Development and Implementation of Quality Rating System (QRS) Measures for Qualified Health Plans (QHPs)

Draft Measure Testing Summary Report:
Drug Testing for Individuals on Chronic Opioid Therapy (COT)
Deliverable #27

Contract Number: HHSM-500-2013-13009I, Task Order: HHSM-500-T0001

MIDS Task Order: 151095

June 15, 2018

Submitted to:

Contract Officer Representative:
Sophia Chan
Centers for Medicare & Medicaid Services
CCSQ/QMVIG/DPMS
7500 Security Boulevard, Mail Stop: S3-02-01
Baltimore, MD 21244-1850
sophia.chan@cms.hhs.gov

Submitted by:

Project Director:
Jensen Chiu, MHA, SCPM
IMPAQ International, LLC
10420 Little Patuxent Parkway
Suite 300
Columbia, MD 21044
www.impaqint.com

EXHIBIT OF CONTENTS

INTRODUCTION

This project, titled Developing and Implementing Quality Rating System Measures for Qualified Health Plans (Task Order: HHSM-500-T0001), is performed under the Measure & Instrument Development and Support (MIDS) Indefinite Delivery Indefinite Quantity contract for the Centers for Medicare & Medicaid Services (CMS).

The purpose of the project is to inform efforts by CMS to deliver meaningful information to health insurance consumers selecting among an array of Qualified Health Plans (QHPs) in the Health Insurance Exchanges, also known as Health Insurance Marketplaces (hereafter referred to as Exchanges). The overall project goal is to 1) identify measurement gaps in the existing Quality Rating System (QRS) for QHPs, and 2) develop four to six new measures to add to the QRS that will enhance the information available to consumers.

This draft measure testing summary report provides testing results for the *de novo* process measure *Drug Testing for Individuals on Chronic Opioid Therapy (COT)*.

The objective of testing the *de novo* measure is to ensure that the measure is feasible and scientifically acceptable for the QRS. Prior to testing and in alignment with the CMS Measures Management System Blueprint,[1] the Measure Developer:

- Conducted an environmental scan to understand the breadth of existing measures for possible consideration in the QRS.
- Identified gaps in quality measurement in the QRS.
- Determined which gaps necessitated *de novo* measure development.
- Gathered input from a 12-member Technical Expert Panel (TEP) on which measure topics to prioritize for *de novo* development, one of which included *Drug Testing for Individuals on Chronic Opioid Therapy (COT)*. See *Technical Expert Panel and Subject Matter Experts,* Appendix A.
- Developed a business case for *Drug Testing for Individuals on Chronic Opioid Therapy (COT)* based on clinical practice guidelines, available research, and subject matter expert input.
- Constructed the approach to test the measure for feasibility and scientific acceptability.

In addition to providing testing results, this report provides a description of the measure development process and the final measure specifications for *Drug Testing for Individuals on Chronic Opioid Therapy (COT)*. This report also describes the results of testing to determine whether the measure adheres to scientific standards for quality measurement (i.e., measure is feasible, reliable, valid, and the measure identifies variation and room for improvement among QHP products).

SUMMARY OF MEASURE JUSTIFICATION

Importance

The measure *Drug Testing for Individuals on Chronic Opioid Therapy (COT)* focuses on individuals on COT who have not received a drug test at least once in the measurement year. The results of drug tests are important sources of information for providers of patients receiving COT. Routine drug screenings can inform providers of aberrant drug-related behaviors, which can then influence referrals for substance use disorder. Such tests can inform providers when more patient education is warranted to prevent potential drug-drug interactions if undisclosed drugs are discovered through test results. Additionally, drug tests can inform the provider if prescribed opioids are not evidenced by test results, indicating potential diversion or the need to change the treatment regimen to optimize patient outcomes when opioids are not used.

The importance of drug testing for patients on COT is supported by five evidence-based clinical practice guidelines that recommend drug testing at the initiation of COT and periodically thereafter. Clinical practice guidelines from the American Association for Clinical Chemistry (AACC) Academy, American Society of Interventional Pain Physicians, US Department of Veterans Affairs, Centers for Disease Control and Prevention, and American Pain Society and the American Academy of Pain Medicine support the need to manage COT through routine drug screening. More information on each of the clinical practice guidelines is provided in *Clinical Practice Guideline Recommendations,* Appendix B. Additionally, feedback on the measure from 16 consumers enrolled in QHPs and five former Exchange navigators supports the importance of drug testing during COT. During interviews conducted between May and August of 2016, 100% of those interviewed (16 members of QHPs and five former Exchange navigators) rated an early version of the measure as important.

Evidence suggests a link between the measure focus and the outcome of interest (i.e., improving patient care in the management of COT). Specifically, a recent study found that high levels of drug tests for COT patients were associated with lower risk of suicide and drug overdose.[2] Another study suggests that the measure focus is actionable. Providers planned to change the treatment plan in 69% of 83 cases in which a patient tested positive for aberrant behavior determined through the use of drug testing.[3] In 52% of cases with planned treatment changes, the documented change was to alter patients' opioid prescriptions. A potential unintended consequence of drug tests is related to patients not returning for follow-up care.[4] Patients engaging in aberrant drug-related behaviors may not return to their provider if they know that they may be drug-tested and the results could affect their access to opioids from that provider. However, given the importance of optimizing the management of patients on COT, the benefits of drug testing outweigh this potential unintended consequence.

Despite the benefits of drug testing patients on COT, drug testing rates are suboptimal and indicate a gap in care. As evidenced by testing conducted during the development of this measure, as many as 69% of patients on COT do not receive a drug test over the course of a year. As suggested by the literature, testing rates were as low as 8% over a four-year study period.[5] Additionally, variations in rates of drug testing for patients on COT exist between similar sites of

care (i.e., three safety-net primary care clinics had drug test rates of 31%, 35%, and 67%),[6] which suggests there is room for improvement in rates of drug testing patients on COT.

As a process of care that is important for patient safety, drug testing is feasible for managing patients on COT and is actionable by health plans and providers. Actionable treatment changes include changing opioid prescribing, instituting more frequent drug tests, and making referrals for evaluation and treatment by behavioral health specialists.[3,7] Treatment changes are associated with improved patient outcomes such as participation in addiction treatment, cessation of the COT prescription when aberrant drug-related behaviors are present, and improved compliance with the treatment plan.[4,8] A study found that providing primary care physicians with additional resources and education resulted in more patients receiving at least one urine drug test at 1-year post-intervention compared to usual care (74.6% versus 57.9%).[9] Among a federally qualified health center and primary care clinics, implementation of a chronic pain protocol and the development of electronic reports to track provider adherence to the protocol led to an 18.3% increase in the number of patients who had a urine drug screen over a 12-month period.[10]

Impact

In the United States, over 11 million adults misuse opioids and 2.4 million have an opioid use disorder.[11] For patients on COT, research indicates that as many as one in five will develop an opioid use disorder.[12] Additionally, opioids are among the five most common drug classes implicated in ADEs resulting in emergency department and hospital utilization.[13] In 2015, 63% of fatal drug overdoses were attributed to opioids and the overall fatality rate due to opioids was 10.3 deaths per 100,000.[14] Monitoring rates of drug testing is one of many tools that health plans and providers can use to improve the management of patients on COT. It will assist providers in identifying patients who need behavioral health services while safely maintaining patients who need COT and adhere to the plan of care.

The measure addresses the White House priority of combating the national opioid crisis,[15-17] and aligns with the CMS Meaningful Measures priority area of Prevention and Treatment of Opioid and Substance Use Disorders.[18] Furthermore, the measure addresses and supports a number of federal initiatives to reduce ADEs, including the National Drug Control Strategy,[19] the CMS Opioid Misuse Strategy,[20] the US Department of Health and Human Services' 2014 National Action Plan for Adverse Drug Event Prevention (ADE Action Plan),[21] and the 2016 update to the ADE Action Plan.[22]

The measure is anticipated to address a gap in the quality of care related to the management of patients on COT. It will assist QHPs in identifying providers who objectively are not following evidence-based clinical practice guidelines that enhance safety for patients on COT. This process will help QHPs and providers to identify patients on COT who engage in aberrant drug-related behaviors and can help identify patients who need referral for opioid use disorder. Ultimately, the measure should lower the risk of ADEs, including substance abuse and drug-related mortality, for patients on COT.

METHODS

This section of the report describes the data and the approach to developing and operationalizing the measure specifications and scoring methodology. This section also includes the approach to assessing the reliability and validity of the measure and determining whether disparities exist between different subpopulations of patients.

Data

Data from four issuers, representing seven QHP products in 2015 and eight products in 2016, were used to test the measure. In this report, the term "issuer" refers to an individual insurance company or insurance organization. The term "product" refers to a package of health coverage benefits that are offered using a particular network type (i.e., health maintenance organization, preferred provider organization, exclusive provider organization, point of service, or indemnity).[23] Unique products for each issuer are referred to using alphabetic labeling (e.g., two unique products from the same issuer are referred to as Product A and Product B).

Patient-level data representing the target population—members enrolled in Affordable Care Act (ACA) Health Insurance Exchange QHP products—were provided to the Measure Developer from one issuer, henceforth Issuer 1. These data were used to calculate all analyses. A data analytic firm provided QHP analytic results for three issuers, henceforth Issuer 2, Issuer 3, and Issuer 4, in lieu of patient-level data.

Additionally, national claims data from Medicare Part B and stand-alone Part D prescription drug plans (PDPs) were used to supplement the QHP analyses since limited QHP data were available for testing. Medicare PDPs were used as a supplement to QHP data because they offer a robust sample for measure testing, such as calculation of measure performance reliability. Medicare PDPs are similar to QHPs in that they are offered by private insurance companies and are responsible for providing safe and effective medication management. If variation in performance is similar among QHP products and Medicare PDPs, we could conclude this measure is generally applicable and reliable at the health plan level.

Testing results presented in this report use data from calendar years (CY) 2015 and 2016. The Measure Developer excluded Exchange products with less than 500 enrollees, in alignment with the QRS requirement that a QHP product have a minimum of 500 members in order to report measure data to CMS. This requirement is based on the 2018 Quality Rating System Measure Technical Specifications.[24] Also for consistency with QRS requirements, throughout the analyses the Measure Developer provided measure-specific results only when there were at least 30 denominator events or members.[24] The 500 member and 30 minimum denominator rules are not part of the measure specifications. The analyses followed these rules to reflect steps that would be taken if the measure were implemented into the Quality Rating System (QHP data). The 500 member and 30 minimum denominator rules were not applied to the Medicare data since the rules are specific to the Quality Rating System (QHP data).

Feasibility

To determine the feasibility of the measure, the Measure Developer examined the availability of the data elements. For measure calculation, the following data elements, which are entirely based on administrative claims, are required: member enrollment data, pharmacy claims, laboratory claims, institutional claims, and non-institutional claims. Laboratory results data can be used but are optional. Eligible members are identified using enrollment and pharmacy claims data. Exclusions for hospice care and cancer diagnoses are identified using institutional and non-institutional claims. Drug tests are identified using laboratory, institutional, and non-institutional claims, or optionally, laboratory results data containing LOINC codes for specified tests.

Expert Input

A workgroup composed of seven of the 12 TEP members met periodically for updates on the progress of measure development and to give feedback on measure specifications. The workgroup members represented the following areas of expertise: performance measurement, medication safety, quality improvement, healthcare delivery, QHPs, and the insurance industry. Additionally, in summer 2017, the Measure Developer included three subject matter experts in the workgroup, including a physician who specializes in pain management, an expert in laboratory testing, and an administrative billing consultant specializing in healthcare administrative coding (see *Technical Expert Panel and Subject Matter Experts,* Appendix A). Three webinar meetings were held with the clinical experts to review the measure specifications and testing results.

Development of the Denominator

Denominator Inclusion Criteria

The proposed denominator includes patients aged 18 years and older as of the end of the measurement year who are prescribed COT. To define the denominator specifications, the Measure Developer reviewed the clinical practice guidelines and the literature, conducted an empirical analysis, compared specifications of existing measures with a similar focus, and consulted with the TEP and subject matter experts.

Denominator Exclusion Criteria

Denominator exclusions were determined by clinical justification and supported by examining the literature and other opioid quality measures that specify chronic use, as well as consulting with the TEP and subject matter experts.

Development of Numerator

The focus of the measure is on drug tests for patients on COT. Clinical practice guidelines were used to determine: 1) what types of drug tests should be used to define the numerator, and 2) when the drug tests should be conducted (see *Clinical Practice Guideline Recommendations,* Appendix B).[25-28] To specify drug tests using administrative claims, code sets for HCPCS, CPT, and LOINC were identified. Prior to 2016, all three code sets allowed for identifying tests for specific drugs/drug classes of interest. This changed in 2016 when the HCPCS drug test codes were

consolidated from 28 drug class-specific codes (Table 1) to four generic, non-specific codes (Table 2) that differ only by the number of drug classes being tested.

Table 1. 2015 HCPCS Drug Testing Codes – Specific to Drug

HCPCS Code	Description	HCPCS Code	Description	HCPCS Code	Description
G6030	Amitriptyline	G6041	Alkaloids, Urine, Quantitative	G6051	Flurazepam
G6031	Benzodiazepines	G6042	Amphetamine or Methamphetamine	G6052	Meprobamate
G6032	Desipramine	G6043	Barbiturates, Not Elsewhere Specified	G6053	Methadone
G6034	Doxepin	G6044	Cocaine or Metabolite	G6054	Methsuximide
G6035	Gold	G6045	Dihydrocodeinone	G6055	Nicotine
G6036	Assay of Imipramine	G6046	Dihydromorphinone	G6056	Opiate(s), Drug and Metabolites, Each Procedure
G6037	Nortriptyline	G6047	Dihydrotestosterone	G6057	Phenothiazine
G6038	Salicylate	G6048	Dimethadione	G6058	Drug Confirmation, Each Procedure
G6039	Acetaminophen	G6049	Epiandrosterone		
G6040	Alcohol (Ethanol); Any Specimen Except Breath	G6050	Ethchlorvynol		

Table 2. 2016 HCPCS Drug Testing Codes – Specific Only to Number of Drug Classes Tested

HCPCS Code	Description
G0480	Drug test(s), definitive, utilizing (1) drug identification methods able to identify individual drugs and distinguish between structural isomers (but not necessarily stereoisomers), including, but not limited to GC/MS (any type, single or tandem) and LC/MS (any type, single or tandem and excluding immunoassays (e.g., IA, EIA, ELISA, EMIT, FPIA) and enzymatic methods (e.g., alcohol dehydrogenase), (2) stable isotope or other universally recognized internal standards in all samples (e.g., to control for matrix effects, interferences and variations in signal strength), and (3) method or drug-specific calibration and matrix-matched quality control material (e.g., to control for instrument variations and mass spectral drift); **qualitative or quantitative, all sources, includes specimen validity testing, per day; 1-7 drug class(es),** including metabolite(s) if performed
G0481	Drug test(s), definitive, utilizing (1) drug identification methods able to identify individual drugs and distinguish between structural isomers (but not necessarily stereoisomers), including, but not limited to GC/MS (any type, single or tandem) and LC/MS (any type, single or tandem and excluding immunoassays (e.g., IA, EIA, ELISA, EMIT, FPIA) and enzymatic methods (e.g., alcohol dehydrogenase), (2) stable isotope or other universally recognized internal standards in all samples (e.g., to control for matrix effects, interferences and variations in signal strength), and (3) method or drug-specific calibration and matrix-matched quality control material (e.g., to control for instrument variations and mass spectral drift); **qualitative**

HCPCS Code	Description
	or quantitative, all sources, includes specimen validity testing, per day; 8-14 drug class(es), including metabolite(s) if performed
G0482	Drug test(s), definitive, utilizing (1) drug identification methods able to identify individual drugs and distinguish between structural isomers (but not necessarily stereoisomers), including, but not limited to GC/MS (any type, single or tandem) and LC/MS (any type, single or tandem and excluding immunoassays (e.g., IA, EIA, ELISA, EMIT, FPIA) and enzymatic methods (e.g., alcohol dehydrogenase), (2) stable isotope or other universally recognized internal standards in all samples (e.g., to control for matrix effects, interferences and variations in signal strength), and (3) method or drug-specific calibration and matrix-matched quality control material (e.g., to control for instrument variations and mass spectral drift); **qualitative or quantitative, all sources, includes specimen validity testing, per day; 15-21 drug class(es),** including metabolite(s) if performed
G0483	Drug test(s), definitive, utilizing (1) drug identification methods able to identify individual drugs and distinguish between structural isomers (but not necessarily stereoisomers), including, but not limited to GC/MS (any type, single or tandem) and LC/MS (any type, single or tandem and excluding immunoassays (e.g., IA, EIA, ELISA, EMIT, FPIA) and enzymatic methods (e.g., alcohol dehydrogenase), (2) stable isotope or other universally recognized internal standards in all samples (e.g., to control for matrix effects, interferences and variations in signal strength), and (3) method or drug-specific calibration and matrix-matched quality control material (e.g., to control for instrument variations and mass spectral drift); **qualitative or quantitative, all sources, includes specimen validity testing, per day; 22 or more drug class(es),** including metabolite(s) if performed

The Measure Developer tested the effect of the 2016 changes in HCPCS codes. A sensitivity analysis was conducted to test the validity of a measure that defines specific drug tests using CPT and LOINC codes and also includes the non-specific drug test codes defined by HCPCS (Table 3).

Table 3. Testing 2016 HCPCS Code Changes on Measure Performance

Type of Codes	Used in measure specifications		Identifies specific types of drug tests	
	Prior to 2016	2016 and Beyond	Prior to 2016	2016 and Beyond
HCPCS	✔	✔	✔	✘
CPT	✔	✔	✔	✔
LOINC	✔	✔	✔	✔

To examine the potential impact of the 2016 HCPCS code changes, 2015 data were used, and the following steps were taken:

1. Calculating measure rates using all of the 2015 codes (HCPCS, CPT, and LOINC) specific to drug classes identified in AACC Academy guidelines for routine screening.[25]
2. Calculating new measure rates using 2015 CPT and LOINC codes specific to the drug classes of interest, and all of the 2015 HCPCS codes (regardless of type of drug/drug class). This set of 2015 HCPCS codes represented the HCPCS code change in 2016 and beyond.
3. Comparing the two versions of the 2015 measure rates to examine the impact of not identifying specific drug classes with HCPCS codes.

Reliability

Measure Performance Reliability

To assess measure precision in the context of the observed variability across products, the Measure Developer used the signal-to-noise approach, which determines how well performance can be distinguished between products. The signal-to-noise ratio was calculated as a function of the variance between products (signal) and the variance within a product (noise). Measure score reliability was estimated using a beta-binomial model. For the QHP data, the mean reliability was calculated across QHP products. Reliability estimates for Medicare PDPs were computed by using the methods of minimum denominator and volume categories, described by Scholle et al. (2008).[29] This method assumes the denominator size in each volume category is equal to the minimum for that category. As such, it provides a more conservative estimate of reliability for each volume category. This difference in approach to the data is due to the limited number of available QHP products.

Reliability scores can range from 0.0 to 1.0. A score of 0.0 implies that all variation is completely attributable to measurement error (i.e., noise or the product variance), whereas a reliability of 1.0 implies that all variation is caused by a real difference in performance across products. Adams showed that differences can be seen at a reliability of 0.7 and significant differences could be seen at a reliability of 0.9.[30] The rationale for the reliability analysis was based on Adams' work, and thus a minimum reliability score of 0.7 was used to indicate sufficient signal strength to discriminate performance between products.

The Measure Developer used the following formula to calculate the reliability of the measure rate for each PDP and QHP product, reflecting a signal-to-noise ratio.

$$Reliability = \frac{\sigma^2_{Between-PDPs}}{\sigma^2_{Between-PDPs} + \sigma^2_{Within-PDPs}}$$

In which $\sigma^2_{Between-PDPs}$ is the variance of scores between PDPs or QHP products and $\sigma^2_{Within-PDPs}$ is the variance within PDPs or QHP products.

Calculating Measure Performance

The Measure Developer determined gaps in performance by calculating the measure rate using 2015 and 2016 data from Issuer 1 and Medicare. To calculate the measure rate, the Measure Developer determined how many members in the denominator did not have at least one drug test during the measurement period. Measure rates for Issuer 2, Issuer 3, and Issuer 4 were calculated by a data analytic firm and provided to the Measure Developer.

Steps for calculating the measure rate included:

1. Calculating the denominator by including all members enrolled for 11 of 12 months during the measurement year, who are 18 years of age and older, with a days' supply of opioid medications of 90 days or more, and who do not have a cancer diagnosis or hospice care.
2. Calculating the numerator by including members from the denominator who do not have any claims for a drug test during the measurement year.

3. Calculating the measure rate as the number of members in the numerator divided by the number of members in the denominator multiplied by 100, with members being attributed to the products in which they were last enrolled during the measurement year. Lower measure rates are indicative of better quality.

Exclusions Analysis

To determine the effect of the exclusions on the measure rates, the Measure Developer calculated the pooled rates with and without each exclusion. Results for Issuer 2, Issuer 3, and Issuer 4 were calculated by a data analytic firm and provided to the Measure Developer.

Disparities Analyses

To assess whether disparities in measure performance exist between subpopulations, the Measure Developer used the method employed by the Agency for Healthcare Research and Quality (AHRQ) for the National Healthcare Quality and Disparities Report.[31] Two criteria were applied to determine meaningful differences between the performance for a reference group and another population group. A group's results may be interpreted as:

- Better than the reference group by at least a 10% relative difference and with a $p<0.05$
- Worse than the reference group by at least a 10% relative difference and with a $p<0.05$

Relative differences were calculated by subtracting the reference group from each demographic group and dividing it by the reference group. Statistical significance of the difference between two proportions was determined using a Z-test. Results for Issuer 2, Issuer 3, and Issuer 4 were calculated by a data analytic firm and provided to the Measure Developer.

Validity

Face validity was used to assess the validity of the measure. Face validity is a subjective assessment by experts of whether the measure results reflect the intent of the measure. In this context, the purpose of evaluating face validity is to determine whether performance scores resulting from the measure as specified can be used to distinguish good from poor quality of care. Face validity of the measure score was obtained by a TEP vote on April 4, 2018. Prior to the vote, the TEP was provided with the final measure specifications and presented the results of testing. After review and discussion, the TEP agreed to vote on the face validity of the measure. The Measure Developer asked the TEP members to vote on whether they agree, disagree, or are unable to rate the following face validity statement:

> *"The performance scores resulting from the measure Drug Testing for Individuals on Chronic Opioid Therapy (COT), as specified, can be used to distinguish good from poor plan-level quality related to the process of administrating at least one drug test during the measurement year among those with chronic opioid therapy."*

RESULTS

Data Description

Among the four QHP issuers in 2015, membership size ranged from 3,354 (Issuer 4) to 289,136 (Issuer 1). There was a similar distribution in membership size among the four issuers in 2016. In 2015, the range of average ages of members across QHP issuers was approximately 43 to 44. In 2015, all QHP issuers had almost an even distribution of males and females with Issuer 1 having slightly more females (51.9%) and the other three QHP issuers having slightly more males (54.1%-56.5%). The age and sex distributions were similar in 2016 in the QHP data.

In 2015, there were more than 32 million Medicare beneficiaries represented in the data. Among the more than 32 million Medicare beneficiaries, there were 56.7% (18,257,146 beneficiaries) who were enrolled in a PDP at some time during 2015 across 67 stand-alone Part D drug plans (PDPs). The average age of Medicare beneficiaries in stand-alone PDPs in 2015 was approximately 71 years and 55.2% were female. This description is similar to the characteristics of Medicare beneficiaries in stand-alone PDPs in 2016. For a full description of all the data, see Table 4 (2015) and Table 5 (2016) below.

Comparing the two populations, there was a similar distribution of males and females. Medicare beneficiaries were older by approximately 30 years compared to QHP members. Although these age differences were notable, the purpose of using the two populations is not to suggest they are comparable. Rather, if measure performance is similar across the two populations, then we can reasonably conclude that the measure is generally applicable at the health plan-level of analysis.

Table 4. Description of Data Used During Testing (2015)

Characteristics	Issuer 1	Issuer 2	Issuer 3	Issuer 4	Medicare PDPs
Total Number of QHP Products or PDPs	3	1	2	1	67
Total Member/ Beneficiary Sample Size Enrolled in a QHP Product/PDP	289,136	49,137	15,671	3,354	18,257,146
Sex n (% of Total Sample)					
Female	150,116 (51.9)	21,399 (43.5)	7,043 (44.9)	1,538 (45.9)	10,071,540 (55.2)
Male	139,020 (48.1)	27,738 (56.5)	8,628 (55.1)	1,816 (54.1)	8,185,606 (44.8)
Age n (% of Total Sample)					
< 18 years	9,584 (3.3)	3,600 (7.3)	1,578 (10.1)	247 (7.4)	111 (0.0)
18-26 years	38,590 (13.4)	3,633 (7.4)	1,640 (10.5)	333 (9.9)	89,804 (0.5)
27-44 years	81,098 (28.0)	12,486 (25.4)	5,671 (36.2)	1,022 (30.5)	864,242 (4.7)
45-64 years	152,252 (52.7)	28,965 (59.0)	6,603 (42.1)	1,711 (51.0)	2,813,147 (15.4)

Characteristics	Issuer 1	Issuer 2	Issuer 3	Issuer 4	Medicare PDPs
≥65 years	7,612 (2.6)	453 (0.9)	179 (1.1)	41 (1.2)	14,489,842 (79.4)
Race n (% of Total Sample)					
White/Caucasian	N/A	N/A	N/A	N/A	15,275,375 (83.7)
African-American	N/A	N/A	N/A	N/A	1,826,519 (10.0)
Hispanic	N/A	N/A	N/A	N/A	368,352 (2.0)
Other	N/A	N/A	N/A	N/A	608,822 (3.3)
Unknown	N/A	N/A	N/A	N/A	178,078 (1.0)

Table 5. Description of Data Used During Testing (2016)

Characteristics	Issuer 1	Issuer 2	Issuer 3	Issuer 4	Medicare PDPs
Total Number of QHP Products or PDPs	3	1	3	1	63
Total Member/ Beneficiary Sample Size Enrolled in a QHP Product/PDP	223,427	33,205	84,255	2,284	18,945,015
Sex n (% of Target Population)					
Female	116,111 (52.0)	14,546 (43.8)	38,433 (45.6)	1,027 (45.0)	10,433,654 (55.1)
Male	107,316 (48.0)	18,659 (56.2)	45,822 (54.4)	1,257 (55.0)	8,511,361 (44.9)
Age n (% of Target Population)					
< 18 years	8,536 (3.8)	3,077 (9.3)	8,618 (10.2)	207 (9.1)	99 (0.0)
18-26 years	27,732 (12.4)	2,445 (7.4)	8,268 (9.8)	236 (10.3)	85,827 (0.5)
27-44 years	58,419 (26.2)	8,584 (25.8)	27,730 (32.9)	724 (31.7)	844,283 (4.5)
45-64 years	121,304 (54.3)	18,756 (56.5)	38,748 (46.0)	1,089 (47.7)	2,801,328 (14.8)
≥65 years	7,436 (3.3)	343 (1.0)	891 (1.1)	28 (1.2)	15,213,478 (80.3)
Race n (% of Target Population)					
White/Caucasian	N/A	N/A	N/A	N/A	15,830,941 (83.6)
African-American	N/A	N/A	N/A	N/A	1,849,827 (9.8)
Hispanic	N/A	N/A	N/A	N/A	385,552 (2.0)
Other	N/A	N/A	N/A	N/A	647,648 (3.4)
Unknown	N/A	N/A	N/A	N/A	231,047 (1.2)

Feasibility

To determine the feasibility of the measure, the Measure Developer examined the availability of the data elements. The data elements (i.e., member enrollment data, pharmacy claims, laboratory claims, institutional claims, and non-institutional claims) that are required to calculate the measure were available in the health plan administrative claims datasets used to test the measure. No data elements were found to be missing. Therefore, the measure is feasible to specify and calculate using administrative claims data at the health plan product level.

Denominator

The proposed denominator includes patients aged 18 years and older as of the end of the measurement year who are prescribed COT.

Denominator Inclusion Analysis and Results

Defining Chronic Opioid Therapy

COT was defined by reviewing clinical practice guidelines, the literature, and other opioid quality measures. Results indicated that opioid prescriptions lasting at least 90 days are routinely considered chronic therapy. Therefore, the definition of COT should be specified as a days' supply of 90 days or more, with days' supply calculated as the sum of the days' supply for every prescription during the measurement year for opioid medications indicated for pain. Members qualify for the measure denominator if this sum is at least 90 days.

The following existing measures use this definition to specify COT:

- NQF 2940 Use of Opioids at High Dosage in Persons Without Cancer
- NQF 2951 Use of Opioids from Multiple Providers and at High Dosage in Persons Without Cancer
- MUC15-1169 Potential Opioid Overuse

Although MIPS measures #408, #412, and #414 specify COT as a prescription of opioids longer than six weeks in duration, the evidence base,[32-35] including the clinical practice guidelines[27,36] define COT as 90 days or more.

Defining Which Patients on Chronic Opioid Therapy To Include

To further define the scope of the denominator, the clinical practice guidelines were reviewed to determine whether only new patients starting COT, only established patients on COT, or both new and established patients taking COT should be included in the measure. Results suggested that drug testing is an important process for both new and established patients; four of five guidelines recommend testing at initiation of COT and all five guidelines recommend continued drug testing (defined differently as one to two times per year, random, ongoing, and periodic). More information on each of the clinical practice guidelines is provided in *Clinical Practice Guideline Recommendations,* Appendix B. Therefore, both new and established patients on COT will be included in the denominator, allowing for optimal clinical discretion as to when drug tests should be given to patients on COT.

Selection of Opioids

Specific opioid formulations were examined by the subject matter experts to determine which are used for chronic management of pain and therefore should be included to define patients on COT. Verified by the subject matter experts' review, the opioids selected for the measure include only those indicated for pain management. All routes of delivery will be included, except intravenous (IV) or epidural (EP) routes. The subject matter experts indicated that opioid formulations administered through IV or EP routes are not typically abused, nor are they typically prescribed to patients with chronic pain who do not have a cancer diagnosis and are not receiving hospice care. Additionally, using Issuer 1 pharmacy claims, an empirical analysis of routes of administering opioids suggests that almost all 2015 opioid pharmacy claims were for oral formulations (Table 6).

Table 6. Frequency of Opioid Routes in 2015 Issuer 1 Pharmacy Claims

Route	Count	%
Oral (PO)	119,977	98.4%
Transdermal (TD)	1,622	1.3%
Sublingual (SL)	314	0.3%
Buccal (BU)	13	0.0%
Nasal (NA)	11	0.0%
Rectal (PR)	1	0.0%
Epidural (EP)	0	0.0%
Intravenous (IV)	0	0.0%
Total	121,938	

Defining Age
Examination of the literature and discussions with the TEP and subject matter experts resulted in the specification of patients aged 18 years and older. In the examined literature where age was specified, no study included patients younger than 18 years of age.[2,4-6,8,9,37-46] Further, the clinical practice guidelines are intended for adults (i.e., patients at least 18 years of age), are based on evidence for adults, and do not provide recommendations for those younger than 18 years of age.[25-28,36]

In an effort to align the measure with other health plan quality measures, the Measure Developer examined HEDIS measures that specified age and used the entire measurement year as the measurement period.[47] Findings indicated that most of these measures define age as of the end of the measurement year. Therefore, this measure includes those 18 years and older as of the end of the measurement year.

Denominator Exclusion Results

Clinical practice guidelines for prescribing opioids routinely define their scope as outpatient treatment of patients with chronic pain, not including patients with active cancer, palliative care, or end-of-life care. The literature examined aligns with this scope, universally excluding cancer patients. Further, the CMS Opioid Misuse Strategy 2016[20] does not address treatment of patients with cancer or hospice care, based on the clinical practice guidelines and current literature. The rationale behind this focus is that such patients require case-by-case decisions made by providers that are based on therapeutic goals and ethical considerations. This rationale was supported by

both the TEP and subject matter experts, who agreed that drug testing in cancer or hospice populations is not currently supported by the evidence. The subject matter experts noted there could be potential in the future to include patients with cancer as more research and clinical practice guideline recommendations become available.

Additionally, these specifications align with other opioid quality measures that exclude these patient populations. The aligned quality measures include:

- NQF 2940 Use of Opioids at High Dosage in Persons Without Cancer
- NQF 2950 Use of Opioids from Multiple Providers in Persons Without Cancer
- NQF 2951 Use of Opioids from Multiple Providers and at High Dosage in Persons Without Cancer
- MUC15-1169 Potential Opioid Overuse

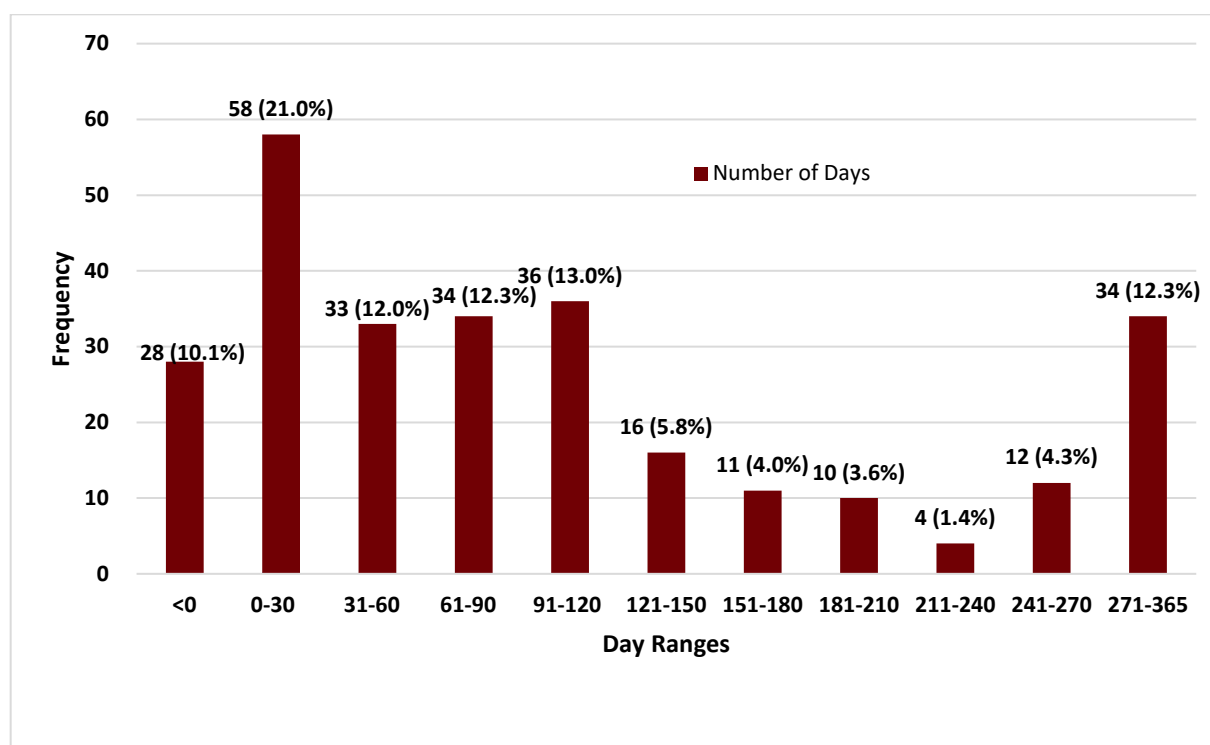## Numerator

Defining the Types of Drug Tests

The focus of the measure is on drug tests for patients on COT. To determine which drug tests should be used to define the numerator, the clinical practice guidelines were reviewed. Results of the review indicated the AACC Academy guideline is the most detailed in recommending specific types of drug screenings to perform.[25] The AACC Academy guideline suggests routine monitoring for the following: amphetamines, barbiturates, benzodiazepines, cannabinoids, cocaine, and opiates/opioids. The TEP and subject matter experts supported this decision, indicating the AACC Academy routine monitoring list contains common drugs of abuse, and screening for these drugs would meet a minimum standard of care. The TEP and subject matter experts also indicated that providers can test for substances beyond this list at their discretion. For instance, a provider could test a patient for a prescribed opioid, which would count towards the measure, and additionally test for alcohol if the patient has a known or suspected alcohol use disorder. Further, the TEP and subject matter experts supported inclusion of both general/presumptive and specific/confirmatory drug test types, which further allows providers to maintain discretion in deciding the types of tests that are appropriate on a case-by-case basis. This definition provides a minimum standard of care and maximizes clinical judgment.

Defining the Timing of Drug Tests

The Measure Developer reviewed clinical practice guidelines to specify when drug tests should be conducted to count toward the numerator. Results found that clinical practice guidelines vary in their recommendations for when to drug test (i.e., before initiation of therapy, one to two times per year, random, ongoing, and periodic). However, all guidelines recommend drug testing at least once annually.[25-28,48] See *Clinical Practice Guideline Recommendations,* Appendix B*,* for specific recommendations from each set of clinical practice guidelines.

Additionally, an empirical analysis was conducted using Issuer 1 data to examine patterns in drug testing. Results of the analysis support specifying the inclusion of drug tests that occur at any point during the measurement year. Figure 1 shows the distribution of the number of days between the first opioid prescription and the nearest drug test. In this figure, the nearest drug test may occur any number of days before or after the first prescription, which does not suggest a clear pattern and illustrates that drug tests are conducted at various times during the use of COT. Therefore, specifying the measure to allow credit for drug tests throughout the measurement year maintains clinical discretion related to the timing of such tests.

Figure 1. Number of Days between First Opioid Prescription and Nearest Drug Screen in 2016 Issuer 1 Data

Code Sets to Identify Drug Tests

To specify drug tests in the claims data, code sets for HCPCS, CPT, and LOINC were identified. As described in the Methods section, prior to 2016 all three code sets allowed for the identification of testing for specific drug classes of interest. In 2016, the HCPCS drug test codes were consolidated from 28 drug class-specific codes to four generic, non-specific codes that differ only according to the number of drug classes being tested. A sensitivity analysis was conducted to determine the impact of the 2016 HCPCS changes. The intent of the analysis was to determine if measure performance in 2015 was affected by not being able to identify only targeted drug classes. To perform the sensitivity analysis, the Measure Developer calculated two measure rates for 2015, using Issuer 1 and Medicare PDP data. Calculation instructions were given to an analytic firm and rates for Issuer 2, Issuer 3, and Issuer 4 were provided to the Measure Developer.

1. **Reference rate:** measure rates using the 2015 HCPCS, CPT, and LOINC codes specific to drug classes identified in AACC Academy guidelines for routine screening.[25] The specific rates represent the ability to specify drug classes within the 2015 HCPCS coding.
2. **Comparison rate:** measure rates using 2015 CPT and LOINC codes specific to the drug classes of interest, along with all 2015 HCPCS codes (regardless of type of drug/drug class). The comparison rates represent the effect the HCPCS drug test coding change would have on measure rates in 2016 and beyond.

The sensitivity analysis found the effect of the HCPCS code change was negligible. Potential false positives due to inclusion of generic codes that do not specify what drug is tested had little effect on measure performance. As shown in Table 7, the results for Issuer 1 found the Product A measure rate did not change, and the Product B measure rate decreased (improved) 0.1%, a non-significant difference (p = 0.918). Product C did not have ≥30 members denominators in 2015, therefore results have been suppressed in alignment with QRS requirements.[24] Results for Issuer 2 show a non-significant, 1.0% decrease (improvement) in the Product A measure rate (p = 0.3273). Measure rates did not change in any products for Issuer 3 and Issuer 4. The results among pooled Medicare PDPs produced a 0.5% decrease (improvement) in national measure rates and a 0.4% decrease (improvement) in mean rate across PDPs, a non-significant difference (p = 0.793). These findings suggest that when drug tests occur, they seldom are billed using HCPCS codes and the effect of HCPCS code changes had a minimal impact on measure rates. The remainder of results in this report uses 2016 data and code sets, which include the generic HCPCS drug test codes.

Table 7. HCPCS Code Change Comparison – Issuer 1 and Medicare 2015 Measure Rates

	Number of Members Without Drug Test: Numerator/Denominator	Percent Without Drug Test
Issuer 1 Product A (Reference)	33/45	73.3%
Issuer 1 Product A (Comparison)	33/45	73.3%
Issuer 1 Product B (Reference)	1,187/1,411	84.1%
Issuer 1 Product B (Comparison)	1,185/1,411	84.0%

	Number of Members Without Drug Test: Numerator/Denominator	Percent Without Drug Test
Issuer 1 Product C (Reference)	--	--
Issuer 1 Product C (Comparison)	--	--
Issuer 2 Product A (Reference)	655/944	69.4%
Issuer 2 Product A (Comparison)	646/944	68.4%
Issuer 3 Product A (Reference)	53/74	71.6%
Issuer 3 Product A (Comparison)	53/74	71.6%
Issuer 3 Product B (Reference)	22/38	57.9%
Issuer 3 Product B (Comparison)	22/38	57.9%
Issuer 4 Product A (Reference)	37/61	60.7%
Issuer 4 Product A (Comparison)	37/61	60.7%
Medicare (Reference)	1,138,225/1,674,121	68.0%
Medicare (Comparison)	1,130,194/1,674,121	67.5%

Because drug tests can no longer be specified in HCPCS codes, the measure is specified to target patients on COT without a drug test at least once in the measurement year. With this specification, lower measure rates are indicative of better quality. This approach continues to align with recommendations from the clinical practice guidelines[25-28] by identifying patients who objectively are not receiving care in adherence with clinical practice guidelines. A limitation of this approach is that the measure will give credit to health plans and providers for a small percentage of patients who may have received a drug test not associated with the management of COT (<1%). However, given that this change will have a negligible impact on the performance of a QHP, the TEP agreed that the benefit of the measure outweighed any limitation associated with the change to the HCPCS codes.

Reliability

Measure Performance Reliability

To assess measure precision in the context of the observed variability across products the Measure Developer used the signal-to-noise approach, which determines how well the performance of one product can be distinguished from another. To examine the reliability of the measure in the QHP data, the Measure Developer determined the reliability of each product using a beta-binomial analysis and then took the average reliability of products with at least 30 denominator members. Among the seven QHP products tested, reliability ranged from 59% to 99% with a mean reliability of 85% indicating the measure can reliably distinguish performance between QHP products. Beta-binomial reliability analyses with data from Medicare PDPs in 2016 showed that 100 beneficiaries are needed in the denominator to achieve reliability of at least 70%. With at least 100 beneficiaries in the denominator, the measure can reliably distinguish between PDPs, given the inherent variability in measure scores within PDPs due to sampling error.

Measure Performance Scores

Performance scores indicate that there is a significant gap in care related to drug tests for patients on COT. These results, in Table 8, show that over 80% of members on COT in Issuer 1, more than 62% of members on COT in Issuer 2, 58% or more of members on COT in Issuer 3 and 4, and 65% of beneficiaries on COT in Medicare PDPs are not receiving drug tests, and are therefore not in alignment with clinical practice guidelines. There was no significant difference in the distribution of rates between QHP products and Medicare PDPs (mean difference 0.7%, Wilcoxon Rank Sum Test p = 0.9433), indicating broad overlap in the distribution of product performance across these two populations.

Table 8. Issuer 1, Issuer 2, Issuer 3, Issuer 4 and Medicare 2016 Measure Rates

	Number of Members Without Drug Test: Numerator/Denominator	Percent Without Drug Test
Issuer 1 Product A	40/49	81.6%
Issuer 1 Product B	1,040/1,299	80.1%
Issuer 1 Product C	39/47	83.0%
Issuer 2 Product A	437/702	62.3%
Issuer 3 Product A	268/462	58.0%
Issuer 3 Product B	233/391	59.6%
Issuer 3 Product C	-	-
Issuer 4 Product A	21/36	58.3%
Medicare PDPs	1,050,828/1,617,089	65.0%

The measure rate distribution among seven QHP products (Table 9) suggests room for improvement as demonstrated by variation in measure performance and suboptimal performance. Among the 51 Medicare PDPs with at least 100 beneficiaries in the denominator (the threshold needed for achieve reliable results; Table 10) also suggests room for improvement as demonstrated by variation in measure performance and suboptimal performance.

Table 9. Measure Rate Distribution Among Issuer 1, Issuer 2, Issuer 3, and Issuer 4

Year	Mean	Minimum	P10	P25	P50	P75	P90	Maximum
2016	69.0%	58.0%	58.0%	58.3%	62.3%	81.6%	83.0%	83.0%

Table 10. Measure Rate Distribution Among Medicare PDPs

Year	Mean	Minimum	P10	P25	P50	P75	P90	Maximum
2016	69.7%	51.9%	58.7%	63.6%	71.3%	75.2%	79.0%	83.3%

Exclusions Analysis and Results

To determine the effect of the exclusions on the measure rates, pooled rates with and without each exclusion were calculated. As shown in Table 11, the impact on measure rates is minimal from excluding patients with cancer and those receiving hospice care. These exclusions will be retained in the measure specifications to align with other NQF-endorsed measures, clinical practices guidelines, and expert recommendations.

Table 11. 2016 COT Measure Rate by Exclusion Status

	Denominator	Numerator	Measure Rate	95% CI
Issuer 1				
No exclusions	1,670	1,341	80.3%	(78.4, 82.2)
Cancer excluded	1,395	1,119	80.2%	(78.1, 82.3)
Hospice excluded	1,664	1,336	80.3%	(78.4, 82.2)
Cancer and Hospice excluded	1,395	1,119	80.2%	(78.1, 82.3)
Issuer 2				
No exclusions	924	592	64.1%	(61.0, 67.2)
Cancer excluded	702	437	62.3%	(58.7, 65.8)
Hospice excluded	920	588	63.9%	(60.8, 67.0)
Cancer and Hospice excluded	702	437	62.3%	(58.7, 65.8)
Issuer 3				
No exclusions	1,173	700	59.7%	(56.8, 62.5)
Cancer excluded	881	517	58.7%	(55.4, 62.0)
Hospice excluded	1162	690	59.4%	(56.5, 62.3)
Cancer and Hospice excluded	880	517	58.8%	(55.4, 62.1)
Issuer 4				
No exclusions	44	24	54.5%	(39.8, 69.3)
Cancer excluded	36	21	58.3%	(42.2, 74.4)
Hospice excluded	44	24	54.5%	(39.8, 69.3)
Cancer and Hospice excluded	36	21	58.3%	(42.2, 74.4)
Medicare PDPs				
No exclusions	2,397,909	1,600,658	66.8%	(66.7, 66.8)
Cancer excluded	1,657,545	1,086,634	65.6%	(65.5, 65.6)
Hospice excluded	2,320,003	1,533,032	66.1%	(66.0, 66.1)
Cancer and Hospice excluded	1,617,089	1,050,828	65.0%	(64.9, 65.1)

Disparities Analyses

Sex and age were the sociodemographic variables available in the datasets to test disparities in care (Tables 12-15). Based on the methodology to determine disparities in care used by the AHRQ for the National Healthcare Quality and Disparities Report, disparities exist when there is a significant difference in measure rates as well as a 10% relative difference. In Tables 12-15, relative differences were reported where feasible; dashes indicate where small sample sizes precluded these results.

In the QHP data, the small sample sizes of many of the products restricted the ability to conduct disparities analysis for sex or age. Overall, QHP measure rates stratified by sex and age did not suggest disparities in care. Although there is a statistically significant difference between the 2016 measure rates for males and females in Issuer 1 Product B (p = 0.0224), the relative difference of 1% is less than the AHRQ standard of 10%, meaning these statistics do not meet the threshold set by AHRQ to be classified as a disparity. Differences in the 27-44 and over-65 age groups from the reference in Issuer 1 Product B are also not classified as disparities due to insufficient relative differences. In Issuer 3 Product B, there was a disparity detected between the 27-44 and 45-64 age groups (relative difference = -22.1%; p = 0.0178). However, additional data from the QHP population is needed to determine whether there are disparities across other age groups and whether drug testing occurs more frequently in younger or older QHP cohorts.

Using data from Medicare PDPs, disparities analyses suggest that females and older adults are tested significantly less often than males and younger cohorts. The presence of disparities in the national Medicare data set supports the need for measurement since we may expect to see the same disparities in the QHP population if we were to have access to a nationally representative data set.

Table 12. 2016 QHP Disparities Analyses - Sex

Product	Sex	Denominator	Numerator	Measure Rate	Relative Difference	P-Value
Issuer 1						
A	Female	-	-	-	Reference	
A	Male	30	24	80.0%	-	-
B	Female	735	591	80.4%	Reference	
B	Male	564	449	79.6%	1.0%	0.0224
C	Female	-	-	-	Reference	
C	Male	-	-	-	-	-
Issuer 2						
A	Female	440	269	61.1%	Reference	
A	Male	262	168	64.1%	4.9%	0.2150
Issuer 3						
A	Female	301	170	56.5%	Reference	
A	Male	161	98	60.9%	7.8%	0.1811
B	Female	234	140	59.8%	Reference	
B	Male	157	93	59.2%	-1.0%	0.4533
C	Female	-	-	-	Reference	
C	Male	-	-	-	-	-
Issuer 4						
A	Female	-	-	-	Reference	
A	Male	-	-	-	-	-

Table 13. 2016 Medicare PDPs Disparities Analyses - Sex

Number of PDPs	Sex	Denominator	Numerator	Measure Rate	Relative Difference	P-Value
62	Female	1,130,728	766,606	67.8%	Reference	
62	Male	486,361	284,222	58.4%	13.9%	0.0008

Table 14. 2016 QHP Disparities Analyses - Age

Product	Age Group	Denominator	Numerator	Measure Rate	Relative Difference	P-Value
Issuer 1						
A	18-26	-	-	-	-	-
A	27-44	-	-	-	-	-
A	45-64	40	33	82.5%	Reference	
A	65+	-	-	-	-	-
B	18-26	-	-	-	-	-
B	27-44	195	155	79.5%	-0.6%	0.0314
B	45-64	1,065	852	80.0%	Reference	
B	65+	33	28	84.9%	6.1%	0.0485
C	18-26	-	-	-	-	-
C	27-44	-	-	84.2%	-	-
C	45-64	-	-	-	Reference	
C	65+	-	-	-	-	-
Issuer 2						
A	18-26	-	-	-	-	-
A	27-44	108	61	56.5%	-10.2%	0.1049
A	45-64	579	364	62.9%	Reference	
A	65+	-	-	-	-	-
Issuer 3						
A	18-26	-	-	-	-	-
A	27-44	71	42	59.2%	2.5%	0.4099
A	45-64	383	221	57.7%	Reference	
A	65+	-	-	-	-	-
B	18-26	-	-	-	-	-
B	27-44	71	34	47.9%	-22.1%	0.0178
B	45-64	314	193	61.5%	Reference	
B	65+	-	-	-	-	-
C	18-26	-	-	-	-	-
C	27-44	-	-	-	-	-
C	45-64	-	-	-	Reference	
C	65+	-	-	-	-	-
Issuer 4						
A	18-26	-	-	-	-	-
A	27-44	-	-	-	-	-
A	45-64	-	-	-	Reference	
A	65+	-	-	-	-	-

Table 15. 2016 Medicare PDPs Disparities Analyses - Age

Number of PDPs	Age Group	Denominator	Numerator	Measure Rate	Relative Difference	P-Value
62	18-26	2,459	1,170	47.6%	36.9%	<0.0001
62	27-44	92,972	36,569	39.3%	47.9%	<0.0001
62	45-64	456,325	209,954	46.0%	39.0%	<0.0001
62	65+	1,065,333	803,135	75.4%	Reference	

Validity

Face validity was used to assess the validity of the measure. The TEP voted on whether they agreed, disagreed, or were unable to rate face validity of the measure. All nine TEP members (100%) voted in agreement that the measure can be used to distinguish good from poor plan-level quality of care related to the patients on COT (Table 16).

Table 16. Workgroup Voting Results

Do you agree with the following statement for the current measure (i.e., reporting the inverse rate)? "The performance scores resulting from the measure *Drug Testing for Individuals on Chronic Opioid Therapy (COT)*, as specified, can be used to distinguish good from poor plan-level quality related to the process of administering at least one drug test during the measurement year among those with chronic opioid therapy."
Yes = 9/9 (100%)
No = 0/9 (0%)
Abstention = 0/9 (0%)

SUMMARY

Measure Information

The following describes the final proposed specifications for the measure, *Drug Testing for Individuals on Chronic Opioid Therapy (COT)*.

Denominator

The target population for this measure is QHP members aged 18 years and older as of the end of the measurement year and prescribed COT during the measurement year. Eligible members must be continuously enrolled in a QHP, i.e., 11 out of 12 months during the measurement year or enrolled with no gaps in enrollment until the month of death, if applicable, in the measurement year. Members are excluded if they have had any claims indicating a cancer diagnosis or hospice care at any time during the measurement year.

The measurement year is defined as 12 consecutive months. COT is defined as at least 90 days' cumulative supply of any combination of opioid medications indicated for pain during the measurement period that are identified using pharmacy claims.

The active ingredient of the opioid medications is limited to formulations indicated for pain and delivered through any route except intravenous (IV) or epidural (EP). Opioid medications are specified in the supplementary materials for this report.

Numerator

The numerator is defined as members in the denominator who do not have at least one claim for a drug test during the measurement year. The entire measurement year in which a member is continuously enrolled is the time period used to identify at least one claim for a drug test.

A drug test is one that is either identified through HCPCS drug test codes or through specified CPT or LOINC codes for presumptive or definitive drug screens/tests for at least one of the following targeted drug classes: amphetamines, barbiturates, benzodiazepines, cannabinoids, cocaine, and opiates/opioids. Drug tests are specified in the supplementary materials for this report.

Measure Implementation

Implementation of the measure *Drug Testing for Individuals on Chronic Opioid Therapy (COT)* can improve adherence to clinical practice guidelines, enhance patient safety, and aid health plans and providers in identifying patients who are engaging in aberrant drug-related behaviors. Providers can use the information gained from drug test results to identify aberrant drug-related behaviors, influence referrals for substance use disorder, initiate patient education to prevent potential drug-drug interactions, and inform the provider if prescribed opioids are not evidenced. Therefore, it is anticipated that the measure will improve care for patients on COT.

Drug testing individuals on COT is a relatively low burden process that is critical for safe management of patients on COT. The burden will vary by patient and can increase depending on the type of drug test used. However, given the magnitude of the misuse of opioids and potential for ADEs related to COT, the benefit of the measure outweighs the associated burden.

TEP members were asked if they agreed with the recommendation to consider implementation of the measure in the QRS and 89% (8/9) agreed, acknowledging that the measure covers a high-priority need in the QRS. One TEP member suggested developing a companion measure to flag high rates of drug testing for the purpose of detecting potential over-testing and billing fraud, which was not an issue in the datasets used in testing but historically has been problematic among a few outliers. Finally, a TEP member suggested including this measure in other measurement programs due to its importance and to prevent unfairly targeting QHP members as needing different levels of drug testing compared with other populations.

Related Existing Measures

Similarities Between COT and Related Measures

With similar target populations (denominator), 11 existing measures are related to the measure *Drug Testing for Individuals on Chronic Opioid Therapy (COT)* through their inclusion of noncancer patients on opioid therapy. It should be noted that the measures described in this section do not include patients with opioid use disorders. Having a diagnosis of opioid use disorder is not the focus of the COT measure. Instead, this measure aims to improve the management of patients on COT. Potentially, it will reduce the number of patients with opioid use disorders by promoting information that can be used to inform treatment decisions while safely maintaining patients who need COT and adhere to the plan of care.

The following list shows existing measures related to opioid therapy:

Related measures with denominators targeting patients on COT (defined as 90 days or more of COT)

- o NQF 2940 Use of Opioids at High Dosage in Persons Without Cancer
- o NQF 2951 Use of Opioids from Multiple Providers and at High Dosage in Persons Without Cancer
- o MUC15-1169 Potential Opioid Overuse

Related measures with denominators targeting patients on COT (definition of COT: six weeks or 180 days)

- o MIPS Quality ID 408: Opioid Therapy Follow-Up Evaluation (not endorsed by NQF)
- o MIPS Quality ID 412: Documentation of Signed Opioid Treatment Agreement (not endorsed by NQF)
- o MIPS Quality ID 414: Evaluation or Interview for Risk of Opioid Misuse (not endorsed by NQF)
- o MUC14-X3376 Consideration of Non-Pharmacologic Interventions

Related measures with denominators targeting patients on opioids (not specific to chronic use)

- o NQF 1617 Patients Treated with an Opioid who are Given a Bowel Regimen
- o NQF 2950 Use of Opioids from Multiple Providers in Persons Without Cancer
- o NQF 3316e Safe Use of Opioids - Concurrent Prescribing of Opioids

       o MUC14-X3770 Overuse of Opioid Containing Medications For Primary Headache Disorders

Differences Between COT and Related Measures

Although the measures are related through similar target populations, only the measure described in this report focuses on drug testing for patients on COT. Comparatively, the existing complementary measures target follow-up evaluations, documentation of signed treatment agreements, high dosage opioid use, opioids from multiple providers, consideration of nonpharmacological care, questionnaire or interview-based evaluations for risk of opioid misuse, and potential overuse.

The COT measure is the only opioid-focused measure that promotes a process of care that gives the provider objective evidence of aberrant behavior. This is important in the management of patients on COT due to evidence suggesting that patients may not be forthcoming about aberrant behaviors.[40] Additionally, this is the only measure that promotes a process of care that gives providers actionable information during the point of care or shortly thereafter, making it more actionable than the other related existing measures.

Benefits of the COT Measure

The measure *Drug Testing for Individuals on Chronic Opioid Therapy (COT)* addresses patient safety at a potentially earlier point in the continuum of care. It will aid health plans and providers in identifying patients who are engaging in aberrant drug-related behaviors, potentially before a substance use disorder or ADE occurs. While follow-up appointments, signed treatment agreements, and use of a risk assessment tool are important, a patient could be misusing opioids or other substances that have the potential to interact with opioid therapy despite attending a follow-up appointment, signing an agreement, or filling out a risk assessment tool.

Furthermore, identifying patients who have opioid prescriptions at high dosages from multiple providers is extremely important information, but the effectiveness of the measure that targets this information is dependent upon health plans relaying the information to the providers and the providers then deciding to act upon it. In comparison, the drug testing for patients on COT measure promotes a process of care that puts clinical information directly into the hands of providers and into patients' medical records. Additionally, identifying patients who have opioid prescriptions at high dosages from multiple providers does not indicate usage of illicit substances, drugs that interact with opioids, or diversion of opioids. Therefore, the measure *Drug Testing for Individuals on Chronic Opioid Therapy (COT)* not only complements related existing measures but also addresses multiple aspects of care related to the management of patients on COT.

CONCLUSION

The measure *Drug Testing for Individuals on Chronic Opioid Therapy (COT)* addresses the Meaningful Measures priority area of Prevention and Treatment of Opioid and Substance Use Disorders,[18] and will address a gap in the quality of care received by members of QHP products. As demonstrated by the analyses conducted in the testing of this measure, 80% or more of members on COT in Issuer 1, 62% or more of members on COT in Issuer 2, 58% or more of members on COT in Issuer 3 and 4, and 65% of beneficiaries on COT in Medicare PDPs are not receiving drug testing in alignment with clinical practice guidelines. Improving rates of drug testing will enhance patient safety by aiding health plans and providers in identifying patients engaging in aberrant drug-related behaviors, which can lower patient risk for ADEs, suicide, and potential addiction as well as improve patient adherence to the plan of care. Finally, the measure aligns where possible with related existing measures and meets the scientific standards for quality measures established by CMS and NQF. In summary, implementation of this measure will be informative to patients, providers, and health plans and is anticipated to lead to improvements in the quality of care COT patients receive.

REFERENCES

1.   Centers for Medicare & Medicaid Services. *Blueprint for the CMS Measures Management System Version 13.0.* Baltimore, MD: US Department of Health and Human Services; 2017. https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/MMS/Downloads/Blueprint-130.pdf. Accessed June 21, 2018.

2.   Brennan PL, Del Re AC, Henderson PT, Trafton JA. Healthcare system-wide implementation of opioid-safety guideline recommendations: The case of urine drug screening and opioid-patient suicide- and overdose-related events in the Veterans Health Administration. *Transl Behav Med.* 2016;6(4):605-612. doi: 10.1007/s13142-016-0423-7.

3.   Morasco BJ, Krebs EE, Cavanagh R, Hyde S, Crain A, Dobscha SK. Treatment changes following aberrant urine drug test results for patients prescribed chronic opioid therapy. *J Opioid Manag.* 2015;11(1):45-51. doi: 10.5055/jom.2015.0251.

4.   Knezevic NN, Khan OM, Beiranvand A, Candido KD. Repeated quantitative urine toxicology analysis may improve chronic pain patient compliance with opioid therapy. *Pain Physician.* 2017;20(2S):S135-S145.

5.   Starrels JL, Becker WC, Weiner MG, Li X, Heo M, Turner BJ. Low use of opioid risk reduction strategies in primary care even for high risk patients with chronic pain. *J Gen Intern Med.* 2011;26(9):958-964. doi: 10.1007/s11606-011-1648-2.

6.   Lange A, Lasser KE, Xuan Z, et al. Variability in opioid prescription monitoring and evidence of aberrant medication taking behaviors in urban safety-net clinics. *Pain.* 2015;156(2):335-340. doi: 10.1097/01.j.pain.0000460314.73358.ff.

7.   Gilbert JW, Wheeler GR, Mick GE, et al. Urine drug testing in the treatment of chronic noncancer pain in a Kentucky private neuroscience practice: the potential effect of Medicare benefit changes in Kentucky. *Pain Physician.* 2010;13(2):187-194.

8.   Barth KS, Becker WC, Wiedemer NL, et al. Association between urine drug test results and treatment outcome in high-risk chronic pain patients on opioids. *J Addict Med.* 2010;4(3):167-173. doi: 10.1097/ADM.0b013e3181c379ed.

9.   Liebschutz JM, Xuan Z, Shanahan CW, et al. Improving Adherence to Long-term Opioid Therapy Guidelines to Reduce Opioid Misuse in Primary Care: A Cluster-Randomized Clinical Trial. *JAMA Intern Med.* 2017;177(9):1265-1272. doi: 10.1001/jamainternmed.2017.2468.

10.   Downes JM, Klepser DG, Foster J, Nelson M. Development of a standardized approach for managing opioids in adults with chronic noncancer pain. *Am J Health Syst Pharm.* 2018;75(5):321-326. doi: 10.2146/ajhp161012.

11.   Substance Abuse and Mental Health Services Administration. *The CBHSQ Report. January 25, 2017.* Rockville, MD: Department of Health and Human Services; 2017. https://www.samhsa.gov/data/sites/default/files/report_2734/Spotlight-2734.pdf. Accessed June 15, 2018.

12.   Washington State Agency Medical Directors' Group. *Interagency Guideline on Prescribing Opioids for Pain.* Olympia, WA: State of Washington Department of Health;

2015. http://www.agencymeddirectors.wa.gov/Files/2015AMDGOpioidGuideline.pdf.
Accessed June 4, 2018.

13.   Budnitz DS, Pollock DA, Weidenbach KN, Mendelsohn AB, Schroeder TJ, Annest JL.
National surveillance of emergency department visits for outpatient adverse drug
events. *JAMA.* 2006;296(15):1858-1866. doi: 10.1001/jama.296.15.1858.

14.   The Council of Economic Advisers. *The Underestimated Cost of the Opioid Crisis.*
Washington, DC: Office of the Press Secretary; 2017.
https://www.whitehouse.gov/sites/whitehouse.gov/files/images/The%20Underestimat
ed%20Cost%20of%20the%20Opioid%20Crisis.pdf. Accessed May 21, 2018.

15.   Trump DJ. Establishing the President's Commission on Combating Drug Addiction and
the Opioid Crisis. *Office of the Federal Register*. 2017;82(62):16283-16285.

16.   Christie C, Baker C, Cooper R, Kennedy PJ, Madras B, Bondi P. *The President's
Commission on Combating Drug Addiction and the Opioid Crisis.* Washington, DC: Office
of the Press Secretary; 2017.
https://www.whitehouse.gov/sites/whitehouse.gov/files/images/Final_Report_Draft_1
1-1-2017.pdf. Accessed June 2, 2018.

17.   Hargan ED. Determination that a public health emergency exists. Washington, DC: Office
of the Press Secretary; 2017.
https://www.hhs.gov/sites/default/files/opioid%20PHE%20Declaration-no-sig.pdf.
Accessed June 8, 2018.

18.   Centers for Medicare & Medicaid Services. CMS Administrator Verma announces new
Meaningful Measures initiative and addresses regulatory reform; promotes innovation
at LAN summit. Baltimore, MD: US Department of Health and Human Services; 2017.
https://www.cms.gov/Newsroom/MediaReleaseDatabase/Press-releases/2017-Press-
releases-items/2017-10-30.html. Accessed June 1, 2018.

19.   Office of National Drug Control Policy. National Drug Control Strategy. 2016.
https://obamawhitehouse.archives.gov/sites/default/files/ondcp/policy-and-
research/2016_ndcs_final_report.pdf. Accessed June 9, 2018.

20.   Centers for Medicare & Medicaid Services. Opioid Misuse Strategy 2016. Baltimore, MD:
US Department of Health and Human Services; 2016. https://www.cms.gov/Outreach-
and-Education/Outreach/Partnerships/Downloads/CMS-Opioid-Misuse-Strategy-
2016.pdf. Accessed May 8, 2018.

21.   US Department of Health and Human Services, Office of Disease Prevention and Health
Promotion. *National Action Plan for Adverse Drug Event Prevention.* Washington, DC: US
Department of Health and Human Services; 2014. https://health.gov/hcq/pdfs/ade-
action-plan-508c.pdf. Accessed June 11, 2018.

22.   Ducoffe AR, York A, Hu DJ, Perfetto D, Kerns RD. National Action Plan for Adverse Drug
Event Prevention: recommendations for safer outpatient opioid use. *Pain Med.* 2016.
doi: 10.1093/pm/pnw106.

23.   Centers for Medicare & Medicaid Services. Federal Definitions for Health Insurance
Products and Plans. Baltimore, MD: US Department of Health and Human Services;

2016. https://www.cms.gov/CCIIO/Resources/Training-Resources/Downloads/product-vs-plan-ppt.pdf. Accessed June 12, 2018.

24. Centers for Medicare & Medicaid Services. 2018 Quality Rating System Measure Technical Specifications. Baltimore. MD: US Department of Health and Human Services; 2018. https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/QualityInitiativesGenInfo/Downloads/Revised_QRS-2018-Measure-Tech-Specs_20170929_508.pdf. Accessed July 13, 2018.

25. Langman L, Jannetto P. *Laboratory Medicine Practice Guidelines: Using Clinical Laboratory Tests to Monitor Drug Therapy in Pain Management Patients.* The American Association for Clinical Chemistry Academy; 2017. https://www.aacc.org/science-and-practice/practice-guidelines/using-clinical-laboratory-tests-to-monitor-drug-therapy-in-pain-management-patients. Accessed May 3, 2018.

26. Manchikanti L, Kaye AM, Knezevic NN, et al. Responsible, Safe, and Effective Prescription of Opioids for Chronic Non-Cancer Pain: American Society of Interventional Pain Physicians (ASIPP) Guidelines. *Pain Physician.* 2017;20(2S):S3-S92.

27. US Department of Veterans Affairs, US Department of Defense. *VA/DoD clinical practice guideline for opioid therapy for chronic pain.* Washington, DC 2017. https://www.healthquality.va.gov/guidelines/Pain/cot/VADoDOTCPG022717.pdf.

28. Dowell D, Haegerich T, Chou R. CDC guideline for prescribing opioids for chronic pain — United States, 2016. Atlanta, GA: Centers for Disease Control and Prevention; 2016. http://www.cdc.gov/mmwr/volumes/65/rr/rr6501e1.htm. Accessed June 6, 2018.

29. Scholle SH, Roski J, Adams JL, et al. Benchmarking physician performance: reliability of individual and composite measures. *Am J Manag Care.* 2008;14(12):833-838.

30. Adams JL, Mehrotra A, Thomas JW, McGlynn EA. Physician cost profiling--reliability and risk of misclassification. *N Engl J Med.* 2010;362(11):1014-1021. doi: 10.1056/NEJMsa0906323.

31. Agency for Healthcare Research and Quality. *2014 National Healthcare Quality and Disparities Report.* Rockville, MD: US Department of Health and Human Services; 2015. http://www.ahrq.gov/sites/default/files/wysiwyg/research/findings/nhqrdr/nhqdr14/2014nhqdr.pdf. Accessed January 26, 2018.

32. Edlund MJ, Martin BC, Russo JE, DeVries A, Braden JB, Sullivan MD. The role of opioid prescription in incident opioid abuse and dependence among individuals with chronic noncancer pain: The role of opioid prescription. *Clin J Pain.* 2014;30(7):557-564. doi: 10.1097/AJP.0000000000000021.

33. Martin BC, Fan MY, Edlund MJ, Devries A, Braden JB, Sullivan MD. Long-term chronic opioid therapy discontinuation rates from the TROUP study. *J Gen Intern Med.* 2011;26(12):1450-1457. doi: 10.1007/s11606-011-1771-0.

34. Huffman KL, Shella ER, Sweis G, Griffith SD, Scheman J, Covington EC. Nonopioid substance use disorders and opioid dose predict therapeutic opioid addiction. *J Pain.* 2015;16(2):126-134. doi: 10.1016/j.jpain.2014.10.011.

35. Vanderlip ER, Sullivan MD, Edlund MJ, et al. National study of discontinuation of long-term opioid therapy among veterans. *Pain.* 2014;155(12):2673-2679. doi: 10.1016/j.pain.2014.09.034.

36. Chou R, Fanciullo GJ, Fine PG, et al. Clinical guidelines for the use of chronic opioid therapy in chronic noncancer pain. *J Pain.* 2009;10(2):113-130. doi: 10.1016/j.jpain.2008.10.008.

37. Sekhon R, Aminjavahery N, Davis CN, Jr., Roswarski MJ, Robinette C. Compliance with opioid treatment guidelines for chronic non-cancer pain (CNCP) in primary care at a Veterans Affairs Medical Center (VAMC). *Pain Med.* 2013;14(10):1548-1556. doi: 10.1111/pme.12164.

38. Morasco BJ, Duckart JP, Dobscha SK. Adherence to clinical guidelines for opioid therapy for chronic pain in patients with substance use disorder. *J Gen Intern Med.* 2011;26(9):965-971. doi: 10.1007/s11606-011-1734-5.

39. Kelly JP, Cook SF, Kaufman DW, Anderson T, Rosenberg L, Mitchell AA. Prevalence and characteristics of opioid use in the US adult population. *Pain.* 2008;138(3):507-513.

40. Schuckman H, Hazelett S, Powell C, Steer S. A validation of self-reported substance use with biochemical testing among patients presenting to the emergency department seeking treatment for backache, headache, and toothache. *Subst Use Misuse.* 2008;43(5):589-595. doi: 10.1081/JA-200030572.

41. Matteliano D, Chang YP. Describing prescription opioid adherence among individuals with chronic pain using urine drug testing. *Pain Manag Nurs.* 2015;16(1):51-59. doi: 10.1016/j.pmn.2014.04.001.

42. Turner JA, Saunders K, Shortreed SM, et al. Chronic opioid therapy risk reduction initiative: impact on urine drug testing rates and results. *J Gen Intern Med.* 2014;29(2):305-311. doi: 10.1007/s11606-013-2651-6.

43. Turner JA, Saunders K, Shortreed SM, et al. Chronic opioid therapy urine drug testing in primary care: prevalence and predictors of aberrant results. *J Gen Intern Med.* 2014;29(12):1663-1671. doi: 10.1007/s11606-014-3010-y.

44. Boscarino JA, Rukstalis M, Hoffman SN, et al. Risk factors for drug dependence among out-patients on opioid therapy in a large US health-care system. *Addiction.* 2010;105(10):1776-1782. doi: 10.1111/j.1360-0443.2010.03052.x.

45. Edlund MJ, Martin BC, Devries A, Fan MY, Braden JB, Sullivan MD. Trends in use of opioids for chronic noncancer pain among individuals with mental health and substance use disorders: the TROUP study. *Clin J Pain.* 2010;26(1):1-8. doi: 10.1097/AJP.0b013e3181b99f35.

46. Chen JH, Hom J, Richman I, Asch SM, Podchiyska T, Johansen NA. Effect of opioid prescribing guidelines in primary care. *Medicine (Baltimore).* 2016;95(35):e4760. doi: 10.1097/MD.0000000000004760.

47. National Committee for Quality Assurance. *HEDIS® 2018 Volume 2 Technical Specifications for Health Plans.* Washington, DC: National Committee for Quality Assurance; 2018.

48.   Chou R, Fanciullo GJ, Fine PG, Miaskowski C, Passik SD, Portenoy RK. Opioids for chronic noncancer pain: prediction and identification of aberrant drug-related behaviors: a review of the evidence for an American Pain Society and American Academy of Pain Medicine clinical practice guideline. *J Pain.* 2009;10(2):131-146. doi: 10.1016/j.jpain.2008.10.009.

Appendix A. TECHNICAL EXPERT PANEL AND SUBJECT MATTER EXPERTS

The Measure Developer would like to thank the Quality Rating System Measures for Qualified Health Plans Technical Expert Panel (TEP) and subject matter experts who provided important insight and feedback during measure development and testing. The Measure Developer would also like to thank the expert workgroup, which provided periodic feedback on the measure specifications. Expert workgroup members are denoted below with an asterisk.

Technical Expert Panel (TEP)

1. Andy Amster*, MSPH; Kaiser Permanente National Office
2. Marybeth Farquhar*, PhD, MSN, RN; URAC
3. Susan Fitzpatrick*, RN, BSN; Cigna Healthcare
4. Aparna Higgins, Duke-Margolis Center for Health Policy; Brandeis University
5. Jon Mark Hirshon*, MD, PhD, MPH; University of Maryland, School of Medicine
6. Christine Hunter, MD; US Office of Personnel Management
7. Carol Keegan, PhD; Patient representative
8. Dana Mukamel, PhD; University of California, Irvine
9. Chinwe Nwosu, MS; America's Health Insurance Plans
10. Derek Robinson*, MD, MBA, FACEP; Health Care Service Corporation
11. Arlene Salamendra*, Patient representative
12. Ted von Glahn*, MSPH; von Glahn Consulting

Subject Matter Experts

1. Paul Jannetto*, PhD; Mayo Clinic
2. Graves T Owen*, MD; Texas Pain Rehabilitation Institute
3. Kara McVey, CPC, CPMA; ILEX Consulting LLC

Appendix B. CLINICAL PRACTICE GUIDELINE RECOMMENDATIONS

Clinical Guideline	Guideline Recommendation	Evidence Level	Strength of Recommendation
The American Association for Clinical Chemistry (AACC) Academy (2017) [25]	"Testing biological specimens for drugs/drug metabolites is effective for detecting the use of relevant over-the-counter, prescribed, and non-prescribed drugs and of illicit substances in pain management patients. Laboratory testing does not specifically identify most other outcomes, but should be used in conjunction with additional information to detect other outcomes in pain management patients."	I - Evidence includes consistent results from well-designed, well-conducted studies in representative populations.	A – The AACC Academy strongly recommends adoption; there is good evidence that it improves important health outcomes, and it concludes that benefits substantially outweigh harms.
	"Based on level II evidence, baseline drug testing should be performed prior to initiation of acute or chronic controlled substance therapy. In addition, random drug testing should be performed at a minimum of one to two times a year for low-risk patients (based on history of past substance abuse/addiction, aberrant behaviors, and opioid risk screening criteria), with increasing frequency for higher-risk patients prescribed controlled substances."	II - Evidence is sufficient to determine effects, but the strength of the evidence is limited by the number, quality, or consistency of the individual studies; generalizability to routine practice; or indirect nature of the evidence.	A – The AACC Academy strongly recommends adoption; there is good evidence that it improves important health outcomes, and it concludes that benefits substantially outweigh harms.
American Society of Interventional Pain Physicians (2017) [26]	Recommended <u>initial</u> steps of opioid therapy include: "UDT must be implemented at initiation of opioid therapy, along with continued adherence monitoring to identify patients who are non-compliant or abusing prescription drugs or illicit drugs."	II – Moderate. Evidence obtained from at least one relevant high quality randomized controlled trial or multiple relevant moderate or low quality randomized controlled trials OR evidence obtained from at least two high quality relevant observational studies or large case series for assessment of preventive measures, adverse consequences,	Moderate — There is moderate confidence that the recommendation reflects best practice. This is based on a) good evidence for a true net effect; b) consistent results with minor and/or few exceptions; c) minor and/or few concerns about study quality; and/or d) the extent of panelists' agreement. Other compelling considerations may also warrant a moderate recommendation.

Clinical Guideline	Guideline Recommendation	Evidence Level	Strength of Recommendation
		and effectiveness of other measures	
	Recommended monitoring for <u>adherence</u> and side effects include: "In order to reduce prescription drug abuse and doctor shopping, adherence monitoring by UDT and PDMPs provide evidence that is essential to the identification of those patients who are non-compliant or abusing prescription drugs or illicit drugs."	I-II – Strong-Moderate. Evidence obtained from multiple relevant high quality randomized controlled trials for effectiveness OR evidence obtained from multiple relevant high quality observational studies or large case series for assessment of preventive measure, adverse consequences, and effectiveness of other measures	Moderate to Strong — There is moderate to high confidence that the recommendation reflects best practice.
U.S. Department of Veterans Affairs (2017)[27]	"We recommend implementing risk mitigation strategies upon initiation of long-term opioid therapy, starting with an informed consent conversation covering the risks and benefits of opioid therapy as well as alternative therapies. The strategies and their frequency should be commensurate with risk factors and include: • Ongoing, random urine drug testing (including appropriate confirmatory testing) • Checking state prescription drug monitoring programs • Monitoring for overdose potential and suicidality • Providing overdose education • Prescribing of naloxone rescue and accompanying education"	Moderate for UDT	Strong for UDT — the Work Group is highly confident that desirable outcomes outweigh undesirable outcomes.
Centers for Disease Control and Prevention (2016) [28]	"When prescribing opioids for chronic pain, clinicians should use urine drug testing before starting opioid therapy and consider urine drug testing at least annually to assess for prescribed medications as well as other controlled prescription drugs and illicit drugs."	4 — Clinical experience and observations, observational studies with important limitations, or randomized clinical trials with several major limitations	B — Individual decision-making needed

Clinical Guideline	Guideline Recommendation	Evidence Level	Strength of Recommendation
The American Pain Society and The American Academy of Pain Medicine (2009) [36]	"In patients on COT who are at high risk or who have engaged in aberrant drug-related behaviors, clinicians should periodically obtain urine drug screens or other information to confirm adherence to the COT plan of care."	Low — Evidence is insufficient to assess effects on health outcomes	Strong for patients at high risk
	"In patients on COT not at high risk and not known to have engaged in aberrant drug-related behaviors, clinicians should consider periodically obtaining urine drug screens or other information to confirm adherence to the COT plan of care."	Low — Evidence is insufficient to assess effects on health outcomes	Weak for patients at low risk