identifying asthma, and that year to year test retest reliability seems to indicate similar patterns of performance when identifying ED visits for asthma, reinforcing the reliability of our operational definitions for identifying eligible children. Our specification provide a sensitive and face valid approach to identifying an unbiased sample of children with ED visits(ensuring we don't bias the results towards the inappropriate by missing those with hospitalization).

Most databases contain consistent elements, are available in a timely manner, provide information about large numbers of individuals, and are relatively inexpensive to obtain and use. Validity of many databases has been established, and their strengths and weaknesses relative to data abstracted from medical records and obtained via survey have been documented (30). Administrative data are supported, if not encouraged by federal agencies, such as NIH, AHRQ, HCFA, and the VA. The Centers for Medicare & Medicaid Services has made clear to the participating AHRQ-CMS CHIPRA Centers of Excellence funded to develop measures in the Pediatric Quality Measures Program that it places a premium on feasibility when assessing those measures that it will most highly recommend to states to complete. The sources of data for the existing measure and other similar measures are typically based upon administrative data as well, providing consensual validation for using administrative data as the primary data source.

Our Kappa results indicated excellent agreement in the reliability of the chart audit. Kappa values over 0.75 are considered excellent, 0.40 to 0.75 as fair to good, and below 0.40 as poor.

2b3. EXCLUSIONS ANALYSIS

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

2b3.1. 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*)

Exclusions were only included if they were endorsed by the expert panel. In studying the denominator we found that a very few percent of potentially eligible children (<=2.5%) were excluded by clinical diagnoses. The use of three months of continuous enrollment was recommended by our multi-stakeholder consortium and avoids the exclusion of more than 20% of otherwise eligible children from the population with identifiable asthma compared to a 12 month requirement.

Denominator Exclusions
Children with concurrent or pre-existing: Chronic Obstructive Pulmonary Disease (COPD) diagnosis (ICD-9 Code: 496), Cystic Fibrosis diagnosis (ICD-9 code 277.0, 277.01. 277.02, 277.03, 277.09), or Emphysema diagnosis (ICD-9 code 492xx).

Children who have not been consecutively enrolled in the reporting plan for at least two months prior to the index reporting month, as well as the index reporting month itself.

There are no numerator exclusions.

2b3.2. 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*)

2b3.3. 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*)

Exclusions are clinical and represent construct validity rather than statistical considerations.

2b4. 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.

2b4.1. 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** 1 **risk categories**

☒ **Other, There are optional categories for stratification of outcomes, such as race/ethnicity that are for descriptive and not risk stratification purposes. The NHLBI guideline clearly articulates a preference for no such stratifications based upon race/ethnicity, insurance, etc.**

2b4.1.1 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.

2b4.2. If an outcome or resource use component measure is not risk adjusted or stratified, provide rationale and analyses to demonstrate that controlling for differences in patient characteristics (case mix) is not needed to achieve fair comparisons across measured entities.

Specifications for this measure requires stratification by age group. Several additional stratifications are optional but may be required by the accountability entity or provided by the reporting entitity. These variables include race/ethnicity, rurality/urbanicity and county level of poverty.

Within age group, we specify a number of stratifications as we have done for all of our CAPQuaM PQMP measure. Absent clear biological evidence that ED visits should be more likely in any of the sub categories we have chosen not to adjust but to report both topline and stratified results.

The NIH NHLBI NAEPP (http://www.nhlbi.nih.gov/health-pro/guidelines/current/asthma-guidelines/full-report) guideline notes that goals of care and definition of successful management are the same regardless of baseline presentation. Hence clinical risk adjustment is not appropriate.

(page 38)

> ***"An important point linking asthma severity, control, and responsiveness is that the goals are identical for all levels of baseline asthma severity.*** *A patient who has severe persistent asthma compared to a patient who has mild persistent asthma, or a patient who is less responsive to therapy may require more intensive intervention to achieve well-controlled asthma; however, the goals are the same: in well-controlled asthma, the manifestations of asthma are minimized by therapeutic intervention."*

High levels of appropriateness suggest that the children in the ED are there because of an immediate clinical need and the ED service is well utilized. Some of these may have been preventable with better quality care prior to the ED visit and some will not. When appropriateness sis high, Asthma ED visit rates represent a strong proxy for asthma clinical outcomes.

Low levels of appropriateness suggest that the cause of many ED visits is not break through asthma or failures of biological asthma management, but insufficient access or quality of care provided that families are seeking care in the ED as preferential to a less acute setting. The good news in such a finding is that clinical asthma outcomes are better than would appear simply by counting the number of ED visits.

The results have independent meaning but from both accountability and improvement perspectives there is synergy in the interpretation of this measure with the CAPQuaM rate of ED visits in asthma measure.

2b4.3. Describe the conceptual/clinical and statistical methods and criteria used to select patient factors (clinical factors or sociodemographic 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 conceptual model is that of CAPQuaM that includes that in pediatrics age is a key predictor and stratification is valuable. We were asked by AHRQ and CMS to include other constructs and we have manifest them as specified, such as race/ethnicity, poverty level in the caregivers county of residence, rurality/urbanicity on the caregiver's county of residence, insurance type and plan type, when variable. We have not added a stratum for children with special health care needs since asthmatics going to the emergency room are highly likely to belong in this category.

2b4.4a. What were the statistical results of the analyses used to select risk factors?
There are statistically significant differences by age group.

2b4.4b. Describe the analyses and interpretation resulting in the decision to select SDS 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)

2b4.5. 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*)

Provide the statistical results from testing the approach to controlling for differences in patient characteristics (case mix) below.
If stratified, skip to 2b4.9

2b4.6. Statistical Risk Model Discrimination Statistics (*e.g., c-statistic, R-squared*)**:**

2b4.7. Statistical Risk Model Calibration Statistics (*e.g., Hosmer-Lemeshow statistic*):

2b4.8. Statistical Risk Model Calibration – Risk decile plots or calibration curves:

2b4.9. Results of Risk Stratification Analysis:

For results of age-stratified analysis, please refer to section 2b4.4a

2b4.10. 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*)

Our medical chart audit found that the measure varies by race/ethnicity.

Appropriateness varied by age (Chi^2=8.2,p=.02), with younger (p=.01) and school aged (p=.01) children each being significantly different; Adolescents experienced a level of appropriateness intermediate to the other two groups and were not significantly different from them when combined (ie comparing Adolescents to All others). We also found racial differences with Hispanics at 44.1% appropriateness, non-Hispanic Blacks at 51.3%, Whites at 56.5% and all others at 72.2%. Chi square with 3 degrees of freedom was 15.4, with p=.0015. The appropriateness of ED visits for Hispanic children was less than for other children (p=.002).

Hispanic children had higher rates of questionable use of the ED (55.9% of visits) when compared to non-Hispanic children (46.8%), p=.002. Black children showed a trend toward more questionable use compared to all other children (53.6% questionable vs 48.7%, p=.10).

2b4.11. 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*)

2b5. IDENTIFICATION OF STATISTICALLY SIGNIFICANT & MEANINGFUL DIFFERENCES IN PERFORMANCE
2b5.1. 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)*

Contingency table analysis with chi square, SAS 9.4 Generalized linear models (Proc GLM) and SAS 9.4 Logistic Regression (Proc Logistic) analyses were performed and were coherent and each illustrated the presence of statistical differences among identifiable subgroups. All of this work was done in a single hospital facility.

2b5.2. 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*)

The proportion of visits found to be appropriate varied by age and there are biological reasons that make plausible such differences not only being related to health services. Therefore we have specified this measure to be reported as stratified by age. Our data showed that:

- For children 2-5: 181 of 335 audits (54.3%) were deemed appropriate.
- For children 6-11: 209 of 477 audits (43.8%) were deemed appropriate.
- Adolescents aged 12-18: 165 of 341 audits (48.4%) were deemed appropriate

based upon information in the chart audit.

Criteria for appropriateness that were met were recorded and did vary by age

The GLM models regressed appropriateness simultaneously on the class variables Age Group, Ethnicity, Gender, and presence or absence of private insurance found that gender (P=.017), Hispanic ethnicity (p=.002), and private insurance (p=.005) were all significantly associated with level of appropriateness, as was age group (p=.009). For this analysis, N=1,188 with a model F value of 6.56 (Pr>F is <0.0001).

To confirm the distinction between what we expected to be strong and weak effects, we substituted day of week for the various demographic variables other than age group. The P value for day of week (as a class variable) was >0.30. The non-zero effect size is consistent with social science literature that suggests that variables such as time of day and day of week are weakly meaningful. Still, the lack of a significant finding in a reasonably good-sized data set demonstrates that spurious significant findings are not likely to be identified as significant.

Differences between major subgroups were statistically significant, including race/ethnicity, age group, and insurance status. We note that this is a stricter test than had the measure been assessed across different entities. These data showed differences by type of insurance, which in this case can serve as proxy for health plan. The F Value for Insurance Status (after controlling for age group) was 3.91 with 4 degrees of freedom, which exceeded the critical value and is associated with a p-value <0.004. This should correlate with excellent capacity to distinguish between health plans.

2b5.3. 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?*)

In sum, we found true signal in social determinants (consistent with the asthma literature) and did not incorrectly identify weak signal as meaningful. The measure distinguishes signal from noise.

The measures are sensitive enough to detect meaningful differences as observed within a population (as described above). Since the sum of squares across populations is expected to be greater in distinct populations, we expect the measure to perform very well when comparing across populations as well. Since the effective sample size of within population comparisons (such as we have conducted) is diminished by an (unmeasured) intraclass correlation coefficient, we would expect greater power for equal sample size to detect differences between entities than we had in our testing of various subpopulations within a single state. This supports the same conclusion. The signal to noise ratio is very strong for this measure.

2b6. COMPARABILITY OF PERFORMANCE SCORES WHEN MORE THAN ONE SET OF SPECIFICATIONS
If only one set of specifications, this section can be skipped.

Note*: This item is directed to measures that are risk-adjusted (with or without SDS 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 SDS 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.***

2b6.1. 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*)

2b6.2. 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*)

2b6.3. 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*)

2b7. MISSING DATA ANALYSIS AND MINIMIZING BIAS

2b7.1. 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*)

Chart review data has been shown to be an accurate method for identifying the presence or absence of conditions required to identify the level of appropriateness of a clinical service. Documentation is a part of the clinical responsibility and failure to document is a quality deficit that is not construed as missing data. Since inclusion requires the affirmative presence of data and we are unaware of any evidence to suggest that there would be differential absence of data between appropriate and non-appropriate visits we are not concerned about introducing bias in our findings. Further, we use random sampling of eligible visits as another means to avoid the introduction of bias. (1, 2)

1. Kahn, KL, Kosecoff, J, Chassin, MR, et al. Measuring the clinical appropriateness of a procedure: Can we do it? Medical Care, 1988, 26:415-422.
2. Kosecoff, J., et al., The appropriateness of using a medical procedure. Is information in the medical record valid? Med Care, 1987. **25**(3): p. 196-201.

While assessing the definition for identifiable asthma, our colleagues at NY State Medicaid conducted a series of iterative analysis using NY State Medicaid Managed care data to assess the importance of our data elements and definitions. These analyses helped to confirm the importance of using, for example, both revenue codes and procedure codes to identify ED visits. These analyses also confirmed that the use of pharmaceutical data to identify children with asthma expanded the pool of these so identified and quantified that statewide doing so added around 10,000 to a total

of around 190,000 children with identifiable asthma in the state. We found no evidence that this was a threat to the measure's validity. The key reason for inclusion of pharmacy data is that our expert panel directed us to use it and it is a slightly more sensitive way to identify asthmatic children from the pool of all children with asthma related claims. The expert panel did not want the absence of pharmacy data to preclude inclusion of a reporting entity in the measure or to exempt any entity from measurement. We do not have either direct access to the data or a copy of all the iterative analyses at this time or we would include more specific data to demonstrate these findings. The analyses were in hand and were incorporated into our decision-making at the time that we developed the specifications.

2b7.2. 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*)

See section above.

2b7.3. 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*)
Not biased.

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), Coded by someone other than person obtaining original information (e.g., DRG, ICD-9 codes on claims), Abstracted from a record by someone other than person obtaining original information (e.g., chart abstraction for quality measure or registry)

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**.

Some data elements are in defined fields in 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).

Detailed clinical data are needed. There are no technical barriers to capturing the necessary data in defined electronic fields in electronic health records. We view NQF endorsement as a step to help us to initiate a conversation to consider such inclusion.

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 a PRO-PM, consider implications for both individuals providing PRO data (patients, service recipients, respondents) and those whose performance is being measured.

We have learned that chart review is a reliable and accepted method of measuring appropriate use. There are no technical barriers to incorporating structured fields to help assess the appropriateness of the visits in conjunction with the criteria outlined above and implemented in this measure, although such fields do not currently exist. We further demonstrated that our measure was able to identify differences in the proportion appropriate, such as those associated with age and race. For example, the overall level of appropriateness for children aged 2-5 was 54%, for children aged 6-1 was 44%, and for adolescents between 12 and 18, 48%. Because of these differences we have chosen to present the measure stratified by age group. We found that use of a clinical database was an inefficient way to identify eligible charts and thus have adapted eligibility criteria that rely on administrative data. Because chart review is relatively time consuming, we have articulated the specifications in a way that

represents a hybrid whereby administrative data can qualify a proportion of numerator events without chart review. Our paper data collection tool underwent a number of revisions for time and data collection efficiency and the chart review team demonstrated excellent agreement in data collection with a group kappa of .87 in identifying numerator events. Although the chart collection tool is not a formal part of this measure, we would be happy sharing a general version (data collection template) of it upon request.

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)*.
None at present.

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)
Not in use	

4a.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

N/A

4a.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?*)
We are awaiting NQF endorsement for use. There are no policies or actions of the developer/steward or accountable entities that would restrict access to performance results of impeded implementation. Some potential users are awaiting NQF endorsement.

The topic of ED asthma overuse was assigned to our measure development project in the Pediatric Quality Measures Program by CMS, by far the largest single third party payer for medical care for children in the US, and by AHRQ. Major federal policy makers have indicated to us that these measures are a priority. This measure has received the imprimatur of the American Academy of Pediatrics as one of its high priority measures that emerged from their joint (with the American Board of Pediatrics) Measurement Alignment and Strategic Selection Work Group.

We have begun a dialogue with the CDC to consider use of this measure to serve their interests.

4a.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.*)

The CAPQuaM team includes multiple stakeholders, including payers and state agencies. Several are interested in using this measure and are awaiting NQF endorsement. As a part of our CAPQuaM work we will disseminate and assist in the implementation of this measure subsequent to endorsement. This measure has been approved for inclusion in the National Quality Measures Clearinghouse.

As noted above, the topic of ED asthma overuse was assigned to our measure development project in the Pediatric Quality Measures Program by CMS, by far the largest single third party payer for medical care for children in the US, and by AHRQ. Major federal policy makers have indicated to us that these measures are a priority. This measure has received the imprimatur of the American Academy of Pediatrics as one of its high priority measures that emerged from their joint (with the American Board of Pediatrics) Measurement Alignment and Strategic Selection Work Group.

We have begun a dialogue with the CDC to consider use of this measure to serve their interests.

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.

4b. **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.
Not in longitudinal use. As noted above, both high and low levels of appropriateness are interpretable and actionable as outcomes of asthma management. This measure of process provides information regarding the outcomes of asthma care – both access to care and quality of management. Its interpretation is synergistic with the CAPQuaM rate of asthma ED visit measure also developed in the PQMP and currently under review at NQF.

4c. 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).

4c.1. **Please explain any unexpected findings (positive or negative) during implementation of this measure including unintended impacts on patients.**
None observed.

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

4d1.1. **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.
The CAPQuaM team includes payers, MCO´s, state health programs, consumers, Accreditors, family advocates, clinicians, hospitals, and others and all have had the opportunity to participate in the dialogue that led to the measure development and to the interpretation of its findings

4d1.2. **Describe the process(es) involved, including when/how often results were provided, what data were provided, what educational/explanatory efforts were made, etc.**

4d2.1. **Summarize the feedback on measure performance and implementation from the measured entities and others described in 4d.1.**
Describe how feedback was obtained.

Steering committee meetings, conference calls, email

4d2.2. **Summarize the feedback obtained from those being measured.**
We don´t distinguish by source of feedback, please see 4d2.3

4d2.3. **Summarize the feedback obtained from other users**
This measure has been received enthusiastically by our stakeholder partners. We highlight feedback in an area for which we received comment in the prior review of this measure.

This measure is unusual in that it offers value and opportunity for improvement regardless of results. Low levels of appropriateness suggest inefficiencies in primary care, as children who probably do not need to be in the ED because of the severity of their disease are nonetheless being brought to the ED by their caregivers. This may represent failures of asthma education or of availability/access/attractiveness of primary care in the context of acute concerns. Low levels of appropriateness also suggest that whatever measures are being used to assess how frequently children are using the ED are overestimating the clinical failure rate of asthma care, since the ED visits that are presumed to be clinical failures may represent something else.

Conversely, high levels of appropriateness suggest that the ED use for asthma is predominately among those children who are experiencing breakthrough in their clinical asthma. If the overall rate of use is high, this would raise concerns about the effectiveness of the asthma management in ambulatory care. But it does not point specifically to availability/access/attractiveness of primary care as does low rates of effectiveness.

Because of this bi-directionality of the measures interpretation, there is lower risk for gaming than with some other measures.

4d.3. **Describe how the feedback described in 4d.2 has been considered when developing or revising the measure specifications or implementation, including whether the measure was modified and why or why not.**
Feedback incorporated directly into the development process.

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?
Yes

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 **Attachment:** Appendix_Asthma_5_12_12_16.docx

Contact Information

Co.1 **Measure Steward (Intellectual Property Owner):** University Hospitals Cleveland Medical Center
Co.2 **Point of Contact:** Lawrence, Kleinman, drlarrykleinman@gmail.com, 216-286-6969-
Co.3 **Measure Developer if different from Measure Steward:** University Hospitals Cleveland Medical Center
Co.4 **Point of Contact:** Lawrence, Kleinman, drlarrykleinman@gmail.com, 216-286-6969-

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.
Role: Expert Panelists
Elizabeth Allen, M.D Ohio State University
Chitra Dinakar, M.D. Department of Pediatrics University of Missouri-Kansas City
Stephen Teach, M.D., M.P.H. Children's National Medical Center
Charles Macias M.D., M.P.H. Pediatrics, Baylor Texas Children's Hospital
Michael Cabana M.D., M.P.H. University of California, San Francisco

Barbara Yawn M.D., M.S. Olmstead Medical Center, University of Minnesota
Joan Connell M.D. University of North Dakota
Delaney Gracy M.D., M.P.H. The Children's Health Fund and Montefiore Children's Hospital
Sharlene Miner M.D. Emergency Medicine Inter Mountain Medical Center

ROLE: Steering Committee and Investigator
Wilson Pace, MD American Academy of Family Physicians – DARTNET Institute - University of Colorado
Lynn Olson, PhD American Academy of Pediatrics
Christina Bethell, PhD, MBA, MPH Child and Adolescent Health Measurement Initiative, Johns Hopkins University (Previous OHSU)
Elizabeth Howell, MD Icahn School of Medicine at Mount Sinai
Harold Kaplan, MD Icahn School of Medicine at Mount Sinai
Lawrence Kleinman, MD, MPH Icahn School of Medicine at Mount Sinai

Rebecca Anderson Mount Sinai Medical Center
Ira Nash, MD Northshore - Long Island Jewish Medical Center (previous Mount Sinai School of Medicine)
Eyal Shemesh, MD Icahn School of Medicine at Mount Sinai
Mary Barton, MD National Committee on Quality Assurance
Charles Homer, MD, MPH US Department of HHS (previous National Institute for Child Health Quality)
Marla Clayman, PhD American Institutes for Research (previous Northwestern University)
Foster Gesten, MD New York State Dept. of Health, Office of Health Insurance Programs
Jerod M. Loeb, PhD The Joint Commission
Robert Rehm National Committee for Quality Assurance
Steve Kairys, MD American Academy of Pediatrics/QuIIN
Erin DuPree, MD The Joint Commission (previous: Icahn School of Medicine at Mount Sinai)
David Baker The Joint Commission (Previous Northwestern)
Beverley Johnson, BSN* Institute for Patient- and Family-Centered Care
Doris Peter Consumers Union

ROLE: Senior Advisory Board Member and Investigator
Shoshanna Sofaer, DrPH American Institutes for Research (previous CUNY Baruch)
Lynne Richardson, MD Icahn School of Medicine at Mount Sinai
Ian Holzman, MD Icahn School of Medicine at Mount Sinai
Marilyn Kacica, MD New York State Dept. of Health, Division of Family Health

ROLE: Senior Advisory Board Member
Marc Lashley, MDAllied Pediatrics
Gary Mirkin, MD Allied Pediatrics
Barbara Kupferman, RN, CMCN, LNC AmeriChoice by United Healthcare
L. Gregory Pawlson, MD, MPH SE Healthcare QUALITY Consultants (previous Blue Cross Blue Shield)
John Santa, MD* (previous Consumers Union)
John Clarke, MD ECRI, PA Patient Safety Authority
Scott Breidbart, MD Empire Blue Cross Blue Shield/ Anthem
Jeff Terry, MBA GE Healthcare
Robert St. Peter, MD, MPH Kansas Health Institute
Ruth Stein, MD Montefiore Children's Hospital
Arthur Aufses, MD Icahn School of Medicine at Mount Sinai
Eric Rose, MD Icahn School of Medicine at Mount Sinai
Wendy Brennan, MS National Alliance on Mental Illness of New York City (NAMI - NYC Metro)
Laurel Pickering, MPH Northeast Business Group on Health
Tony Hope, BM BCh, PhD Oxford University, The Ethox Center
Martin Hatlie, JD Partnership 4 Pt. Safety, Consumers Advancing Patient Safety
Paul Wise, MD Stanford University
Mark Chassin, MD, MPH, MPP The Joint Commission
Ira Nash, MD Icahn School of Medicine at Mount Sinai
ROLE: Investigators/Key Staff
Rusty McLouth, MS AAFP
Kasey McCracken, MPH CAHMI
Elise Barrow, MPH Icahn School of Medicine at Mount Sinai
Kasey Coyne, BA Icahn School of Medicine at Mount Sinai
Natalia Egorova, PhD Icahn School of Medicine at Mount Sinai
Elizabeth Howell, MD Icahn School of Medicine at Mount Sinai
Harold Kaplan, MD Icahn School of Medicine at Mount Sinai
Melissa Manice, MPH Icahn School of Medicine at Mount Sinai
Joshua Oppenheimer, BA Icahn School of Medicine at Mount Sinai
Eve Oster, MD Pediatric Partners of Gwinnett (previous Icahn School of Medicine at Mount Sinai)
Daniel O'Sullivan, BS University of Virginia (previous Icahn School of Medicine at Mount Sinai)

Michael Parides, PhD	Icahn School of Medicine at Mount Sinai
Barbara Rabin, MHA	Icahn School of Medicine at Mount Sinai
Carolyn Rosen, MD	Icahn School of Medicine at Mount Sinai
Melissa Saperstein, MSW	Icahn School of Medicine at Mount Sinai
Eyal Shemesh, MD	Icahn School of Medicine at Mount Sinai
Andrew Ting, MD	Icahn School of Medicine at Mount Sinai
Adam Vella, MD	Icahn School of Medicine at Mount Sinai
Virginia Walther, MSW	Icahn School of Medicine at Mount Sinai
Jennifer Zeitlin, SD	Icahn School of Medicine at Mount Sinai, French Institute of Health and Medical Research (INSERM)
Marianne McPherson, PhD, MS	NICHQ
Joseph Anarella, MPH	NYS DOH, OHIP
Lee Sanders, MD, MPH	Stanford University
Kasey Coyne	Northwestern
Victoria Wagner	New York State Health Department Office Quality Patient Safety
Laura Soloway	New York State Health Department Office Quality Patient Safety
Kevin Malloy	New York State Health Department Office Quality Patient Safety
Judy Stribling	Icahn School of Medicine at Mount Sinai
Linda Radecki	AAP
Keri Thiessen	AAP – QuIIN
Laura Morris	New York State Department of Health
Olivia Lindly	OHSU
Kasey McCracken	OHSU
Sandeep Sharma	Icahn School of Medicine at Mount Sinai
Ann Nevar	UH Rainbow Babies and Children's Hospital/ CAPQuaM
Suzanne Lo	UH Rainbow Babies and Children's Hospital/ CAPQuaM
Amy Balbierz	Icahn School of Medicine at Mount Sinai
Samantha Raymond	Icahn School of Medicine at Mount Sinai
Allisyn Vachon	Icahn School of Medicine at Mount Sinai
Shannon Weber	Icahn School of Medicine at Mount Sinai
ROLE: Contractors	
Mary Fialkowski, MBA	The Joint Commission
Linda Hanold, MHSA	The Joint Commission
Elvira Ryan, MBA, BSN, RN	The Joint Commission
Sharon Sprenger, RHIA, CPHQ, MPA	The Joint Commission
Ann Watt, MBA, RHIA	The Joint Commission
Amy Slott	The Joint Commission
Denise Krusenoski	The Joint Commission
Elvira Ryan	The Joint Commission
Tasha Mearday	The Joint Commission

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:**

Appendix

Figure 1

FIGURE 1

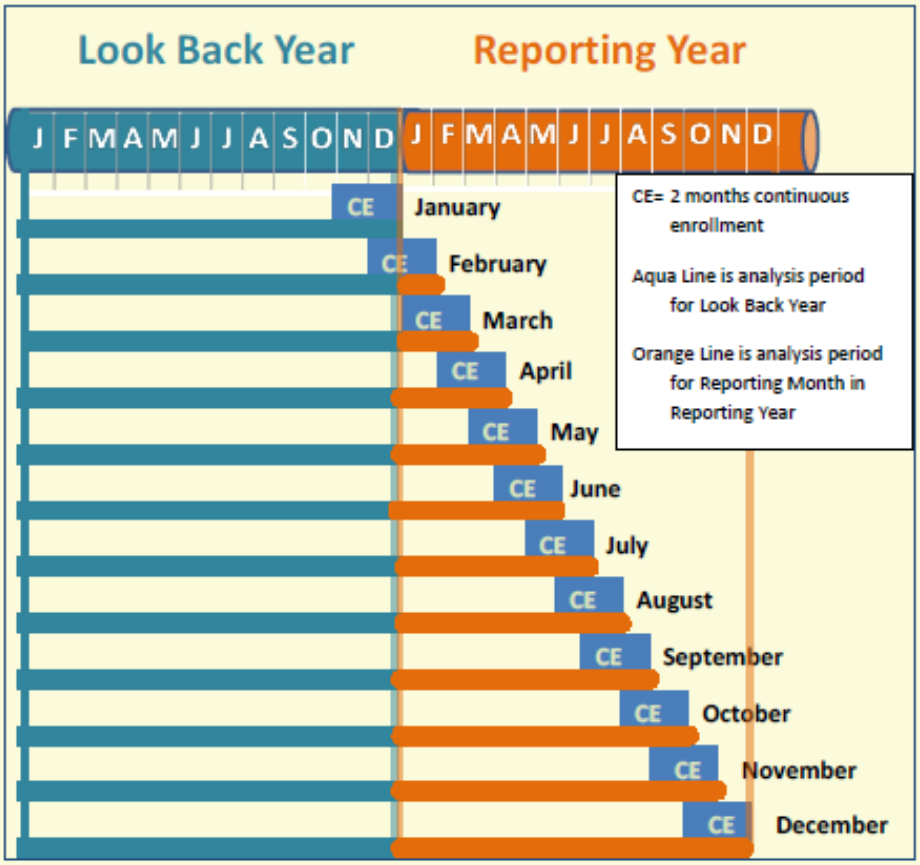


Diagram of Eligibility Requirements for Asthma Measure

Description of Diagram Components

Aqua: Look Back Year
Orange: Reporting Year
Blue: Continuous Enrollment (CE) period must be met to be eligible in a given reporting month
Black: Label of Reporting Month

Description of Eligibility Criteria Assessment

Eligibility Criteria are assessed month by month, using both the entire Look Back Year and the Reporting Year.

For any given reporting month, assess eligibility on 2 criteria. Eligible children are those that meet both of the following:

1) Has the child been continuously enrolled (CE) for the two months prior to the Reporting Month (3 months of continuous enrollment including the current Reporting Month)
2) Does the child meet the criteria for having been managed for persistent asthma in the assessment period (Aqua + Orange Lines)

Figure 2
Roadmap and overview of sample selection and measure calculation

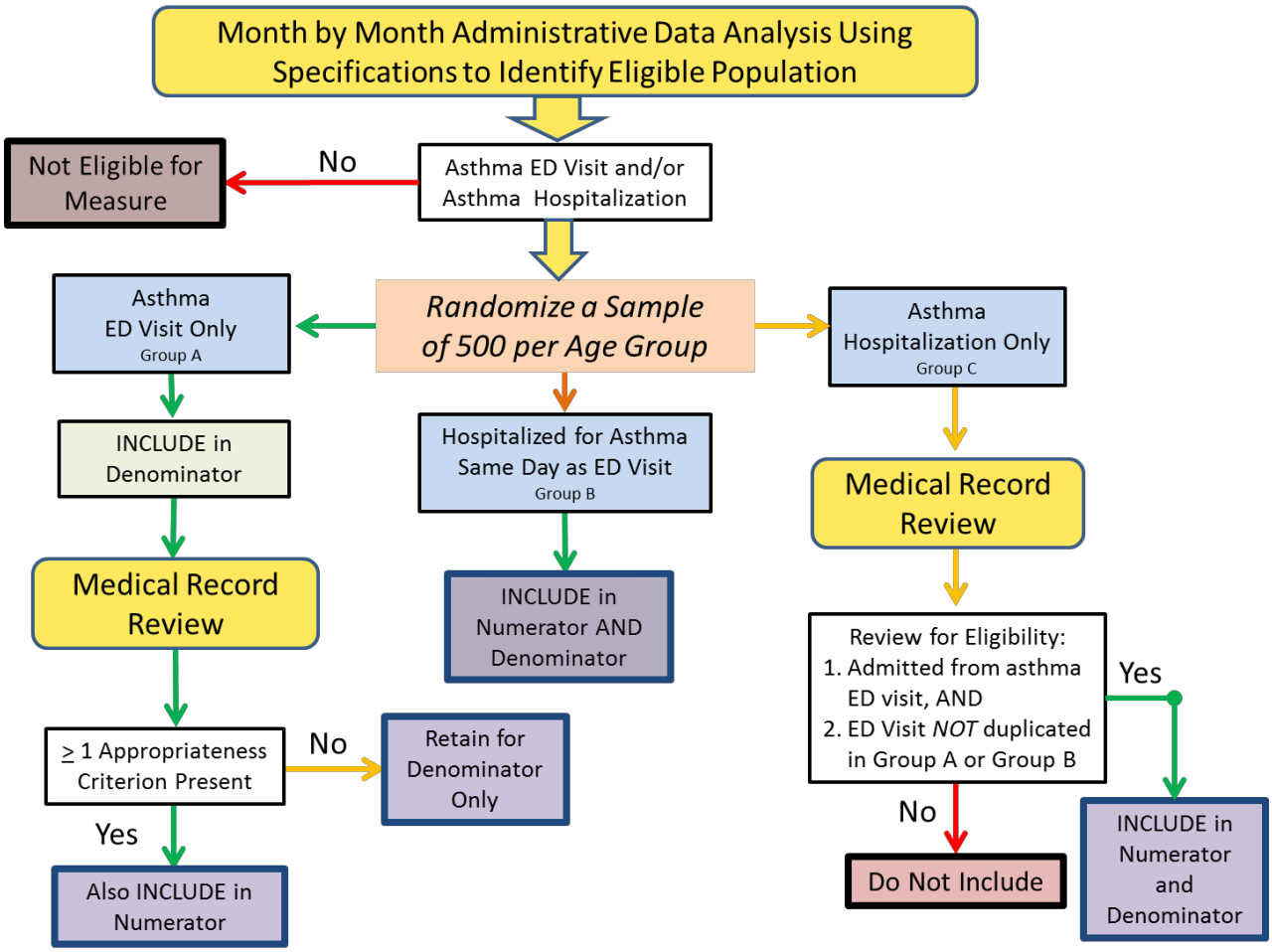


Figure 2 Notes:

- Determining eligibility for sample selection precedes determining eligibility for measure.
- On the basis of the Administrative Data Analysis, children who are potentially eligible for the measure will be identified and segregated into Groups A, B, and C (the blue boxes above).
- Children are eligible for Group B if three things are found in the administrative data: ED Visit; Hospitalization on same day and same institution; and Hospital discharge is after date of ED visit.
- National and NY State data suggest that approximately ¾ of childhood asthma hospitalizations are admitted from ED, that about 1 in 9 childhood asthma ED visits result in hospitalization and that children admitted from the ED may not have their ED visit coded in administrative data.
- Medical record review determines eligibility for numerator among the Group A children, all of whom have already qualified to be included in the denominator.
- Group B children are eligible for both the numerator and the denominator on the basis of administrative data analysis alone and do not require chart review.
- Medical record review determines eligibility for inclusion in the measure (denominator!) for Group C children. If they are eligible for the denominator (i.e. that have been admitted directly from an unduplicated ED visit) then they are also qualified for the numerator.

[i] The *k* value indicates a near perfect agreement (*k*: 0.81-1.0 between coded data and chart review data), and substantial agreement (*k*: 0.61-0.80).

[ii] We note that 1799 and 1800 are not directly applicable because they were tested at the score level. However, the scores were dependent upon definitions which use the same data element level as our measure and thus provide indirect evidence of the capacity of a measure using such data elements to produce valid scores.

Thus we cite them not as specific evidence of our score level performance of the submitted measure, but as evidence that the HEDIS measures that rely on the same administrative data elements for their denominator have the capacity to distinguish signal to noise at a very high level. While the evidence is indirect it is dispositive. That is, we assert that had the data elements been inadequate it would result in non-differential misclassification error which is a major bias towards the null thus introducing noise and reducing signal. That this does not happen to an appreciable degree specifically implies that the data elements function well – indeed this could be one rationale for why NQF allows the use of performance score level analysis in the first place. These findings provide strong indirect evidence of the validity of our approach to capturing the measure's denominator.

There is nearly complete overlap of the denominator codes and there is overlap of the denominator elements. Where codes differ it is specific to decisions made by the CAPQuaM expert panel which was aware of the NCQA measures. Review of the medication lists for 0036 reveal that all medication used by the submitted CAPQuaM measure are also in the HEDIS measure. The CAPQuaM measure excludes specifically short acting beta agonists and leukotriene inhibitors at the specific direction of the CAPQuaM expert panel. We also specify exclude indacaterol from the list of "asthma specific medications" since it is a long acting beta agonist which is only indicated in the USA for treatment of COPD, which is a specific exclusion criterion for this measure.

Further, we identify asthma visits and medications using the same data that an insurance company or Medicaid would use for payment, including ICD codes, CPT codes, and revenue codes. We have had conversations with expert coders and New York State Department of Health Office of Health Insurance Programs to confirm our choices. Our literature review found that while there is moderate agreement (kappa = 0.45 - 0.50) when comparing administrative data regarding the presence of constructs such as recent asthma attacks, use of asthma medications, attack or medication, attack and medication, using 1 year of administrative claims data to parent report, the agreement improves from 0.55 to 0.60 when using two years of data.(1) We expect that these kappas would be significantly higher were the analyses restricted to children with disease that met our construct criteria for identifiable asthma.

The literature further supports our work as highlighted above in the table and in more detail in our testing form 2b2.3 (validity testing).