

NATIONAL QUALITY FORUM

Measure Submission and Evaluation Worksheet 5.0

This form contains the information submitted by measure developers/stewards, organized according to NQF's measure evaluation criteria and process. The evaluation criteria, evaluation guidance documents, and a blank online submission form are available on the [submitting standards web page](#).

NQF #: 0022 NQF Project: Patient Safety Measures-Complications Project
(for Endorsement Maintenance Review) Original Endorsement Date: Aug 10, 2009 Most Recent Endorsement Date: Aug 10, 2009
BRIEF MEASURE INFORMATION
De.1 Measure Title: Use of High Risk Medications in the Elderly
Co.1.1 Measure Steward: National Committee for Quality Assurance
De.2 Brief Description of Measure: a: Percentage of Medicare members 65 years of age and older who received at least one high-risk medication. b: Percentage of Medicare members 65 years of age and older who received at least two different high-risk medications. For both reates, a lower rate represents better performance.
2a1.1 Numerator Statement: a: At least one prescription dispensed for any high-risk medication during the measurement year. b: At least two prescriptions dispensed for different high-risk medications during the measurement year.
2a1.4 Denominator Statement: All patients ages 65 years and older as of December 31 of the measurement year.
2a1.8 Denominator Exclusions: N/A
1.1 Measure Type: Process 2a1. 25-26 Data Source: Administrative claims, Electronic Clinical Data, Electronic Clinical Data : Pharmacy 2a1.33 Level of Analysis: Clinician : Group/Practice, Clinician : Individual, Health Plan, Integrated Delivery System
1.2-1.4 Is this measure paired with another measure? No
De.3 If included in a composite, please identify the composite measure (title and NQF number if endorsed): N/A

STAFF NOTES <i>(issues or questions regarding any criteria)</i>
Comments on Conditions for Consideration:
Is the measure untested? Yes <input type="checkbox"/> No <input type="checkbox"/> If untested, explain how it meets criteria for consideration for time-limited endorsement:
1a. Specific national health goal/priority identified by DHHS or NPP addressed by the measure (check De.5): 5. Similar/related endorsed or submitted measures (check 5.1): Other Criteria:
Staff Reviewer Name(s):

1. IMPACT, OPPORTUNITY, EVIDENCE - IMPORTANCE TO MEASURE AND REPORT
Importance to Measure and Report is a threshold criterion that must be met in order to recommend a measure for endorsement. All three subcriteria must be met to pass this criterion. See guidance on evidence .

Measures must be judged to be important to measure and report in order to be evaluated against the remaining criteria. (evaluation criteria)

1a. High Impact: H M L I

(The measure directly addresses a specific national health goal/priority identified by DHHS or NPP, or some other high impact aspect of healthcare.)

De.4 Subject/Topic Areas (Check all the areas that apply): Prevention

De.5 Cross Cutting Areas (Check all the areas that apply): Care Coordination, Overuse, Palliative Care and End of Life Care, Population Health, Safety, Safety : Complications

1a.1 Demonstrated High Impact Aspect of Healthcare: Affects large numbers, A leading cause of morbidity/mortality, Frequently performed procedure, High resource use, Patient/societal consequences of poor quality, Severity of illness

1a.2 If "Other," please describe:

1a.3 Summary of Evidence of High Impact (Provide epidemiologic or resource use data):

Certain medications are associated with increased risk of harms from drug side-effects and drug toxicity and pose a concern for patient safety. There is clinical consensus that these drugs pose increased risks in the elderly. Studies link prescription drug use by the elderly with adverse drug events that contribute to hospitalization, increased length of hospital stay, increased duration of illness, nursing home placement and falls and fractures that are further associated with physical, functional and social decline in the elderly.

Despite widely-accepted medical consensus that certain drugs increase the risk of harm to the elderly and should generally be avoided, (Fick, 2003) these drugs are still frequently prescribed to the elderly. Studies have found that 21% to almost 37% of elderly patients filled at least one potentially inappropriate prescription and more than 15% filled at least two. (Curtis, 2004; Simon, 2005) A study of elderly managed care patients found that almost 29% receive at least one potentially inappropriate medication (Simon, 2005).

1a.4 Citations for Evidence of High Impact cited in 1a.3: Fick, DM, et al.Updating the Beers criteria for potentially inappropriate medication use in older adults. Arch Intern Med 2003; 163:2716-2724.

Curtis, LH, et al. Inappropriate Prescribing for Elderly Americans in a Large Outpatient Population Arch Intern Med 2004; 164:1621-1625.

Simon, SR, et al. Potentially Inappropriate Medication Use by Elderly Persons in U.S. Health Maintenance Organizations, Journal of the American Geriatrics Society, 2005, Volume 53, Issue 2, 227-232.

1b. Opportunity for Improvement: H M L I

(There is a demonstrated performance gap - variability or overall less than optimal performance)

1b.1 Briefly explain the benefits (improvements in quality) envisioned by use of this measure:

Lowering the use of high-risk medications in the elderly population should decrease morbidity and mortality associated with adverse drug reactions.

1b.2 Summary of Data Demonstrating Performance Gap (Variation or overall less than optimal performance across providers):

[For Maintenance – Descriptive statistics for performance results for this measure - distribution of scores for measured entities by quartile/decile, mean, median, SD, min, max, etc.]

Medicare

One prescription	2009	2008	2007
N	294	278	244
MEAN	23	23.4	23.2
STDEV	8.99	9.03	9.26
STDERR	0.52	0.54	0.59
MIN	3.87	2.48	0.53
MAX	56.2	54.9	58.9

NQF #0022 Use of High Risk Medications in the Elderly

P10	13.3	12.8	12.1
P25	16.6	16.7	16.6
P50	22.2	22.9	22.8
P75	28.6	28.7	29.2
P90	35.2	33.9	34.8

Medicare

At least 2 prescriptions	2009	2008	2007
N	294	278	244
MEAN	5.68	6.04	5.98
STDEV	4.23	4.38	4.45
STDERR	0.25	0.26	0.28
MIN	0.22	0	0
MAX	27	28.5	29
P10	1.6	1.72	1.8
P50	4.67	5.13	5.15
P75	7.69	7.79	8.25
P90	10.7	11	10.8

1b.3 Citations for Data on Performance Gap: [For Maintenance – Description of the data or sample for measure results reported in 1b.2 including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included]
 Section 1b.2 references data from the most recent three years of measurement for this measure. The data in section 1b.2 includes percentiles, mean, min, max, standard deviations and standard errors. There were 816 submissions each for both portions of this measure.

1b.4 Summary of Data on Disparities by Population Group: [For Maintenance –Descriptive statistics for performance results for this measure by population group]

This measure is not stratified to detect disparities. NCQA has participated with IOM and others in attempting to include information on disparities in measure data collection. However, at the present time, this data, at all levels (claims data, paper chart review, and electronic records), is not coded in a standard manner, and is incompletely captured. There are no consistent standards for what entity (physician, group, plan, employer) should capture and report this data. While “requiring” reporting of the data could push the field forward, it has been our position that doing so would create substantial burden with inability to use the data because of its inconsistency. At the present time, we agree with the IOM report that disparities are best considered by the use of zip code analysis which has limited applicability in most reporting situations. At the health plan level, for HEDIS health plan data collection, NCQA does have extensive data related to our use of stratification by insurance status (Medicare, Medicaid and private-commercial) and would strongly recommend this process where the data base supporting the measurement includes this information. However, we believe that the measure specifications should NOT require this since the measure is still useful where the data needed to determine disparities cannot be ascertained from the data available.

1b.5 Citations for Data on Disparities Cited in 1b.4: [For Maintenance – Description of the data or sample for measure results reported in 1b.4 including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included]

N/A

1c. Evidence (Measure focus is a health outcome OR meets the criteria for quantity, quality, consistency of the body of evidence.)
Is the measure focus a health outcome? Yes No If not a health outcome, rate the body of evidence.

Quantity: H M L I **Quality:** H M L I **Consistency:** H M L I

Quantity	Quality	Consistency	Does the measure pass subcriterion 1c?
M-H	M-H	M-H	Yes <input type="checkbox"/>
L	M-H	M	Yes <input type="checkbox"/> IF additional research unlikely to change conclusion that benefits to patients outweigh harms: otherwise No <input type="checkbox"/>

NQF #0022 Use of High Risk Medications in the Elderly

M-H	L	M-H	Yes <input type="checkbox"/> IF potential benefits to patients clearly outweigh potential harms: otherwise No <input type="checkbox"/>
L-M-H	L-M-H	L	No <input type="checkbox"/>
Health outcome – rationale supports relationship to at least one healthcare structure, process, intervention, or service		Does the measure pass subcriterion1c? Yes <input type="checkbox"/> IF rationale supports relationship	
<p>1c.1 Structure-Process-Outcome Relationship (<i>Briefly state the measure focus, e.g., health outcome, intermediate clinical outcome, process, structure; then identify the appropriate links, e.g., structure-process-health outcome; process- health outcome; intermediate clinical outcome-health outcome</i>):</p> <p>This Patient Safety measure addresses medication management to prevent the harms associated with certain medications in the elderly.</p> <p>Panels of experts in pharmacology and geriatrics have compiled lists of medications to avoid prescribing for patients 65 years of age or older. The most commonly used list is the Beers criteria, which was introduced in 1991 to serve researchers evaluating prescribing quality in nursing homes. The Beers criteria were updated in 1997 and again in 2003 to include 48 "potentially inappropriate medications" (PIMs) for which, according to the consensus panel, there are more effective or safer alternatives for older patients. (Rothberg)</p> <p>Reducing prescriptions of high-risk drugs in the elderly also represents an opportunity to reduce the costs associated with the harm from medications (e.g., hospitalizations from drug toxicity) and encourage clinicians to consider safer, alternative medications. Reducing unnecessary prescribing will also help to reduce cost, given that the elderly population represent one third of all prescription drug expenditures in the U.S. but comprises only 13 percent of the population. (Families USA)</p> <p>While expenditures for prescription drugs in the US are disproportionately clustered among those 65 years and older, (Families USA) this population is twice as likely as those below age 65 to experience adverse drug events and is almost seven times as likely to be hospitalized. (Budnitz 2006) Important factors increasing the risk of adverse drug events in the elderly include prescription of drugs that are generally inappropriate for the elderly, interactions between drugs and pre-existing conditions, and interactions between contra-indicated drugs.</p> <p>While some drugs are generally appropriate to prescribe in the elderly, the side-effects commonly associated with these drugs pose an extra risk to elderly people with certain pre-existing conditions. For example, the unsteadiness (ataxia) frequently associated with antidepressants may be a particular danger for elderly patients with a history of falls. Clinical guidelines identify drugs that are generally inappropriate for the elderly, as well as drugs that are inappropriate for elderly populations with specific diagnoses or conditions. (Fick)</p> <p>In 2005, rates of potentially inappropriate medication use in the elderly were as large or larger than in a 1996 national sample, highlighting the need for progress in this area. (Simon)</p> <p>While some adverse drug events are not preventable, studies estimate that between 30% and 80% of adverse drug events in the elderly are preventable. (MacKinnon)</p> <p>Reducing the number of inappropriate prescriptions can lead to improved patient safety and significant cost savings. Conservative estimates of extra costs due to potentially inappropriate medications in the elderly average \$7.2 billion a year. (Fu)</p> <p>Medication use by older adults will likely increase further as the U.S. population ages, new drugs are developed, and new therapeutic and preventive uses for medications are discovered. (Rothberg)</p> <p>By the year 2030, nearly 1 in 5 U.S. residents is expected to be aged 65 years or older; this age group is projected to more than double in number from 38.7 million in 2008 to more than 88.5 million in 2050.^{1,2} Likewise, the population aged 85 years or older is expected to increase almost 4-fold, from 5.4 million to 19 million between 2008 and 2050.¹ As the elderly population continues to grow, the number of older adults who present with multiple medical conditions for which several medications are prescribed continues to increase, resulting in polypharmacy. (Gray)</p>			
<p>1c.2-3 Type of Evidence (<i>Check all that apply</i>):</p>			

Selected individual studies (rather than entire body of evidence), Systematic review of body of evidence (other than within guideline development)

1c.4 Directness of Evidence to the Specified Measure (State the central topic, population, and outcomes addressed in the body of evidence and identify any differences from the measure focus and measure target population):

Evidence base is the original Beers study, the Zahn study and the Fick update to the Beers list in 2003.

1c.5 Quantity of Studies in the Body of Evidence (Total number of studies, not articles): 2

1c.6 Quality of Body of Evidence (Summarize the certainty or confidence in the estimates of benefits and harms to patients across studies in the body of evidence resulting from study factors. Please address: a) study design/flaws; b) directness/indirectness of the evidence to this measure (e.g., interventions, comparisons, outcomes assessed, population included in the evidence); and c) imprecision/wide confidence intervals due to few patients or events): Systematic synthesis of research and expert opinion = Low

1c.7 Consistency of Results across Studies (Summarize the consistency of the magnitude and direction of the effect): The studies consistently mention similar drugs. Since the bodies of evidence all relate to the original Beers list, they maintain consistency in process.

1c.8 Net Benefit (Provide estimates of effect for benefit/outcome; identify harms addressed and estimates of effect; and net benefit - benefit over harms):

Each updated study contributes to the strength of the measure by updating the medication lists.

1c.9 Grading of Strength/Quality of the Body of Evidence. Has the body of evidence been graded? No

1c.10 If body of evidence graded, identify the entity that graded the evidence including balance of representation and any disclosures regarding bias: N/A

1c.11 System Used for Grading the Body of Evidence: Other

1c.12 If other, identify and describe the grading scale with definitions: N/A

1c.13 Grade Assigned to the Body of Evidence: N/A

1c.14 Summary of Controversy/Contradictory Evidence: N/A

1c.15 Citations for Evidence other than Guidelines(Guidelines addressed below):

Families USA, Cost Overdose: Growth in Drug Spending for the Elderly, 1992-2010. (Washington, DC. Families USA), July 2000, p. 2.

Budnitz, D et al. National surveillance of emergency department visits for outpatient adverse drug events. JAMA 2006;296:1858-1866.

Zhan, C, et al. Potentially inappropriate medication use in the community-dwelling elderly. JAMA 2001; 286(22):2823-2868

Beers, M.H. Explicit criteria for determining potentially inappropriate medication use by the elderly. Arch Intern Med 1997; 157:1531-1536.

Fick, DM, et al.Updating the Beers criteria for potentially inappropriate medication use in older adults. Arch Intern Med 2003; 163:2716-2724.

Curtis, LH, et al. Inappropriate Prescribing for Elderly Americans in a Large Outpatient Population Arch Intern Med 2004; 164:1621-1625.

Simon, SR, et al. Potentially Inappropriate Medication Use by Elderly Persons in U.S. Health Maintenance Organizations, 2000-200, Journal of the American Geriatrics Society, 2005, Volume 53, Issue 2, 227-232

Fu AZ, et al. Inappropriate Medication Use and Health Outcomes in the Elderly, Journal of the American Geriatrics Society 2004; Volume 52, Issue 11, 1934-9.

MacKinnon NJ, et al. Indicators of Preventable Drug-related Morbidity in Older Adults: Use Within a Managed Care Organization. J Managed Care Pharm. 2003; 9:134-41.

Kaufman MB, et al. Effect of Prescriber Education on the Use of Medications Contraindicated in Older Adults in a Managed Medicare Population. J Manag Care Pharm 2005 April/May; 11(3):211-219.

Rothberg MB, Perkow PS, Liu F, et al. Potentially inappropriate medication use in hospitalized elders. J Hosp Med. 2008;3:91-102.

Gray, PharmD, Gardner, MD. Adverse Drug Events in the Elderly: An Ongoing Problem, Journal of Managed Care Pharmacy Sep. 2009; Vol. 15, No. 7.

AHRQ, National Quality Measures Clearinghouse. www.qualitymeasures.ahrq.gov (Accessed Web page: October 12, 2009)

1c.16 Quote verbatim, the specific guideline recommendation (Including guideline # and/or page #):
N/A

1c.17 Clinical Practice Guideline Citation: N/A

1c.18 National Guideline Clearinghouse or other URL: N/A

1c.19 Grading of Strength of Guideline Recommendation. Has the recommendation been graded? No

1c.20 If guideline recommendation graded, identify the entity that graded the evidence including balance of representation and any disclosures regarding bias:

1c.21 System Used for Grading the Strength of Guideline Recommendation: Other

1c.22 If other, identify and describe the grading scale with definitions: N/A

1c.23 Grade Assigned to the Recommendation: N/A

1c.24 Rationale for Using this Guideline Over Others: N/A

Based on the NQF descriptions for rating the evidence, what was the developer's assessment of the quantity, quality, and consistency of the body of evidence?

1c.25 Quantity: High 1c.26 Quality: High 1c.27 Consistency: High

Was the threshold criterion, *Importance to Measure and Report*, met?

(1a & 1b must be rated moderate or high and 1c yes) Yes No

Provide rationale based on specific subcriteria:

For a new measure if the Committee votes NO, then STOP.

For a measure undergoing endorsement maintenance, if the Committee votes NO because of 1b. (no opportunity for improvement), it may be considered for continued endorsement and all criteria need to be evaluated.

2. RELIABILITY & 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. (evaluation criteria)

Measure testing must demonstrate adequate reliability and validity in order to be recommended for endorsement. Testing may be conducted for data elements and/or the computed measure score. Testing information and results should be entered in the appropriate field. Supplemental materials may be referenced or attached in item 2.1. See [guidance on measure testing](#).

S.1 Measure Web Page (*In the future, NQF will require measure stewards to provide a URL link to a web page where current detailed specifications can be obtained*). Do you have a web page where current detailed specifications for this measure can be obtained? No

S.2 If yes, provide web page URL:

2a. RELIABILITY. Precise Specifications and Reliability Testing: H M L I

2a1. Precise Measure Specifications. (*The measure specifications precise and unambiguous.*)

2a1.1 Numerator Statement (*Brief, narrative description of the measure focus or what is being measured about the target population, e.g., cases from the target population with the target process, condition, event, or outcome*):

a: At least one prescription dispensed for any high-risk medication during the measurement year.

b: At least two prescriptions dispensed for different high-risk medications during the measurement year.

2a1.2 Numerator Time Window (*The time period in which the target process, condition, event, or outcome is eligible for inclusion*):
The measurement year.

2a1.3 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, codes with descriptors, and/or specific data collection items/responses*):

Antianxiety (includes combination drugs)
aspirin-meprobamate and meprobamate

Antiemetics
scopolamine and trimethobenzamide

Analgesics (includes combination drugs)
acetaminophen-diphenhydramine, diphenhydramine-magnesium salicylate, and ketorolac

Antihistamines (includes combination drugs)
APAP/dextromethorphan/diphenhydramine, APAP/diphenhydramine/phenylephrine, APAP/diphenhydramine/pseudoephedrine, acetaminophen-diphenhydramine, atropine/CPM/hyoscyamine/PE/PPA/scopolamine, carbetapentane/diphenhydramine/phenylephrine, codeine/phenylephrine/promethazine, codeine-promethazine, cyproheptadine, dexchlorpheniramine, dexchlorpheniramine/dextromethorphan/PSE, dexchlorpheniramine/guaifenesin/PSE, dexchlorpheniramine/hydrocodone/phenylephrine, dexchlorpheniramine/methscopolamine/PSE, dexchlorpheniramine-pseudoephedrine, dextromethorphan-promethazine, diphenhydramine, diphenhydramine/hydrocodone/phenylephrine, diphenhydramine-magnesium salicylate, diphenhydramine-phenylephrine, diphenhydramine-pseudoephedrine, hydroxyzine hydrochloride, hydroxyzine pamoate, phenylephrine-promethazine, promethazine, tripeleminamine

Antipsychotic, typical
mesoridazine and thioridazine

Amphetamines
amphetamine-dextroamphetamine, benzphetamine, dexamphetamine, dextroamphetamine, diethylpropion, methamphetamine, methylphenidate, pemoline, phendimetrazine, phentermine

Barbiturates
amobarbital, butabarbital, mephobarbital, pentobarbital, Phenobarbital, and secobarbital

Long-acting benzodiazepines (includes combination drugs)
amitriptyline-chlordiazepoxide, chlordiazepoxide, chlordiazepoxide-clidinium, diazepam, and flurazepam

Calcium channel blockers
nifedipine—short-acting only

Gastrointestinal anti-spasmodics
dicyclomine and propantheline

Belladonna alkaloids (includes combination drugs)
atropine, atropine/CPM/hyoscyamine/PE/scopolamine, atropine/hyoscyamine/PB/scopolamine, atropine-difenoxin, atropine-diphenoxylate, atropine-edrophonium, belladonna, belladonna/caffeine/ergotamine/pentobarbital, belladonna/ergotamine/phenobarbital, butabarbital/hyoscyamine/phenazopyridine, digestive enzymes/hyoscyamine/phenyltoloxamine, hyoscyamine, hyoscyamine/methenam/m-blue/phenyl salicyl, and hyoscyamine-phenobarbital

Skeletal muscle relaxants (includes combination drugs)
ASA/caffeine/orphenadrine, ASA/carisoprodol/codeine, aspirin-carisoprodol, aspirin-meprobamate, aspirin-methocarbamol, carisoprodol, chlorzoxazone, cyclobenzaprine, metaxalone, methocarbamol, orphenadrine

Oral estrogens (includes combination drugs)
conjugated estrogen, conjugated estrogen-medroxyprogesterone, esterified estrogen, esterified estrogen-methyltestosterone, estropipate

Oral hypoglycemic: chlorpropamide

Narcotics (includes combination drugs)
ASA/caffeine/propoxyphene, acetaminophen-pentazocine, acetaminophen-propoxyphene, belladonna-opium, meperidine, meperidine-promethazine, naloxone-pentazocine, pentazocine, propoxyphene hydrochloride, and propoxyphene napsylate

Vasodilators
cyclandelate, dipyridamole—short-acting only, ergot mesyloid, isoxsuprine

Others (including androgens and anabolic steroids, thyroid drugs, urinary anti-infectives)
methyltestosterone, nitrofurantoin, nitrofurantoin macrocrystals, nitrofurantoin macrocrystals-monohydrate, thyroid desiccated

2a1.4 Denominator Statement (Brief, narrative description of the target population being measured):
All patients ages 65 years and older as of December 31 of the measurement year.

2a1.5 Target Population Category (Check all the populations for which the measure is specified and tested if any): **Adult/Elderly Care, Populations at Risk, Special Healthcare Needs**

2a1.6 Denominator Time Window (The time period in which cases are eligible for inclusion):
December 31 of the measurement year

2a1.7 Denominator Details (All information required to identify and calculate the target population/denominator such as definitions, codes with descriptors, and/or specific data collection items/responses):
Use administrative data for eligible population

2a1.8 Denominator Exclusions (Brief narrative description of exclusions from the target population):
N/A

2a1.9 Denominator Exclusion Details (All information required to identify and calculate exclusions from the denominator such as definitions, codes with descriptors, and/or specific data collection items/responses):
N/A

2a1.10 Stratification Details/Variables (All information required to stratify the measure results including the stratification variables, codes with descriptors, definitions, and/or specific data collection items/responses):

N/A

2a1.11 Risk Adjustment Type (Select type. Provide specifications for risk stratification in 2a1.10 and for statistical model in 2a1.13): **No risk adjustment or risk stratification** **2a1.12 If "Other," please describe:**

2a1.13 Statistical Risk Model and Variables (Name the statistical method - e.g., logistic regression and list all the risk factor variables. Note - risk model development should be addressed in 2b4.):

N/A

2a1.14-16 Detailed Risk Model Available at Web page URL (or attachment). Include coefficients, equations, codes with descriptors, definitions, and/or specific data collection items/responses. Attach documents only if they are not available on a webpage and keep attached file to 5 MB or less. NQF strongly prefers you make documents available at a Web page URL. Please supply login/password if needed:

2a1.17-18. Type of Score: [Rate/proportion](#)

2a1.19 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](#)

2a1.20 Calculation Algorithm/Measure Logic(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; aggregating data; risk adjustment; etc.):

[Step 1. Determine the eligible population. The eligible population is all members who satisfy all specified criteria, including any age, continuous enrollment, benefit, event, or anchor date enrollment requirement.](#)

[Step 2. Search administrative systems to identify numerator events for all members in the eligible population.](#)

[Step 3. If applicable, for members for whom administrative data do not show a positive numerator event, search administrative data for an exclusion to the service/procedure being measured. Note: This step applies only to measures for which optional exclusions are specified and for which the organization has chosen to search for exclusions. The organization is not required to search for optional exclusions.](#)

[Step 4. Exclude from the eligible population members from step 3 for whom administrative system data identified an exclusion to the service/procedure being measured.](#)

[Step 5. Calculate the rate.](#)

2a1.21-23 Calculation Algorithm/Measure Logic Diagram URL or attachment:

2a1.24 Sampling (Survey) Methodology. If measure is based on a sample (or survey), provide instructions for obtaining the sample, conducting the survey and guidance on minimum sample size (response rate):

N/A

2a1.25 Data Source (Check all the sources for which the measure is specified and tested). If other, please describe:

[Administrative claims, Electronic Clinical Data, Electronic Clinical Data : Pharmacy](#)

2a1.26 Data Source/Data Collection Instrument (Identify the specific data source/data collection instrument, e.g. name of database, clinical registry, collection instrument, etc.): [Healthcare Effectiveness Data and Information Set \(HEDIS\)](#)

2a1.27-29 Data Source/data Collection Instrument Reference Web Page URL or Attachment:

2a1.30-32 Data Dictionary/Code Table Web Page URL or Attachment:

2a1.33 Level of Analysis (Check the levels of analysis for which the measure is specified and tested): Clinician : Group/Practice, Clinician : Individual, Health Plan, Integrated Delivery System

2a1.34-35 Care Setting (Check all the settings for which the measure is specified and tested): Ambulatory Care : Clinician Office, Pharmacy

2a2. Reliability Testing. (Reliability testing was conducted with appropriate method, scope, and adequate demonstration of reliability.)

2a2.1 Data/Sample (Description of the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included):
HEDIS Health Plan performance data from 2010

2a2.2 Analytic Method (Describe method of reliability testing & rationale):

Reliability was estimated by 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.

2a2.3 Testing Results (Reliability statistics, assessment of adequacy in the context of norms for the test conducted):

Reliability for this measure was calculated as 0.99882 for one prescription and 0.99819 for two or more prescriptions.

2b. VALIDITY. Validity, Testing, including all Threats to Validity: H M L I

2b1.1 Describe how the measure specifications (measure focus, target population, and exclusions) **are consistent with the evidence cited in support of the measure focus** (criterion 1c) **and identify any differences from the evidence:**

The measure focuses on reducing risk of adverse drug events in the elderly population. The evidence is consistent with the focus and scope of this measure.

2b2. Validity Testing. (Validity testing was conducted with appropriate method, scope, and adequate demonstration of validity.)

2b2.1 Data/Sample (Description of the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included):

The measure aligns with current evidence.

2b2.2 Analytic Method (Describe method of validity testing and rationale; if face validity, describe systematic assessment):

NCQA tested the measure for face validity using a panel of stakeholders with specific expertise in measurement and women’s health. This panel included representatives from key stakeholder groups geriatricians, health plans, Medicare officials and researchers. Experts reviewed the results of the field test and assessed whether the results were consistent with expectations, whether the measure represented quality care, and whether we were measuring the most important aspect of care in this area.

2b2.3 Testing Results (Statistical results, assessment of adequacy in the context of norms for the test conducted; if face validity, describe results of systematic assessment):

This measure was deemed valid by the expert panel.

POTENTIAL THREATS TO VALIDITY. (All potential threats to validity were appropriately tested with adequate results.)

2b3. Measure Exclusions. (Exclusions were supported by the clinical evidence in 1c or appropriately tested with results demonstrating the need to specify them.)

2b3.1 Data/Sample for analysis of exclusions (Description of the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included):

NCQA currently allows health plans for optional exclusion to their results. NCQA does not conduct the annual analysis applied to a sample. In measure development, field testing and any re-analysis for update, we investigate and validate the effect reliability exclusion applied to the eligible denominator.

2b3.2 Analytic Method (Describe type of analysis and rationale for examining exclusions, including exclusion related to patient preference):

N/A

2b3.3 Results (Provide statistical results for analysis of exclusions, e.g., frequency, variability, sensitivity analyses):

N/A

2b4. Risk Adjustment Strategy. (For outcome measures, adjustment for differences in case mix (severity) across measured entities was appropriately tested with adequate results.)

2b4.1 Data/Sample (Description of the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included):

N/A

2b4.2 Analytic Method (Describe methods and rationale for development and testing of risk model or risk stratification including selection of factors/variables):

N/A

2b4.3 Testing Results (Statistical risk model: Provide quantitative assessment of relative contribution of model risk factors; risk model performance metrics including cross-validation discrimination and calibration statistics, calibration curve and risk decile plot, and assessment of adequacy in the context of norms for risk models. Risk stratification: Provide quantitative assessment of relationship of risk factors to the outcome and differences in outcomes among the strata):

N/A

2b4.4 If outcome or resource use measure is not risk adjusted, provide rationale and analyses to justify lack of adjustment: The measure assesses the use of high-risk medication in a general elderly population; risk adjustment is not indicated.

2b5. Identification of Meaningful Differences in Performance. (The performance measure scores were appropriately analyzed and discriminated meaningful differences in quality.)

2b5.1 Data/Sample (Describe the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included):

Data analysis demonstrates that methods for scoring and analysis of the specified measure allow for identification of statistically significant and practically/clinically meaningful differences in performance.

2b5.2 Analytic Method (Describe methods and rationale to identify statistically significant and practically/meaningfully differences in performance):

Comparison of means and percentiles; analysis of variance against established benchmarks: if sample size is >400, we would use an analysis of variance.

NQF #0022 Use of High Risk Medications in the Elderly

2b5.3 Results (Provide measure performance results/scores, e.g., distribution by quartile, mean, median, SD, etc.; identification of statistically significant and meaningful differences in performance):

One prescription	2009	2008	2007
N	294	278	244
MEAN	23	23.4	23.2
STDEV	8.99	9.03	9.26
STDERR	0.52	0.54	0.59
MIN	3.87	2.48	0.53
MAX	56.2	54.9	58.9
P10	13.3	12.8	12.1
P25	16.6	16.7	16.6
P50	22.2	22.9	22.8
P75	28.6	28.7	29.2
P90	35.2	33.9	34.8

At least 2 prescriptions	2009	2008	2007
N	294	278	244
MEAN	5.68	6.04	5.98
STDEV	4.23	4.38	4.45
STDERR	0.25	0.26	0.28
MIN	0.22	0	0
MAX	27	28.5	29
P10	1.6	1.72	1.8
P50	4.67	5.13	5.15
P75	7.69	7.79	8.25
P90	10.7	11	10.8

2b6. Comparability of Multiple Data Sources/Methods. (If specified for more than one data source, the various approaches result in comparable scores.)

2b6.1 Data/Sample (Describe the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included):

For the field test, NCQA required participating plans to provide data beyond what would normally be necessary to compute these measures. For purposes of the field test, the measurement year was 2002 and 2003. For each measure, the participating plans were asked to provide patient enrollment data and pharmacy data from administrative data systems for the entire measure eligible population.

NOTE: At the time of field testing, the measure was called "Drugs to be avoided in the elderly: a. Patients who receive at least one drug to be avoided, b. Patients who receive at least two different drugs to be avoided."

2b6.2 Analytic Method (Describe methods and rationale for testing comparability of scores produced by the different data sources specified in the measure):

The purpose of field testing is to determine:

- The validity of the administrative algorithm to identify the target population (denominator) based upon the measurement period, continuous enrollment /exclusionary criteria
- The validity of administrative data to accurately capture medical processes delivered (i.e. tests) or diagnoses by comparing administrative results with data from a sample of medical records
- The feasibility of the measure specifications to identify the quality problem and to discriminate performance between health plans for the purposes of HEDIS public reporting.
- The reliability and feasibility of the measure specifications so that all health plans can capture the required data elements and can conduct programming

Based upon the field test results, NCQA made necessary revisions to the measure specifications so that it meets the Desirable Attributes of a HEDIS measure.

2b6.3 Testing Results (Provide statistical results, e.g., correlation statistics, comparison of rankings; assessment of adequacy in the context of norms for the test conducted):

Percentage of members with at least one prescription:

Overall prescribing rates are high (92.6%) among elderly members (continuously enrolled for 12 months) with at least one prescription for any drug to be avoided based on Zahn or Beers list. However, this number may be inflated because the Beers list was more comprehensive and includes drugs considered to be “low severity” for potential patient safety, such as estrogen.

Drug “risk” categories:

On average, 45.2% of enrolled elderly are prescribed at least 1 of the drugs classified by Zahn in one of the 3 high-risk drug categories (plan range 18.7% - 86.7%).

Prescribing rates for drugs classified by Zahn into three high-risk drug categories are:

- About 7.7% of elderly enrollees get at least 1 “never appropriate” drug (plan range 2.4% to 9.1%)
- 27.5% get at least 1 “rarely appropriate” drugs (plan range 10.8% - 44.1%)
- 10% get at least 1 “sometimes indicated” drug (plan range 5.5% - 33.5%).

A more meaningful rate may be to show that on average 35.2% of members are prescribed at least 1 of the drugs from the “never appropriate” or “rarely appropriate” Zahn’s high risk drug categories, (plan range 13.2% - 53.2%).

Another clinically meaningful quality indicator is to look at the percentage of members who received at least 2 prescriptions of different therapeutic classes for drugs to be avoided in the elderly. This represents a subset of members who are at increased risk of adverse drug events and patient safety from additional receipt of harmful drugs. Plan performance on this rate was 6%, plan range 1.1% - 9.3%. Women were more likely than men to receive 2 or more drugs (6.9% vs 4.2%), and older elderly patients ages 85 and older were less likely than younger elderly patients ages 65-74 years (4.2% vs 6.1%). These show specific areas for plans to target improvement.

When the Beers “high and low” severity drug risk categories are used, the extent of the quality problem appears to be much worse, although this classification includes a much broader group of drugs:

- Nearly three quarters (72.8%) of elderly enrollees get at least 1 prescription for a drug considered “high severity” and
- Over half (52.2%) of elderly enrollees get at least 1 prescription for a “low severity” drug.
- Less than 1% of elderly enrollees received a drug, Phenobarbital that is not classified by Beers, but considered “never appropriate” by Zahn. Note: Phenobarbital is also targeted by NCQA for a measure which requires annual drug level monitoring due to potential harms from drug toxicity.

On average, 74% of members are prescribed at least one of the drugs that are on the Beers list but not classified by Zahn (plan range 24.9% - 86.7%).

Number of prescriptions per member per year

NCQA calculated the average number of prescriptions for drugs to be avoided in elderly members per member per year, which included members who were not continuously enrolled for a full year. This method of reporting the information shows a wider range of performance between plans. This may be because some members receive multiple prescriptions for these drugs over the course of their membership. These figures are corroborated by the mean number of prescriptions for members who are enrolled for 12 months.

Based on the data, in 2002, each member could be receiving between 5 and 11 of the drugs on the combined Zahn and Beers list. This PMPY rate is calculated based on the number of drugs received by a member and the number of months he/she is enrolled and is reported as a number of prescriptions per year. This accounts for the time of potential exposure to the “risk” of receiving the drug while enrolled at the health plan. Are these actual calculations...may want to put an transition sentence.

Using the Beers categories (high/low severity), on average a member receives:

- 3 - 6 “high severity” drug prescriptions per year
- 2 - 4.5 “low severity” drug prescriptions per year.

Using the Zhan category, on average a member receives:

NQF #0022 Use of High Risk Medications in the Elderly

- 0.9 – 3.6 prescriptions per year of any drugs in the 3 Zhan risk categories
- 0.1 - 0.3 prescriptions per year for “never appropriate” drugs
- 0.6 - 1.6 prescriptions per year for “rarely appropriate” drugs
- 0.5 - 1.7 prescriptions per year for “sometimes indicated” drugs,
- 3.8 - 7.0 prescriptions per year for drugs not classified by Zahn but which are “high” or “low” severity on the Beers list.
- 0.7 – 2.9 prescriptions per year for “never” or “rarely” appropriate drugs

2c. Disparities in Care: H M L I NA (If applicable, the measure specifications allow identification of disparities.)

2c.1 If measure is stratified for disparities, provide stratified results (Scores by stratified categories/cohorts): The measure is not stratified to detect disparities. NCQA has participated with IOM and others in attempting to include information on disparities in measure data collection. However, at the present time, this data, at all levels (claims data, paper chart review, and electronic records), is not coded in a standard manner, and is incompletely captured. There are no consistent standards for what entity (physician, group, plan, employer) should capture and report this data. While “requiring” reporting of the data could push the field forward, it has been our position that doing so would create substantial burden with inability to use the data because of its inconsistency. At the present time, we agree with the IOM report that disparities are best considered by the use of zip code analysis which has limited applicability in most reporting situations. At the health plan level, for HEDIS health plan data collection, NCQA does have extensive data related to our use of stratification by insurance status (Medicare, Medicaid and private-commercial) and would strongly recommend this process where the data base supporting the measurement includes this information. However, we believe that the measure specifications should NOT require this since the measure is still useful where the data needed to determine disparities cannot be ascertained from the data available.

2c.2 If disparities have been reported/identified (e.g., in 1b), but measure is not specified to detect disparities, please explain:
N/A

2.1-2.3 Supplemental Testing Methodology Information:

Steering Committee: Overall, was the criterion, *Scientific Acceptability of Measure Properties*, met?
(Reliability and Validity must be rated moderate or high) Yes No
Provide rationale based on specific subcriteria:

If the Committee votes No, STOP

3. USABILITY

Extent to which intended audiences (e.g., consumers, purchasers, providers, policy makers) can understand the results of the measure and are likely to find them useful for decision making. (**evaluation criteria**)

C.1 Intended Purpose/ Use (Check all the purposes and/or uses for which the measure is intended): Public Reporting, Quality Improvement with Benchmarking (external benchmarking to multiple organizations)

3.1 Current Use (Check all that apply; for any that are checked, provide the specific program information in the following questions): Public Reporting, Regulatory and Accreditation Programs, Quality Improvement with Benchmarking (external benchmarking to multiple organizations)

3a. Usefulness for Public Reporting: H M L I
(The measure is meaningful, understandable and useful for public reporting.)

3a.1. Use in Public Reporting - disclosure of performance results to the public at large (If used in a public reporting program, provide name of program(s), locations, Web page URL(s)). If not publicly reported in a national or community program, state the reason AND plans to achieve public reporting, potential reporting programs or commitments, and timeline, e.g., within 3 years of endorsement: **[For Maintenance –** If not publicly reported, describe progress made toward achieving disclosure of performance results to the public at large and expected date for public reporting; provide rationale why continued endorsement should be

considered.]

This measure is used in public reporting for plans only through Healthcare Effectiveness Data and Information Set (HEDIS) and is reported through venues such as the annual State of Healthcare Quality report, Quality Compass, America's Best Health Plans.

3a.2. Provide a rationale for why the measure performance results are meaningful, understandable, and useful for public reporting. If usefulness was demonstrated (e.g., focus group, cognitive testing), describe the data, method, and results: [HEDIS measures adhere to the desirable attributes of scientific acceptability, feasibility and usability. The measures provide performance rates that are audited for consistency and accuracy.](#)

3.2 Use for other Accountability Functions (payment, certification, accreditation). If used in a public accountability program, provide name of program(s), locations, Web page URL(s): [It is used in NCQA's Health Plan Accreditation program.](#)

3b. Usefulness for Quality Improvement: H M L I

(The measure is meaningful, understandable and useful for quality improvement.)

3b.1. Use in QI. If used in quality improvement program, provide name of program(s), locations, Web page URL(s):

[For Maintenance – *If not used for QI, indicate the reasons and describe progress toward using performance results for improvement].*

[This measure is a measure in the Healthcare Effectiveness Data and Information Set \(HEDIS\) and is used in NCQA's Health Plan Accreditation program.](#)

3b.2. Provide rationale for why the measure performance results are meaningful, understandable, and useful for quality improvement. If usefulness was demonstrated (e.g., QI initiative), describe the data, method and results:

[Upon review of public comment results, the Committee on Performance Measurement approved the NCQA staff recommendation to add the measure to HEDIS. After reviewing first-year analysis results, the CPM approved the staff recommendation to publicly report the measure. The measure was deemed usable and feasible.](#)

Overall, to what extent was the criterion, Usability, met? H M L I

Provide rationale based on specific subcriteria:

4. FEASIBILITY

Extent to which the required data are readily available, retrievable without undue burden, and can be implemented for performance measurement. (**evaluation criteria**)

4a. Data Generated as a Byproduct of Care Processes: H M L I

4a.1-2 How are the data elements needed to compute measure scores generated? *(Check all that apply).*

Data used in the measure are:

[generated by and used by healthcare personnel during the provision of care, e.g., blood pressure, lab value, medical condition, Coded by someone other than person obtaining original information \(e.g., DRG, ICD-9 codes on claims\)](#)

4b. Electronic Sources: H M L I

4b.1 Are the data elements needed for the measure as specified available electronically *(Elements that are needed to compute measure scores are in defined, computer-readable fields):* [ALL data elements are in a combination of electronic sources](#)

4b.2 If ALL data elements are not from electronic sources, specify a credible, near-term path to electronic capture, OR provide a rationale for using other than electronic sources:

4c. Susceptibility to Inaccuracies, Errors, or Unintended Consequences: H M L I

4c.1 Identify susceptibility to inaccuracies, errors, or unintended consequences of the measurement identified during testing and/or operational use and strategies to prevent, minimize, or detect. If audited, provide results:

[All measures that are used in NCQA programs are audited.](#)

4d. Data Collection Strategy/Implementation: H M L I

A.2 Please check if either of the following apply (*regarding proprietary measures*): [Proprietary measure](#)
4d.1 Describe what you have learned/modified 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 (*e.g., fees for use of proprietary measures*):
[NCQA's multi-stakeholder advisory panels examined an analysis of the measure after its first year of reporting. The measure was deemed appropriate for public reporting. NCQA has processes to ensure coding and specifications are clear and updated when needed.](#)

Overall, to what extent was the criterion, *Feasibility*, met? H M L I
 Provide rationale based on specific subcriteria:

OVERALL SUITABILITY FOR ENDORSEMENT

Does the measure meet all the NQF criteria for endorsement? Yes No

Rationale:

If the Committee votes No, STOP.

If the Committee votes Yes, the final recommendation is contingent on comparison to related and competing measures.

5. COMPARISON TO RELATED AND 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 before a final recommendation is made.

5.1 If there are related measures (either same measure focus or target population) or competing measures (both the same measure focus and same target population), list the NQF # and title of all related and/or competing measures:

5a. Harmonization

5a.1 If this measure has EITHER the same measure focus OR the same target population as [NQF-endorsed measure\(s\)](#): Are the measure specifications completely harmonized?

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

5b. Competing Measure(s)

5b.1 If this measure has 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):

CONTACT INFORMATION

Co.1 Measure Steward (Intellectual Property Owner): [National Committee for Quality Assurance, 1100 13th Street NW, Suite 1000, Washington, District Of Columbia, 20005](#)

Co.2 Point of Contact: [Bob, Rehm, Assistant Vice President, Performance Measurement, Rehm@ncqa.org, 202-955-1728-](#)

Co.3 Measure Developer if different from Measure Steward: [National Committee for Quality Assurance, 1100 13th Street NW, Suite 1000, Washington, District Of Columbia, 20005](#)

Co.4 Point of Contact: [Bob, Rehm, Assistant Vice President, Performance Measurement, Rehm@ncqa.org, 202-955-1728-](#)

Co.5 Submitter: Dawn, Alayon, MPH, CPH, Senior Health Care Analyst, alayon@ncqa.org, 202-955-3533-, National Committee for Quality Assurance

Co.6 Additional organizations that sponsored/participated in measure development:

Co.7 Public Contact: Bob, Rehm, Assistant Vice President, Performance Measurement, Rehm@ncqa.org, 202-955-1728-, National Committee for Quality Assurance

ADDITIONAL INFORMATION

Workgroup/Expert Panel involved in measure development

Ad.1 Provide a list of sponsoring organizations and workgroup/panel members' names and organizations. Describe the members' role in measure development.

Geriatric Measurement Advisory Panel

GMAP Members

Wade Aubry, BCBS Association

Arlene Bierman, University of Toronto and St. Michael's Hospital

Joyce Dubow, AARP

Peter Hollmann, BCBS of Rhode Island

Jerry Johnson, University of Pennsylvania

David Martin, Ovations

Steven Phillips, Sierra Health Services, Inc.

Scott Sarran, BCBS of Illinois

Eric G Tangalos, Mayo Clinic

Joan Weiss, Health Resources and Services Administration

Neil Wenger, UCLA Division of General Internal Medicine and RAND

CMS/AHRQ Liaisons

Marsha Davenport

Jeffrey Kelman

Elizabeth Goldstein

Morgot Blige Holloway

Rosemary Lee

Alice Lee Martin

Chris Haffer

Sonya Bowen

Mary B. Barton

Ad.2 If adapted, provide title of original measure, NQF # if endorsed, and measure steward. Briefly describe the reasons for adapting the original measure and any work with the original measure steward:

Measure Developer/Steward Updates and Ongoing Maintenance

Ad.3 Year the measure was first released: 2006

Ad.4 Month and Year of most recent revision: 05, 2010

Ad.5 What is your frequency for review/update of this measure? Approximately every 3 years, sooner if the clinical guidelines have changed significantly.

Ad.6 When is the next scheduled review/update for this measure?

Ad.7 Copyright statement: © 2011 by the National Committee for Quality Assurance

1100 13th Street, NW, Suite 1000

Washington, DC 20005

Ad.8 Disclaimers:

Ad.9 Additional Information/Comments:

Date of Submission (MM/DD/YY): 09/14/2011