



U.S. Department of Health & Human Services  
**Office of Inspector General**

**Toolkit for Calculating Opioid Levels  
and Identifying Patients At Risk of  
Misuse or Overdose: R and SQL  
Programming Code**



May 2020  
OEI-02-17-00561



## **Toolkit for Calculating Opioid Levels and Identifying Patients At Risk of Misuse or Overdose: R and SQL Programming Code**

### **What is the toolkit?**

This toolkit provides detailed steps for using prescription drug claims data to analyze patients' opioid levels and identify certain patients who are at risk of opioid misuse or overdose. The toolkit includes R and Structured Query Language (SQL) programming code. R is a free programming language and software program for statistical computing. SQL is a programming language used in many software programs.

This toolkit is a companion to the previously released Office of Inspector General (OIG) *Toolkit: Using Data Analysis To Calculate Opioid Levels and Identify Patients At Risk of Misuse or Overdose*, which includes SAS programming code. Similar to the previous one, this toolkit is based on the methodology that OIG developed for its extensive work on opioid use in Medicare Part D.

This toolkit provides highly technical information to assist our public and private sector partners—such as Medicare Part D plan sponsors, private health plans, and State Medicaid Fraud Control Units—with analyzing their own prescription drug claims data to help combat the opioid crisis.

### **Why did OIG create the toolkit?**

The opioid crisis remains a public health emergency. In 2018 alone, there were 46,802 opioid-related overdose deaths in the United States. As one of the lead Federal agencies fighting health care fraud, OIG is committed to supporting our public and private partners in their efforts to curb the opioid epidemic. This toolkit builds on OIG's previous toolkit by providing programming code in two additional languages—R and SQL. In response to the first toolkit, our partners expressed interest in OIG providing this code in these additional programming languages. This toolkit and the accompanying code can be used to analyze claims data for prescription drugs and identify patients who may be misusing or abusing prescription opioids and may be in need of additional case management or other followup. The toolkit and accompanying code can also be used to answer research questions about opioid utilization.

OIG most recently analyzed opioid levels in Medicare Part D in a data brief entitled, *Opioid Use Decreased in Medicare Part D, While Medication-Assisted Treatment Increased* (OEI-02-19-00390). The data brief identified almost 49,000 Part D beneficiaries who were at serious risk of misuse or overdose. Some of these beneficiaries received extreme amounts of opioids. Others appeared to be “doctor shopping”—i.e., receiving high amounts of opioids from multiple prescribers and multiple pharmacies. The analysis identified beneficiaries who are at risk by calculating their opioid levels using Part D prescription drug data.

## **What does the toolkit include?**

This toolkit provides steps to calculate patients' average daily morphine equivalent dose (MED), which converts various prescription opioids and strengths into one standard value. This measure is also called morphine milligram equivalent (MME). The toolkit includes a detailed description of the analysis and programming code in two different programming languages (R and SQL) that can be applied to the user's own data. The R code and the SQL code provide the same data. These data can be used to identify certain patients who are at risk of opioid misuse or overdose. Users can also modify the code to meet their needs, such as identifying patients at varying levels of risk.

The additional code for the toolkit includes the following four chapters:

- 1. Analysis of Prescription Drug Claims Data**
- 2. Explanation of Programming Code To Conduct the Analysis**
- 3. Programming Code for R**
- 4. Programming Code for SQL**

# BACKGROUND

The opioid crisis remains a public health emergency. In 2018, there were 46,802 opioid-related overdose deaths in the United States.<sup>1</sup> In addition to the risk of death from overdose, opioids carry a number of other health risks, including respiratory depression, confusion, increased drug tolerance, and physical dependence. Identifying patients who are at risk of misuse or overdose is integral to curbing this epidemic.

OIG has made fighting the opioid epidemic a top priority, forming a multidisciplinary team dedicated to protecting beneficiaries from prescription drug abuse and misuse. As a part of its efforts, OIG has conducted extensive data analysis, which is the basis of this toolkit. OIG developed this toolkit—and a previous toolkit—to assist our partners, such as Medicare Part D plan sponsors, private health plans, and Medicaid Fraud Control Units, with analyzing their own prescription drug claims data to help combat the opioid crisis.

Both this toolkit and the previously released *Toolkit: Using Data Analysis To Calculate Opioid Levels and Identify Patients At Risk of Misuse or Overdose* provide detailed steps for using data from prescription drug claims to analyze patients' opioid levels and identify certain patients who are at risk of opioid misuse or overdose.<sup>2</sup> The previous OIG toolkit includes SAS programming code, whereas this toolkit includes code in two additional programming languages, R and Structured Query Language (SQL). Both toolkits are based on the methodology that OIG developed for its extensive work on opioid use in Medicare Part D.

OIG's most recent analysis of opioid use in Medicare Part D found that nearly 3 in 10 Medicare beneficiaries received an opioid in 2018.<sup>3</sup> The analysis also identified almost 49,000 beneficiaries who were at serious risk of opioid misuse or overdose. This included about 40,000 beneficiaries who received extreme amounts of opioids for the entire year and almost 8,800 beneficiaries who appeared to be doctor shopping—i.e., they received high amounts of opioids from multiple prescribers and pharmacies. (See Exhibit 1.) In 2018, there were fewer Part D beneficiaries receiving high amounts of opioids and fewer beneficiaries at serious risk of opioid misuse

<sup>1</sup> CDC, *Drug Overdose Deaths in the United States—1999–2018*, <https://www.cdc.gov/nchs/data/databriefs/db356-h.pdf>. Accessed on January 30, 2020.

<sup>2</sup> OIG, *Toolkit: Using Data Analysis To Calculate Opioid Levels and Identify Patients At Risk of Misuse or Overdose* (OEI-02-17-00560), June 2018.

<sup>3</sup> OIG, *Opioid Use Decreased in Medicare Part D, While Medication-Assisted Treatment Increased* (OEI-02-1-00390), July 2019.

or overdose than in the previous 2 years. Despite this apparent progress, concerns remain.

Patterns like these raise concern and warrant further scrutiny. They may indicate that a patient is receiving poorly coordinated care and requires additional case management. They may also indicate that a patient is addicted to opioids and at risk of overdose. Alternatively, they may indicate that a patient is receiving medically unnecessary drugs, which could be diverted for resale.

For this analysis, we identified beneficiaries on the basis of each patient's average daily morphine equivalent dose (MED), which is also known as morphine milligram equivalent (MME). MED is a unit of measurement that converts various prescription opioids and strengths into one standard value. The Centers for Disease Control and Prevention (CDC) recommends that physicians consider a patient's MED level when prescribing opioids. Specifically, CDC recommends that when prescribing opioids to patients with chronic pain, prescribers use caution at any dosage; carefully consider any increases in daily dosages to 50 milligram (50 mg) MED or more; and avoid increasing daily dosages to 90 mg MED or more.<sup>4, 5</sup>

---

<sup>4</sup> The CDC guideline provides recommendations for prescribing opioids for chronic pain outside of cancer treatment, palliative care, and end-of-life care. The guideline recommends that prescribers avoid increasing opioids to an MED of greater than or equal to 90 mg a day or carefully justify the decision to increase to this level. See CDC, *CDC Guideline for Prescribing Opioids for Chronic Pain—United States, 2016*, March 18, 2016, <https://www.cdc.gov/mmwr/volumes/65/rr/rr6501e1.htm>. Accessed on January 29, 2020. See also CDC, *Calculating Total Daily Dose of Opioids for Safer Dosages*, [https://www.cdc.gov/drugoverdose/pdf/calculating\\_total\\_daily\\_dose-a.pdf](https://www.cdc.gov/drugoverdose/pdf/calculating_total_daily_dose-a.pdf). Accessed on January 30, 2020.

<sup>5</sup> For patients who are already taking high dosages of opioids, the CDC guideline recommends that prescribers offer the patients the opportunity to re-evaluate their continued use of these dosages, and prescribers offer to work with patients to taper their opioids to safer dosages. The Department recently issued a guide for clinicians on how to appropriately reduce or discontinue long-term opioid use. For more information on tapering, see Department of Health and Human Services (HHS), *HHS Guide for Clinicians on the Appropriate Dosage Reduction or Discontinuation of Long-Term Opioid Analgesics*, September 2019, [https://www.hhs.gov/opioids/sites/default/files/2019-10/8-Page%20version\\_HHS%20Guidance%20for%20Dosage%20Reduction%20or%20Discontinuation%20of%20Opioids.pdf](https://www.hhs.gov/opioids/sites/default/files/2019-10/8-Page%20version_HHS%20Guidance%20for%20Dosage%20Reduction%20or%20Discontinuation%20of%20Opioids.pdf). Accessed on October 21, 2019.

## **Exhibit 1: Opioids in Medicare Part D, 2018**

### **Nearly 3 in 10 Medicare Part D beneficiaries received a prescription opioid, a decrease from the previous 2 years**

- In total, 13.4 million beneficiaries received opioids.

### **More than 350,000 beneficiaries received high amounts of opioids in 2018**

- Each of these beneficiaries had an average daily MED greater than 120 mg a day for at least 3 months and did not have cancer or a hospice stay in 2017.

### **Almost 49,000 beneficiaries are at serious risk of misuse or overdose, fewer than in the previous 2 years**

- About 40,000 beneficiaries received extreme amounts of opioids. Each had an average daily MED that exceeded 240 mg for the entire year. This amount is more than two-and-a-half times the dose that CDC recommends avoiding.
- Almost 8,800 beneficiaries appear to be doctor shopping. Each received a high amount of opioids (an average daily MED of 120 mg for at least 3 months) AND had four or more prescribers and four or more pharmacies during the year.
- None of these beneficiaries had cancer or were in hospice care in 2017.

Source: OIG, *Opioid Use Decreased in Medicare Part D, While Medication-Assisted Treatment Increased* (OEI-02-19-00390), July 2019.

Like OIG, CMS uses MED to identify Part D beneficiaries who are at high risk of overutilizing opioids and may be in need of case management.<sup>6</sup> This is the focus of CMS's Overutilization Monitoring System. CMS shares a list of these beneficiaries with Part D plan sponsors on a quarterly basis. Sponsors are instructed to conduct reviews of these beneficiaries to determine which beneficiaries require greater oversight or coordination of care.

This toolkit provides steps for calculating patients' average daily MED and gives a detailed description of the analysis as well as R and SQL programming code that can be applied to the user's own data. Users can modify the code to meet their needs, such as identifying patients at varying levels of risk. The toolkit includes the following four chapters: (1) Analysis of Prescription Drug Claims Data; (2) Explanation of the Programming Code To

<sup>6</sup> CMS uses the term "MME" to refer to MED. As noted earlier, they are equivalent.

Conduct the Analysis; (3) Programming Code for R; and (4) Programming Code for SQL.

### Standards

We created this toolkit in accordance with the *Quality Standards for Inspection and Evaluation* issued by the Council of the Inspectors General on Integrity and Efficiency.

# 1. ANALYSIS OF PRESCRIPTION DRUG DATA

This chapter provides an overview of the steps that are needed to conduct the analysis. The chapter contains the following sections.

## Sections

### 1.1 Collecting Prescription Drug Claims Data

### 1.2 Merging Data for MED Calculation

### 1.3 Conducting Quality Control Checks

### 1.4 Reviewing Patients With Certain Conditions

### 1.5 For SQL Users, Creating a Calendar Table

### 1.6 Identifying At-Risk Patients

## 1.1 Collecting Prescription Drug Claims Data

The first step is to collect the data needed for the analysis.

To calculate a patient's average daily MED over time, five data fields from the prescription drug claims are necessary:

- Unique Patient Identifier;
- Prescription Fill Date;
- Quantity Dispensed, or the number of units dispensed;
- Days Supply; and
- National Drug Code (NDC).

The Prescription Fill Date field must be formatted so that each calendar date is equal to one unit (e.g., YYYYMMDD). If the field uses another format (i.e., date time), make the necessary adjustments.

It is also valuable to combine the MED analysis with additional patient-level data to further identify patients and patterns of concern. Two additional fields that may be particularly helpful are Unique Prescriber Identifier and Unique Pharmacy Identifier.<sup>7</sup>

---

<sup>7</sup> The code in this toolkit includes an optional section that counts the number of pharmacies and the number of prescribers from which each patient received opioids. If unique identifiers for either the prescribers or pharmacies are not available, that portion of the code can be skipped.

Other data that may be helpful include patient-level medical data (e.g., data on diagnoses and medical services) and demographic data (e.g., date of birth and geographic location). These data can be used to conduct additional analyses to better understand the at-risk patient population. The data can also be used to identify—or to exclude from the analysis—beneficiaries who have cancer or are in hospice care.

Next, the timeframe for the analysis needs to be selected. A timeframe of at least 6 months is recommended.<sup>8</sup> In order to capture all prescriptions that a patient used during the timeframe, the analysis should include data for certain prescriptions dispensed before the timeframe. For example, for a timeframe of January 1, 2019, to December 31, 2019, we included prescriptions that have a Prescription Fill Date in 2018 but had at least 1 day of supply in 2019. To identify these prescriptions, calculate the last day of supply for each 2018 prescription using this calculation:

$$\textit{Last day of supply} = \textit{Prescription Fill Date} + \textit{Days supply} - 1$$

If the last day of supply was in 2019, we included this prescription in our analysis.

## 1.2 Merging Data for MED Calculation

The next step is to merge the prescription drug data with CDC's MME conversion data based on the NDC. CDC's data include two fields—MME conversion factor and strength per unit—that are needed for the MED calculation. In addition, if users have prescription data that include opioids and other types of prescription drugs, merging the data with CDC's data will identify which prescriptions are opioids.

The MME conversion factor reflects the potency of each opioid prescription drug. For example, hydrocodone has a conversion factor of 1, meaning that its potency is similar to that of morphine. Hydromorphone, on the other hand, has a conversion factor of 4, meaning that it is about 4 times more potent than morphine.

The "strength per unit" field contains a numeric value that indicates the strength of the opioid in each drug. For example, a 30-mg tablet of oxycodone hydrochloride has a strength per unit of 30. For combination drugs in this analysis, the strength per unit is the strength of the opioid ingredient. For example, what is called a "10-325 mg" tablet of oxycodone-acetaminophen contains 10 mg of oxycodone and 325 mg of acetaminophen, so the strength per unit is 10.

---

<sup>8</sup> The analysis is conducted over 90-day periods; a 6-month timeframe allows the user to detect patterns using a substantive amount of data.

CDC publishes these data on its website. The data are available for download in both SAS and Excel versions at <https://www.cdc.gov/drugoverdose/resources/data.html>.<sup>9</sup>

The Documentation tab in the CDC's Excel spreadsheet contains important information about the conversion factors. A few things to note:

- CDC updates the conversion factor data periodically. These updates add new NDCs and occasionally change some drugs' conversion factors.
- The CDC's spreadsheet excludes certain opioids, such as some cough and cold formulations; injectable opioids; intravenous opioids; and drugs not typically used in outpatient settings. As a result, claims for these opioids will not match with an MME conversion factor and will therefore need to be excluded from the MED analysis.
- The CDC's spreadsheet includes buprenorphine products; however, in the 2017 and 2018 versions, CDC removed the MME conversion factor for these drugs. Although buprenorphine is an opioid, a number of the products containing buprenorphine are indicated for the treatment of opioid use disorder rather than pain.<sup>10</sup> (Such treatment is known as medication-assisted treatment.) Depending on the objectives of their analysis, users should consider options for how to handle buprenorphine products.
- CDC's spreadsheet includes an MME conversion factor for methadone. Methadone is indicated both for pain and for treatment of opioid use disorder. Users should determine whether these drugs are included in their data and then assess whether to include them in their analysis.

CDC's spreadsheet also contains information (e.g., NDCs, drug names) for benzodiazepines, muscle relaxants, stimulants, and zolpidem.<sup>11</sup> These data may be helpful for additional analysis. Benzodiazepines, for example, are known potentiator drugs—i.e., drugs that increase euphoria when mixed with opioids and increase the risk of overdose.

---

<sup>9</sup> The files are called "Oral MMEs—SAS Data File" and "Oral MMEs—Excel Data File." They can be merged using the NDC.

<sup>10</sup> A diagnosis of opioid use disorder indicates that an individual has a problematic pattern of opioid use that leads to clinically significant impairment or distress. Opioid use disorder is sometimes referred to as "opioid addiction."

<sup>11</sup> Zolpidem is a sedative sold under the brand name Ambien.

### 1.3 Conducting Quality Control Checks

Next, it is important to conduct quality control checks on the prescription drug claims data. There are many ways to check for errors in the data. The most appropriate methods depend on the data. Two steps that may be helpful are to check for missing or zero values and to check for erroneous values that may impact the MED calculation.<sup>12</sup>

**Missing/Zero Values:** Missing or zero values may underestimate a patient's average daily MED. We addressed this issue by excluding prescription drug claims data that had missing or zero values for the following fields that affect the MED calculation: Unique Patient Identifier, Prescription Fill Date, Quantity Dispensed, and Days Supply.

**Erroneous Values:** Erroneous data may cause inaccuracies in the MED calculation. Users should assess their own data and determine the best way to identify erroneous values. One method is to calculate the ratio of the Quantity Dispensed field to the Days Supply field and identify outliers. Both of these fields impact the MED calculation.

Once these outlier values have been identified, it is important to review the data and determine the next steps. If the outliers appear to be erroneous values, one way to address this is to exclude them from the analysis. Excluding such values reduces the risk of inappropriately flagging some patients; however, excluding value may lead to underestimating the amounts of opioids taken by these patients.

### 1.4 Reviewing Patients With Certain Conditions

It is important to review patient diagnosis codes and medical claims data—if available—to identify patients with certain conditions. Patients at the end of their lives or those with certain conditions, such as active cancer-related pain, may have different patterns of opioid use. It may be helpful to exclude these types of patients from the analysis or to flag them for separate analysis.

For instance, we excluded patients who were receiving hospice care. To identify these patients, we looked for hospice claims in the timeframe. We also excluded patients who had diagnoses of cancer. In its Overutilization Monitoring System, CMS excludes patients whom it identifies as having hospice care or a diagnosis of cancer. The Pharmacy Quality Alliance provides ICD-10 codes that can be used to identify patients with cancer.<sup>13</sup> See Exhibit 2.

<sup>12</sup> OIG conducts a number of quality control checks. These two steps are not a complete list.

<sup>13</sup> ICD-10 diagnosis codes became effective in October 2015. For analysis of data prior to October 2015, ICD-9 diagnosis codes are needed.

## Exhibit 2: ICD-10 Codes That Can Be Used To Identify Patients With Cancer

|         |  |
|---------|--|
| C00–C14 | Malignant neoplasms of lip, oral cavity, and pharynx                             |
| C15–C26 | Malignant neoplasms of digestive organs  |
| C30–C39 | Malignant neoplasms of respiratory and intrathoracic organs                      |
| C40–C41 | Malignant neoplasms of bone and articular cartilage                              |
| C43     | Malignant melanoma of skin   |
| C4A     | Merkel cell carcinoma  |
| C45–C49 | Malignant neoplasms of mesothelial and soft tissue                               |
| C50     | Malignant neoplasms of breast  |
| C51–C58 | Malignant neoplasms of female genital organs                                     |
| C60–C63 | Malignant neoplasms of male genital organs                                       |
| C64–C68 | Malignant neoplasms of urinary tract   |
| C69–C72 | Malignant neoplasms of eye, brain, and other parts of central nervous system     |
| C73–C75 | Malignant neoplasms of thyroid and other endocrine glands                        |
| C7A     | Malignant neuroendocrine tumors  |
| C7B     | Secondary neuroendocrine tumors  |
| C76–C80 | Malignant neoplasms of ill-defined, other secondary and unspecified sites        |
| C81–C96 | Malignant neoplasms of lymphoid, hematopoietic and related tissue                |
| D37–D48 | Neoplasms of uncertain behavior, polycythemia vera and myelodysplastic syndromes |
| D49     | Neoplasms of unspecified behavior  |
| Q85.0   | Neurofibromatosis (nonmalignant)   |

Source: Pharmacy Quality Alliance ICD Code Value Sets, Cancer Exclusion, 2018.

### 1.5 For SQL Users, Creating a Calendar Table

SQL users need a calendar table to use the code. If the database does not have a standard calendar, create one using the steps below. R users do not need to take this step.

The calendar table can be created on the user's data system or created in an external application (such as in a spreadsheet) and then imported into the user's system. Before creating the table, first define the timeframe of the analysis. For example, a timeframe of January 1, 2019, to December 31, 2019, has 365 calendar dates. Then take the following steps.

- Create a new table named A\_CALENDAR.

- Create one field named CALENDAR\_DATE. The field format needs to use a date format, which has each calendar date equal to one unit (e.g., YYYYMMDD).
- Fill each record in the CALENDAR\_DATE field with one calendar date, starting with the first date of the timeframe and ending with the last date. For a timeframe of 2019, CALENDAR\_DATE will list each date from January 1, 2019, to December 31, 2019, and the table will have 365 records.

## 1.6 Identifying At-Risk Patients

The next step is to select the opioid levels for review. The resulting data can be used to identify certain patients who are at risk of opioid misuse or overdose. Users can also modify the code to meet their needs, such as identifying patients at varying levels of risk.

Depending on the purpose of the analysis, users may opt to identify patients at different opioid levels. For example, if the purpose of the analysis is to identify patients who would benefit from additional case management, the user may choose an MED level that is lower than if the purpose of the analysis were to identify patients to be investigated for possible drug diversion. A lower MED level may also be appropriate when identifying patients who are mixing opioids with potentiator drugs, such as benzodiazepines.

To identify at-risk patients in Medicare Part D, OIG uses the following three measures: (1) patients who received high amounts of opioids (i.e., an average daily MED greater than 120 mg) for any 90-day period during the year; (2) patients who received extreme amounts of opioids (average daily MED greater than 240 mg) for the entire year; and (3) patients who appear to be doctor shopping (i.e., those with high amounts of opioids (as described above) AND four or more prescribers and four or more pharmacies during the year). There may be other patients who are at risk but do not meet these criteria.

OIG's third measure uses criteria similar to those that CMS's Overutilization Monitoring System used to identify high-risk patients from 2013 to 2017. For those years, CMS considered patients to be "high-risk" if they had a daily MED exceeding 120 mg for at least 90 consecutive days and received

opioids from four or more prescribers and four or more pharmacies in 12 months.<sup>14</sup>

After identifying at-risk patients, it is important to conduct additional followup. For example, this followup may include contacting the patient's prescriber or conducting additional data analyses to determine the best course of action.

---

<sup>14</sup> CMS has modified its Overutilization Monitoring System criteria in recent years. For more information, see CMS, *Announcement of Calendar Year (CY) 2020 Medicare Advantage Capitation Rates and Medicare Advantage and Part D Payment Policies and Final Call Letter*, <https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Downloads/Announcement2020.pdf>. Accessed on January 13, 2019.

## 2. EXPLANATION OF PROGRAMMING CODE

This chapter explains the code. Either the R or SQL code can be used. Both follow similar logic and yield the same datasets and results.

The code creates three patient-level datasets:

- 1) Overview MED dataset, which provides an overview of each patient's opioid utilization;
- 2) Daily MED dataset, which provides detailed daily MED amounts for selected patients; and
- 3) At-Risk MED dataset, which identifies at-risk patients who had certain MED levels.

### Sections

#### 2.1 Creating an Overview of Each Patients' Opioid Utilization

#### 2.2 Creating Detailed Patient Data

#### 2.3 Identifying At-Risk Patients

### 2.1 Creating an Overview of Patients' Opioid Utilization

The code creates the "Overview MED" dataset. For each Unique Patient Identifier, it creates three fields:

1. Number of Days of Opioid Use (*OPIOID\_DAYS*),
2. Average Daily MED for the entire study timeframe (*ALL\_AVG\_MED*), and
3. Highest Average Daily MED for any 90 days (*MAX\_AVG\_MED\_90*).

First, the code calculates each patient's "Daily MED" for each day within the timeframe. The Daily MED includes all opioid prescriptions a patient received on a given day. The code creates a field for each day in the study timeframe and begins by assigning a value of zero to each day.

For each of the patient's prescriptions, the code calculates the MED. The MED converts opioids of different ingredients, strengths, and forms into equivalent milligrams of morphine; this is calculated using the following equation (Exhibit 3).

### Exhibit 3: MED Calculation for a Prescription

$$MED = \frac{(Strength\ per\ unit) \times (Quantity\ dispensed) \times (MME\ conversion\ factor)}{(Days\ supply)}$$

The code assigns the MED to the appropriate day based on the Prescription Fill Date and Days Supply.<sup>15</sup> If the day already has an MED value greater than zero, the code sums the two MED values.

Second, the code calculates the days of opioid use by counting the number of days during the timeframe that the patient has a daily MED that is greater than zero.

Third, the code calculates the average daily MED over the entire timeframe for the patient. See Exhibit 4.

### Exhibit 4: Calculation for Patient's Average Daily MED

$$Average\ daily\ MED = \frac{Total\ MED\ of\ all\ prescriptions\ in\ the\ timeframe}{Total\ number\ of\ days\ in\ the\ timeframe}$$

Fourth, the code calculates the average daily MED for the patient over every 90-day time period. The denominator for this calculation is 90, which is the number of days in the timeframe. (The denominator is *not* the number of days within the 90-day period on which the patient received opioids.) The macro then outputs the highest average of any 90-day period. See Exhibit 5.

### Exhibit 5: Determining Each Patient's Highest Average Daily MED for 90 Days

1. For the first 90-day period, the average daily MED over 90 days is:

$$90\text{-Day\ Average\ MED} = \frac{D1 + D2 + D3 + \dots + D88 + D89 + D90}{90}$$

2. For the second 90-day period, the average daily MED over 90 days is:

$$90\text{-Day\ Average\ MED} = \frac{D2 + D3 + D4 + \dots + D89 + D90 + D91}{90}$$

3. This repeats until the last possible 90-day combination. Then, the highest average for any 90 days outputs as *MAX\_AVG\_MED\_90*.

<sup>15</sup> The MED calculation assumes that the patient took the full prescription as prescribed each day starting on the Prescription Fill Date until the last day of the supply.

## 2.2 Creating Detailed Daily Patient Data

The code also creates the Daily MED dataset. This dataset provides the daily MED for each day in the study timeframe for selected patients. This code can be used to look more closely at the daily MED for selected patients with concerning patterns. For example, if a patient has an average daily MED that is extremely high, this data will show on which days during the timeframe the patient received the highest amounts. The code can also be used to identify the drugs associated with those days. For an analysis with 365 days, the data would contain the daily MED for each day. See Exhibit 6.

### Exhibit 6: Example of Daily MED Output

| Patient ID | Calendar Date     | MED |
|------------|-------------------|-----|
| X12345     | January 1, 2019   | 525 |
| X12345     | January 2, 2019   | 525 |
| X12345     | January 3, 2019   | 525 |
| X12345     | ...               | ... |
| X12345     | December 30, 2019 | 495 |
| X12345     | December 31, 2019 | 495 |

## 2.3 Identifying At-Risk Patients

Using the data created with the step above, the next section of the code identifies patients who—based on their average daily MED—are at risk of opioid misuse or overdose.

The code identifies patients who met the criteria for the measures used in OIG's data brief *Opioid Use Decreased in Medicare Part D, While Medication-Assisted Treatment Increased* (OEI-02-19-00390). These patients include the following.

- **Patients who received high amounts of opioids:** Patients with an average daily MED greater than 120 mg for any 90-day period and had at least 90 days of opioid use (*HIGH\_MED*).
- **Patients who received extreme amounts of opioids:** Patients with an average daily MED that was greater than 240 mg for the entire year and had at least 360 days of opioid use (*EXTREME\_MED*).
- **Patients who appear to be doctor shopping:** Patients with high amounts of opioids—i.e., an average daily MED greater than 120 mg for any 90-day period—and who received opioid prescriptions from four or

more prescribers AND four or more pharmacies during the year  
(*DOCTORSHOP\_MED*).<sup>16</sup>

There may be other patients who are at risk but do not meet these criteria. This section of the code can be adjusted depending on each user's needs. For example, if the purpose of the analysis is to identify patients who may benefit from additional case management, the MED levels can be lowered. If the purpose is to refer incidents of possible drug diversion to law enforcement, the MED levels can be increased. Similarly, if the user's analysis is based on a timeframe shorter or longer than 1 year, the days of opioid use can be changed.

---

<sup>16</sup> To identify these patients, the code includes an optional section that calculates the total number of prescribers and pharmacies that dispensed opioids to each patient during the timeframe.

## 3. PROGRAMMING CODE FOR R

This R code can also be downloaded at <https://oig.hhs.gov/oei/reports/oei-02-17-00560.asp>.

```
#####  
# U.S. Department of Health and Human Services  
# Office of Inspector General (OIG)  
# Office of Evaluation and Inspections  
  
# Toolkit for Calculating Opioid Levels and Identifying Patients At Risk of Misuse  
# or Overdose: R and SQL Programming Code  
# OEI-02-17-00561, May 2020  
  
# This code uses prescription drug data to analyze patients' opioid levels and  
# identify patients who are at risk of opioid misuse or overdose. It is based on  
# the methodology OIG developed for its extensive work on opioid use in Medicare  
# Part D, including its recent data brief entitled, Opioid Use Decreased in  
# Medicare Part D, While Medication-Assisted Treatment Increased (OEI-02-19-00390).  
  
# This code is being shared as part of the OIG's Toolkit for Calculating Opioid  
# Levels and Identifying Patients At Risk of Misuse or Overdose: R or SQL  
# Programming Code (OEI-02-17-00561). Please read the full toolkit. It provides  
# important information about preparing the data for the analysis, including how to  
# merge data from the Centers for Disease Control and Prevention. The toolkit may  
# be accessed at www.oig.hhs.gov/oei/reports/oei-02-17-00560.asp.  
  
# OIG has also released versions of this code using Structured Query Language (SQL)  
# and SAS. For the code, see www.oig.hhs.gov/oei/reports/oei-02-17-00560.asp.  
  
# The code calculates patients' average daily morphine equivalent dose (MED), which  
# converts various prescription opioids and strengths into one standard value:  
# equivalent milligrams of morphine. This measure is also called morphine  
# milligram equivalent (MME). Users can use these data to identify patients at  
# varying levels of risk.  
  
# The code creates three patient-level datasets:  
# 1) OVERVIEW_MED dataset provides an overview of each patient's opioid  
# utilization.  
# 2) DAILY_MED dataset provides detailed daily MED amounts for selected patients.  
# 3) AT_RISK_MED dataset identifies at-risk patients who had certain MED levels.  
  
# Please have this information ready to enter into the code below:  
# • The file directory of the data and the name of the dataset.  
# • The number of days in the timeframe of the analysis, the first date of the  
# timeframe, the date format, and the number of days in the timeframe. For  
# example, a timeframe of calendar year 2019 begins on January 1, 2019, uses  
# %d%m%Y date format, and has 365 days in the timeframe.  
# • The field names in the data for: prescription fill date, patient ID,  
# quantity dispensed, days supply, strength per unit, morphine milligram  
# equivalent conversion factor, pharmacy ID (optional), and prescriber ID  
# (optional).  
# • The number of days in the period of MED analysis. The default is set at  
# 3 months, or 90 days.
```

```

#####
### BEFORE RUNNING CODE INSTALL REQUIRED R PACKAGES

libs <- c("readr","dplyr","tictoc","purrr","RcppRoll","data.table")
new.libs <- libs[!(libs %in% installed.packages()[,"Package"])]
if(length(new.libs)) install.packages(new.libs)
lapply(libs, require, character.only=TRUE)

## INSERT THE FILE DIRECTORY OF THE DATA HERE ##
setwd("LINK TO DATA")

#####
### 1) OVERVIEW_MED: AN OVERVIEW OF EACH PATIENT'S OPIOID UTILIZATION

# This section of the code provides an overview of each patient's opioid
# utilization. The output data contains the following fields for each patient:

# PATIENT_ID Unique patient identifier.
# OPIOID_DAYS Days of opioid use, or the number of days during the
# timeframe when the patient has a daily MED that is greater
# than zero.
# ALL_AVG_MED Average daily MED for the entire study timeframe. For
# example, for a timeframe of 1 year, this is the patient's
# average daily MED for that year.
# MAX_AVG_MED_90 Highest average daily MED for any 90 days. This is the
# maximum average daily MED over any 90-day period in the
# timeframe.

# Brief overview of the code:
# 1a. Establish function parameters.
# 1b. Establish global variables for use through the code.
# 1c. Read and prepare the prescription data.
# 1c1. Map parameter values to internal functional variables.
# 1c2. Calculate values:
# • RX_FILL_DATE: prescription fill date
# • RX_END_DATE: prescription end date is the last date of prescription.
# This is the prescription fill date + days supply - 1.
# • AVG_DAILY_MED: average daily MED
# • PT0: the relative number of days before or after the timeframe start
# date that a prescription was filled plus one. Negative values
# indicate that the prescriptions were filled before the first day of
# the timeframe.
# • PT1: the relative number of days before or after the timeframe start
# date that a prescription ended. Negative values indicate that the
# prescriptions ended before the first day of the timeframe.
# 1c3. Keep prescriptions within the timeframe.
# 1c4. Set day indexes for the day of the timeframe that a prescription starts
# and ends.
# • T0_IDX: the day index within the timeframe a prescription starts.
# Prescriptions starting before the first day of the timeframe are
# assigned a day index of 1.
# • T1_IDX: the day index within the timeframe a prescription ends.
# Prescriptions ending after the last day of the timeframe are
# assigned a day index equivalent to the total days in timeframe.
# 1d. Create a list of dataframes. Each dataframe is specific to a single
Toolkit for Calculating Opioid Levels and Identifying Patients At Risk of Misuse or Overdose: R and SQL
Programming Code, OEI-02-17-00561

```

```

#       patient ID. Each dataframe is ordered by the prescription date. Each
#       prescription contains the AVG_DAILY_MED, T0_, and T1_ day indexes.
# 1e. Initialize a vector (INIT_VEC) with an element length equal to the number
#     of days in timeframe with each element set to 0.
# 1f. Define an empty list (RESULTS_LIST) with a vector length equivalent to the
#     total number of patient IDs. This list will be populated with summarized
#     opioid metrics for each patient ID.
# 1g. Loop through each prescription. For each prescription in each dataframe in
#     the ANALYSIS_LIST:
#     • Define a parameter vector with the AVG_DAILY_MED and T1_/T0_ day indexes.
#     • Create a copy of INIT_VEC named MED_VEC. Insert the AVG_DAILY_MED into
#       the range bookended by T1_ and T0 day indexes.
#     • Sum the resulting list of vectors to get the AVG_DAILY_MED for each day
#       of the timeframe and store it in the vector (MED_VEC).
#     • Define variables for each patient ID and compute opioid variables.
#       Combine them into a vector inserted iteratively into RESULTS_LIST.
# 1h. Convert RESULTS_LIST to a rectangular dataframe (OVERVIEW_MED).
# 1i. Write the overview dataset to CSV file.

```

```
# 1a
```

```

overview_med <- function(data_cleaned_csv,
                        days_in_timeframe,
                        window,
                        first_date,
                        date_format,
                        patient_id,
                        prescriber_id,
                        pharmacy_id,
                        prescription_fill_date,
                        quantity_dispensed,
                        days_supply,
                        strength_per_unit,
                        mme_conversion_factor) {

```

```
# 1b
```

```

timeframe_start_date <- as.Date(first_date, format = date_format)
time_frame_end_date <- timeframe_start_date + days_in_timeframe - 1
first_date <- first_date
days_in_timeframe <- days_in_timeframe
date_format <- date_format

```

```
# 1c
```

```
prescriptions_df <- fread(paste0("./",data_cleaned_csv)) %>%
```

```
# 1c1
```

```

rename('pat_id' = !! patient_id,
       'rx_fill_date' = !! prescription_fill_date,
       'quantity' = !! quantity_dispensed,
       'days' = !! days_supply,
       'strength' = !! strength_per_unit,
       'mme_con_factor' = !! mme_conversion_factor,
       'presc_id' = !! prescriber_id,
       'pharm_id' = !! pharmacy_id) %>%

```

```
# 1c2
```

```

mutate(
  rx_fill_date = as.Date(rx_fill_date, format = date_format),
  rx_end_date = as.Date(rx_fill_date) + days - 1,
  avg_daily_med = ((strength * mme_con_factor * quantity) / days),

```

```

    pt0_idx = as.numeric(rx_fill_date - timeframe_start_date) + 1,
    pt1_idx = as.numeric(pt0_idx) + days - 1) %>%

# 1c3
filter(rx_end_date >= timeframe_start_date) %>%
filter(rx_fill_date <= time_frame_end_date) %>%

# 1c4
mutate(
  t0_idx = ifelse(rx_fill_date <= timeframe_start_date, 1, rx_fill_date -
    timeframe_start_date + 1),
  t1_idx = ifelse(rx_end_date > time_frame_end_date, days_in_timeframe,
    rx_end_date - timeframe_start_date + 1)) %>%

select(presc_id, pharm_id, pat_id, rx_fill_date, rx_end_date, days,
  avg_daily_med, pt0_idx, pt1_idx, t0_idx, t1_idx) %>%

arrange(t0_idx)

# 1d
analysis_list <- prescriptions_df %>%
  mutate(rx_rank = as.numeric(as.Date(rx_fill_date,
    format = date_format))) %>%

  group_by(pat_id) %>%
  arrange(rx_rank) %>%
  mutate(rx_idx = row_number()) %>%
  ungroup() %>%f
arrange(pat_id, rx_rank) %>%
  select(rx_idx, pat_id, avg_daily_med, t0_idx, t1_idx) %>%
  filter_all(all_vars(!is.infinite(.))) %>%
  split(.$pat_id)

# 1e
init_vec <- numeric(days_in_timeframe)

# 1f
results_list <- vector("list", length = length(analysis_list))

# 1g
for (i in seq_along(analysis_list)) {
  param_vecs <- pmap(analysis_list[[i]][c("avg_daily_med", "t0_idx", "t1_idx")],
    c, use.names = FALSE)
  med_vec <- param_vecs %>% map(function(x){
    init_vec[x[2]:x[3]] <- x[1]
    return(init_vec)
  }) %>% reduce(`+`)

  pat_id <- analysis_list[[i]][1,2]
  opioid_days <- length(med_vec[med_vec > 0])
  all_avg_med <- sum(med_vec) / days_in_timeframe
  max_avg_med <- max(roll_mean(med_vec, window))

  results_list[[i]] <- c(pat_id, opioid_days, all_avg_med, max_avg_med)
}

# 1h
overview_med_df <- do.call(rbind.data.frame, results_list) %>% set_names(
  "patient_id", "opioid_days", "all_avg_med", "max_avg_med")

```

```

# li
overview_med_df %>%
  set_names("patient_id", "opioid_days", paste0("all_avg_med_",
        days_in_timeframe), paste0("max_avg_med_", window)) %>%
  fwrite("./overview_med.csv", row.names = FALSE, col.names = TRUE)
}

## ENTER TIMEFRAME, DATASET NAMES, AND FIELD NAMES ##
#   Enter the timeframe, the name of the dataset that contains the prescription
#   drug data, and the field names, as described below.

tic("Overview MED: An Overview of Each Patient's Opioid Utilization")
overview_med(
  ## ENTER FIRST DATE OF THE TIMEFRAME ##
  #   For example, for a timeframe of calendar year 2019 (January 1,
  #   2019 to December 31, 2019), the first date is January 1, 2019.
  first_date = "01JAN2019",

  ## SPECIFY THE DATE FORMAT ##
  #   The date format needs to use a format where each calendar date
  #   is equal to 1. This is required for the calculations to work
  #   properly. If the field uses another format (i.e., date time),
  #   some manipulation of the data may be needed to make it a date
  #   format. Search 'R date formats' online for more information.
  date_format = "%d%b%Y",

  ## ENTER THE NUMBER OF DAYS IN THE TIMEFRAME ##
  #   For example, a timeframe of calendar year 2019 (January 1, 2019
  #   to December 31, 2019) has 365 days.
  days_in_timeframe = 365,

  ## ENTER NAME OF CLEANED DATASET OF OPIOID PRESCRIPTION DATA ##
  #   Please note that the code does not exclude cancer and hospice
  #   patients. Please see the toolkit text for more information
  #   about excluding these patients from the analysis.
  data_cleaned_csv = "data_cleaned.csv",

  ## ENTER PRESCRIPTION FILL DATE FIELD NAME ##
  #   Again, the date format needs to use a format where each
  #   calendar date is equal to 1. This is required for the
  #   calculations to work properly. If the field uses another
  #   format (i.e., date time), some manipulation of the data may
  #   be needed to make it a date format. Search 'R date formats'
  #   online for more information.
  prescription_fill_date = 'prescription_fill_date',

  ## ENTER UNIQUE PATIENT IDENTIFIER FIELD NAME ##
  #   Note: the analysis relies on each patient having a unique
  #   identifier. If the prescription drug data do not contain a
  #   unique identifier for each patient, additional steps should be
  #   taken to create one.
  patient_id = 'patient_id',

  ## ENTER QUANTITY DISPENSED FIELD NAME ##
  quantity_dispensed = 'quantity_dispensed',

  ## ENTER DAYS SUPPLY FIELD NAME ##
  days_supply = 'days_supply',

```

```

## ENTER STRENGTH PER UNIT FIELD NAME ##
strength_per_unit = 'strength_per_unit',

## ENTER MORPHINE MILLIGRAM EQUIVALENT (MME) CONVERSION FACTOR
# FIELD NAME ##
mme_conversion_factor = 'MME_Conversion_Factor',

## If available: ENTER PHARMACY UNIQUE IDENTIFIER FIELD NAME ##
# Note: This field is optional.
pharmacy_id = 'pharmacy_id',

## If available: ENTER PRESCRIBER UNIQUE IDENTIFIER FIELD NAME ##
# Note: This field is optional.
prescriber_id = 'prescriber_id',

## ENTER NUMBER OF DAYS IN PERIOD OF MED ANALYSIS ##
# The default is set at 3 months, or 90 days.
window = 90)

```

```
toc()
```

```

#####
### 2) DAILY_MED: DETAILED DAILY MED AMOUNTS FOR SELECTED PATIENTS

```

```

# This code provides the daily MED for each day in the study timeframe for selected
# patients. This code can be used to look more closely at the daily MED for
# selected patients with concerning patterns. For example, if a patient has an
# extremely high average daily MED, these data will show which days during the
# timeframe the patient received the highest amounts.

```

```

# The code works similar to OVERVIEW_MED. Brief overview of the code:
# 2a. Initialize a vector (MED_VEC) with an element length equal to the number of
# days in timeframe with each element set to 0.
# 2b. Subset ANALYSIS_LIST by select PATIENT_ID dataframe and loop through
# prescriptions:
# • Define a parameter vector with the AVG_DAILY_MED and T1_/T0_ day indexes.
# • Insert the AVG_DAILY_MED into the range bookended by T1_ and T0_ day
# indexes.
# • Sum the resulting list of vectors to get the AVG_DAILY_MED for each day
# of the timeframe and store it in the vector (MED_VEC).
# • Define variables for each patient ID and compute opioid variables.
# Combine them into a vector inserted iteratively into MED_VEC.
# 2c. Generate a dataframe with sequential rows for the date in timeframe and the
# corresponding total average daily MED.
# 2d. Write the patient MED dataset to CSV file.

```

```

daily_med <- function(patient_id) {

  # 2a
  med_vec <- numeric(days_in_timeframe)

  # 2b
  param_vecs <- pmap(analysis_list[[patient_id]][c("avg_daily_med",
                                                    "t0_idx", "t1_idx")], c, use.names = FALSE)
  med_vec <- param_vecs %>% map(function(x){
    med_vec[x[2]:x[3]] <- x[1]
  })
}

```

```

    return(med_vec)
  }) %>% reduce(`+`)

# 2c
patient_med_df <- data.frame(patient_id = patient_id,
  date = seq(as.Date(first_date, format = date_format), as.Date(first_date,
    format = date_format) + (days_in_timeframe - 1), "days"), daily_med = med_vec)

# 2d
patient_med_df %>% fwrite(paste0("./patient_daily_med_",patient_id,".csv"),
  row.names = FALSE, col.names = TRUE)
}

```

```

tic("Daily MED: Detailed Daily MED Amounts for Selected Patients")
daily_med(patient_id = c(
  ## ENTER PATIENT ID HERE ##
  'PATIENT_ID'))
toc()

```

```

#####
### 3) AT_RISK_MED: AT-RISK PATIENTS WHO HAD CERTAIN MED LEVELS

```

```

# This code identifies patients who, based on their average daily MED, are at
# risk of opioid misuse or abuse. The code identifies patients who met the
# criteria for the measures used in a series of OIG data briefs, including Opioid
# Use Decreased in Medicare Part D, While Medication-Assisted Treatment Increased
# (OEI-02-19-00390).

```

```

# This section of the code can be adjusted depending on each user's needs. For
# example, if the purpose of the analysis is to identify patients who may
# benefit from additional case management, the MED thresholds can be lowered.
# If the purpose is to refer incidences of possible drug diversion to law
# enforcement, the MED thresholds can be increased. Similarly, if the user's
# analysis is based on a shorter or longer timeframe than 1 year, the days of
# opioid use can be changed.

```

```

# The code identifies patients who appear to be doctor shopping. This analysis
# requires unique pharmacy identifiers and unique prescriber identifiers in the
# opioid prescription data. For each patient, it counts the number of distinct
# pharmacies and prescribers that had at least 1 opioid prescription with a fill
# date in the timeframe. As noted below, this section of the code is optional
# and should be skipped if these unique identifiers are not available.

```

```

# The output dataset contains the following fields in addition to the fields
# from the OVERVIEW_MED dataset:

```

```

# HIGH_MED          Patients who received high amounts of opioids. Patients
#                   with an average daily MED greater than 120 mg for any
#                   90-day period and had at least 90 days of opioid use.
#                   1=Yes, 0=No.
# EXTREME_MED      Patients who received extreme amounts of opioids.
#                   Patients with an average daily MED that was greater than
#                   240 mg for the entire year and had at least 360 days of
#                   opioid use. 1=Yes, 0=No.
# DOCTORSHOP_MED   Patients who appear to be doctor shopping. Patients with
#                   a high amount of opioids (i.e., average daily MED
#                   greater than 120 mg for any 90-day period) and who

```

```

#           received opioid prescriptions from four or more
#           prescribers and four or more pharmacies during the year.
#           1=Yes, 0=No.

# Brief overview of the code:
# 3a. Count the number of distinct pharmacies and prescribers that had at least
#     1 opioid prescription with a fill date in the timeframe. Do not run this
#     if unique pharmacy identifier or unique patient identifier data are not
#     available.
# 3b. Join prescriber and pharmacy counts and apply conditions to flag patients.
#     If unique pharmacy identifier or unique patient identifier data are not
#     available, then remove join and the DR_SHOPPER parts.
# 3c. Write at-risk overview data to CSV file.

```

```

at_risk_med <- function(high_opioid_days,
                       extreme_opioid_days,
                       high_med_threshold,
                       extreme_med_threshold,
                       prescriber_count,
                       pharmacy_count) {

# 3a
count_pharms_presc <- prescriptions_df %>%
  filter(as.Date(rx_fill_date, format = date_format) >= as.Date(first_date,
                                                                format = date_format)) %>%
  filter(as.numeric(as.Date(rx_fill_date, format = date_format)) <=
         as.numeric(as.Date(rx_fill_date, format = date_format)) +
         days_in_timeframe - 1) %>%
  rename('patient_id' = pat_id) %>%
  group_by(patient_id) %>%
  mutate(prescriber_cnt = n_distinct(presc_id, na.rm = TRUE),
         pharmacy_cnt = n_distinct(pharm_id, na.rm = TRUE)) %>%
  select(patient_id, prescriber_cnt, pharmacy_cnt) %>%
  distinct()

# 3b
at_risk_overview <-<- overview_med %>%
  left_join(count_pharms_presc) %>%

  mutate(high_med = case_when((opioid_days >= high_opioid_days &
                               max_avg_med > high_med_threshold) ~ 1,
                              TRUE ~ 0),
         extreme_med = case_when((opioid_days >= extreme_opioid_days &
                                   all_avg_med > extreme_med_threshold) ~ 1,
                                  TRUE ~ 0),
         dr_shopper = case_when((high_med == 1 &
                                   pharmacy_cnt >= pharmacy_count &
                                   prescriber_cnt >= pharmacy_count) ~ 1,
                                  TRUE ~ 0)) %>%

  select(patient_id, opioid_days, all_avg_med, max_avg_med, pharmacy_cnt,
         prescriber_cnt, high_med, extreme_med, dr_shopper) %>%

  replace(., is.na(.), 0)

# 3c
at_risk_overview %>%
  set_names("patient_id", "opioid_days", paste0("all_avg_med_",
                                                days_in_timeframe), paste0("max_avg_med_", high_opioid_days),

```

```

        "pharmacy_count", "prescriber_count", "high_med", "extreme_med",
        "doctorshop_med") %>%
    fwrite("./at-risk-med.csv", row.names = FALSE, col.names = TRUE)
}

```

```

tic("At-Risk Patients: At-Risk Patients Who had Certain MED Levels ")
at_risk_med(

```

```

    ## The MED values (MED_THRESHOLDS), days of opioid use
    #   (OPIOID_DAYS), prescriber count, and pharmacy count may be
    #   adjusted as needed.
    high_med_threshold = 120,
    extreme_med_threshold = 240,
    high_opioid_days = 90,
    extreme_opioid_days = 360,
    prescriber_count = 4,
    pharmacy_count = 4)

```

```

toc()

```

```

### END OF CODE

```

```

#####
# Note that the Office of Inspector General (OIG) is providing this toolkit,
# including the associated programming code, to assist users in analyzing large
# datasets of prescription drug claims to identify individuals at risk of
# potential opioid abuse or misuse. This toolkit was prepared as a technical
# resource and is not intended to, and does not, create any rights, privileges,
# or benefits, substantive or procedural, enforceable by a party against the
# United States; its agencies or instrumentalities; its officers or employees; or
# any other person. The toolkit is provided in "as-is" condition, and OIG and
# its employees, agents, and staff disclaim any express or implied
# representation, warranty, or guarantee, including, but not limited to, the
# implied warranties of merchantability and fitness for a particular purpose.
# In particular, no representation is made that the information included in the
# toolkit, or any data the toolkit produces, is error free. The toolkit should
# not be used as the sole basis to determine whether an individual is abusing or
# overdosing on opioids or other prescription drugs, or in any determinations of
# coverage or dispensing by an insurer, pharmacy, provider, or other individual
# or organization. The toolkit is not intended to be used to determine
# compliance with any laws, regulations or other guidance. In no event shall
# OIG or its employees, agents, or staff be liable for any claim, damages, or
# liability, whether in an action of contract, tort or otherwise, and including
# direct, indirect, incidental, special, exemplary, or consequential damages,
# however caused, and on any theory of liability, arising in any way out of the
# use of this toolkit or its associated code, even if advised of the possibility
# of such damage. Compatibility of the toolkit with any user systems is not
# guaranteed, and any manipulation or alteration of the code is the sole
# responsibility of the user.
#####

```

## 4. PROGRAMMING CODE FOR SQL

This SQL code can also be downloaded at <https://oig.hhs.gov/oei/reports/oei-02-17-00560.asp>.

```
/*  
U.S. Department of Health and Human Services  
Office of Inspector General (OIG)  
Office of Evaluation and Inspections
```

```
Toolkit for Calculating Opioid Levels and Identifying Patients At Risk of Misuse or  
Overdose: R and SQL Programming Code  
OEI-02-17-00561, May 2020
```

This code uses prescription drug data to analyze patients' opioid levels and identify patients who are at risk of opioid misuse or overdose. It is based on the methodology OIG developed for its extensive work on opioid use in Medicare Part D, including its recent data brief entitled, Opioid Use Decreased in Medicare Part D, While Medication-Assisted Treatment Increased (OEI-02-19-00390).

This code is being shared as part of the OIG Toolkit: Calculating Opioid Levels and Identifying Patients At Risk of Misuse or Overdose Using R or SQL (OEI-02-17-00561). Please read the full toolkit. It provides important information about preparing the data for the analysis, including how to merge data from the Centers for Disease Control and Prevention. The toolkit may be accessed at [www.oig.hhs.gov/oei/reports/oei-02-17-00560.asp](http://www.oig.hhs.gov/oei/reports/oei-02-17-00560.asp).

OIG has also released versions of this code using R and SAS. For the code, see [www.oig.hhs.gov/oei/reports/oei-02-17-00560.asp](http://www.oig.hhs.gov/oei/reports/oei-02-17-00560.asp).

This code calculates patients' average daily morphine equivalent dose (MED), which converts various prescription opioids and strengths into one standard value: equivalent milligrams of morphine. This measure is also called morphine milligram equivalent (MME). Users can use these data to identify patients at varying levels of risk.

The code creates three patient-level tables:

- 1) OVERVIEW\_MED table provides an overview of each patient's opioid utilization.
- 2) DAILY\_MED table provides detailed daily MED amounts for selected patients.
- 3) AT\_RISK\_MED table identifies at-risk patients who had certain MED levels.

```
/*  
DATA PREPARATION
```

In addition to reading the toolkit and taking the necessary steps to prepare the data, be sure to complete these two additional steps that are necessary for the SQL code:

1. Rename the prescription drug data file and fields
2. Create a calendar table

```
1. RENAME THE PRESCRIPTION DRUG DATA FILE AND FIELDS
```

Toolkit for Calculating Opioid Levels and Identifying Patients At Risk of Misuse or Overdose: R and SQL Programming Code, OEI-02-17-00561

Before using this code, ensure that your prescription drug data and the code have the same file and field names. As needed, rename the prescription drug data so that the file and field names are as follows:

FILE NAME:

- DATA\_CLEANED: Ensure that the file name of the cleaned prescription drug data is DATA\_CLEANED. Note that the code does not exclude cancer and hospice patients. See the toolkit text for more information about excluding these patients from the analysis.

FIELD NAMES:

- PATIENT\_ID: Unique patient identifier field name. Note: The analysis relies on each patient having a unique identifier. If the prescription drug data do not contain a unique identifier for each patient, additional steps should be taken to create one.
- PRESCRIPTION\_FILL\_DATE: Prescription fill date field name. The field format needs to use a date format (e.g., YYYYMMDD), which has each calendar date equal to 1 unit. This is required for the calculations to work properly. If the field uses another format (i.e., datetime or timestamp), some manipulation of the data may be needed to make it into a date format.
- QUANTITY\_DISPENSED: Quantity dispensed field name.
- DAYS\_SUPPLY: Days supply field name.
- MME\_CONVERSION\_FACTOR: Morphine milligram equivalent (MME) conversion factor field name.
- STRENGTH\_PER\_UNIT: Strength per unit field name.
- PHARMACY\_ID: Pharmacy unique identifier field name. Note: This field is optional.
- PRESCRIBER\_ID: Prescriber unique identifier field name. Note: This field is optional.

## 2. CREATE A CALENDAR TABLE

The next step is to create a calendar table. The calendar table can be created on the user's data system or created in an external application (such as a spreadsheet) and then imported into the user's system.

Before creating the table, first define the timeframe of the analysis. For example, a timeframe of the 2019 calendar year, or January 1, 2019 to December 31, 2019, has 365 calendar dates. Then take the following steps:

- Create a new table named A\_CALENDAR.
- Create one field named CALENDAR\_DATE. The field format needs to use a date format, which has each calendar date equal to 1 unit (e.g., YYYYMMDD).
- Fill each record in the CALENDAR\_DATE field with one calendar date starting with the first date of the timeframe and ending with the last date. For example, for a timeframe of 2019, CALENDAR\_DATE will list each date from January 1, 2019 to December 31, 2019, and the table will have 365 records.

```
/*  
1) OVERVIEW_MED: AN OVERVIEW OF EACH PATIENT'S OPIOID UTILIZATION
```

This section of the code provides an overview of each patient's opioid utilization. The output table contains the following fields for each patient:

|             |  |
|-------------|--|
| PATIENT_ID  | Unique patient identifier.   |
| OPIOID_DAYS | Days of opioid use, or the number of days during the<br>timeframe when the patient has a daily MED that is greater<br>than zero. |

ALL\_AVG\_MED Average daily MED for the entire study timeframe. For example, for a timeframe of 1 year, this is the patient's average daily MED for that year.

MAX\_AVG\_MED\_90 Highest average daily MED for any 90 days. This is the maximum average daily MED over any 90-day period in the timeframe.

Brief overview of the code:

First, the code cross joins A\_CALENDAR to each patient ID. This step creates a table (called B\_LEDGER) of calendar dates in the timeframe of the analysis for each patient.

Second, the code calculates the daily MED and maps it to its calendar date for all of a patient's opioid prescriptions in a table called C\_DAY\_MED.

Third, the code joins the daily MED values from C\_DAY\_MED onto B\_LEDGER by patient ID and calendar date and sums the daily MED values for each date and creates a table called D\_DAILY\_MED. This table gives each patient's daily MED for each date in the timeframe.

Fourth, the code creates a table called E\_AVG\_DAILY\_MED when it inner joins D\_DAILY\_MED with itself to calculate the average daily MED for 90 days. The number of days in the period of MED analysis is set at 3 months, or 90 days. This number of days can be changed if needed.

Last, the code calculates all the overview fields listed above for each patient and creates the table OVERVIEW\_MED. \*/

```

create table b_ledger as
select  t1.PATIENT_ID,
        t2.calendar_date
from    (select distinct PATIENT_ID
        from    DATA_CLEANED) t1
        cross join a_calendar t2

create table c_day_med as
select  t1.PATIENT_ID,
        t2.calendar_date,
        (coalesce(t1.STRENGTH_PER_UNIT,0)
         * coalesce(t1.QUANTITY_DISPENSED,0)
         * coalesce(t1.MME_CONVERSION_FACTOR,0)) /
        case when t1.DAYS_SUPPLY = 0
        then NULL /* Use appropriate missing value here. */
        else t1.DAYS_SUPPLY
        end as day_med
from    DATA_CLEANED t1
        inner join a_calendar t2
        on t2.calendar_date between
        t1.PRESCRIPTION_FILL_DATE and
        t1.PRESCRIPTION_FILL_DATE + t1.DAYS_SUPPLY - 1

create table d_daily_med as
select  t1.PATIENT_ID,
        t1.calendar_date,
        sum(coalesce(t2.day_med,0)) as daily_med
from    b_ledger t1
        left outer join c_day_med t2
        on t1.PATIENT_ID = t2.PATIENT_ID
        and t1.calendar_date = t2.calendar_date
group by t1.PATIENT_ID, t1.calendar_date

create table e_avg_daily_med as

```

```

select  t1.PATIENT_ID,
        t1.calendar_date,
        t1.daily_med,
        sum(t2.daily_med)/90 as avg_med_90
from    d_daily_med t1
        inner join d_daily_med t2
        on t1.PATIENT_ID = t2.PATIENT_ID
        and t2.calendar_date between
           t1.calendar_date          /* This is set to a 90 day period. */
           and t1.calendar_date + 89 /* This can be changed. */
group by t1.PATIENT_ID, t1.calendar_date, t1.daily_med
order by t1.PATIENT_ID, t1.calendar_date

```

```

create table overview_med as
select  PATIENT_ID,
        sum(case when daily_med > 0 then 1 else 0 end) as opioid_days,
        avg(daily_med) as all_avg_med,
        max(avg_med_90) as max_avg_med_90
from    e_avg_daily_med
group by PATIENT_ID

```

```

/*****
2) DAILY_MED: DETAILED DAILY MED AMOUNTS FOR SELECTED PATIENTS

```

This code provides the daily MED for each day in the study timeframe for selected patients. This code can be used to look more closely at the daily MED for selected patients with concerning patterns. For example, if a patient has an extremely high average daily MED, these data will show which days during the timeframe the patient received the highest amounts.

When using this code, insert the patient ID(s) to output data for selected patient(s). \*/

```

create table daily_med as
select  PATIENT_ID,
        calendar_date,
        daily_med
from    e_avg_daily_med
where   PATIENT_ID = /* INSERT PATIENT ID HERE */

```

```

/*****
3) AT_RISK_MED: AT-RISK PATIENTS WHO HAD CERTAIN MED LEVELS

```

This code identifies patients who, based on their average daily MED, are at risk of opioid misuse or abuse. The code identifies patients who met the criteria for the measures used in a series of OIG data briefs, including Opioid Use Decreased in Medicare Part D, While Medication-Assisted Treatment Increased (OEI-02-19-00390).

This section of the code can be adjusted depending on each user's needs. For example, if the purpose of the analysis is to identify patients who may benefit from additional case management, the MED thresholds can be lowered. If the purpose is to refer incidences of possible drug diversion to law enforcement, the MED thresholds can be increased. Similarly, if the user's analysis is based on a shorter or longer timeframe than 1 year, the days of opioid use can be changed.

The code identifies patients who appear to be doctor shopping. This analysis requires unique pharmacy identifiers and unique prescriber identifiers in the opioid prescription data. For each patient, it counts the number of distinct pharmacies and prescribers that had at least 1 opioid prescription with a fill date in the timeframe. As noted below, this section of the code is optional and should be skipped if these unique identifiers are not available.

The output table contains the following fields in addition to the fields from the OVERVIEW\_MED table:

|                |   |
|----------------|---|
| HIGH_MED       | Patients who received high amounts of opioids. Patients with an average daily MED greater than 120 mg for any 90-day period and had at least 90 days of opioid use. 1=Yes, 0=No.  |
| EXTREME_MED    | Patients who received extreme amounts of opioids. Patients with an average daily MED that was greater than 240 mg for the entire year and had at least 360 days of opioid use. 1=Yes, 0=No.   |
| DOCTORSHOP_MED | Patients who appear to be doctor shopping. Patients with a high amount of opioids (i.e., average daily MED greater than 120 mg for any 90-day period) and who received opioid prescriptions from four or more prescribers and four or more pharmacies during the year. 1=Yes, 0=No. |

If unique pharmacy identifiers or unique patient identifiers are not available, do not run the PHARM\_PRESC\_COUNTS step and remove pharmacy and prescriber counts from AT\_RISK\_MED. \*/

```

create table pharm_presc_counts as
select  t1.PATIENT_ID,
        count(distinct t1.PHARMACY_ID) as pharmacy_count,
        count(distinct t1.PRESCRIBER_ID) as prescriber_count
from    DATA_CLEANED t1
        inner join a_calendar t2
        on t1.PRESCRIPTION_FILL_DATE = t2.calendar_date
group by PATIENT_ID
order by PATIENT_ID

create table at_risk_med as
select  t1.PATIENT_ID,
        t1.opioid_days,
        t1.all_avg_med,
        t1.max_avg_med_90,
        coalesce(t2.pharmacy_count,0) as pharmacy_count,
        coalesce(t2.prescriber_count,0) as prescriber_count,
        case when t1.max_avg_med_90 > 120 and t1.opioid_days >= 90 then 1
             else 0 end as high_med,
        case when t1.all_avg_med > 240 and t1.opioid_days >= 360 then 1
             else 0 end as extreme_med,
        case when t1.max_avg_med_90 > 120 and t1.opioid_days >= 90 and
             t2.pharmacy_count >= 4 and t2.prescriber_count >= 4 then 1
             else 0 end as doctorshop_med
from    overview_med t1
        left join pharm_presc_counts t2
        on t1.PATIENT_ID = t2.PATIENT_ID
        order by PATIENT_ID

/* END OF CODE */

```

```
/*  
Note that the Office of Inspector General (OIG) is providing this toolkit,  
including the associated programming code, to assist users in analyzing large  
datasets of prescription drug claims to identify individuals at risk of potential  
opioid abuse or misuse. This toolkit was prepared as a technical resource and is  
not intended to, and does not, create any rights, privileges, or benefits,  
substantive or procedural, enforceable by a party against the United States; its  
agencies or instrumentalities; its officers or employees; or any other person. The  
toolkit is provided in "as-is" condition, and OIG and its employees, agents, and  
staff disclaim any express or implied representation, warranty, or guarantee,  
including, but not limited to, the implied warranties of merchantability and  
fitness for a particular purpose. In particular, no representation is made that  
the information included in the toolkit, or any data the toolkit produces, is error  
free. The toolkit should not be used as the sole basis to determine whether an  
individual is abusing or overdosing on opioids or other prescription drugs, or in  
any determinations of coverage or dispensing by an insurer, pharmacy, provider, or  
other individual or organization. The toolkit is not intended to be used to  
determine compliance with any laws, regulations or other guidance. In no event  
shall OIG or its employees, agents, or staff be liable for any claim, damages, or  
liability, whether in an action of contract, tort or otherwise, and including  
direct, indirect, incidental, special, exemplary, or consequential damages, however  
caused, and on any theory of liability, arising in any way out of the use of this  
toolkit or its associated code, even if advised of the possibility of such damage.  
Compatibility of the toolkit with any user systems is not guaranteed, and any  
manipulation or alteration of the code is the sole responsibility of the user. */
```

---

# ACKNOWLEDGMENTS

Miriam Anderson served as the team leader for this study. Others in the Office of Evaluation and Inspections (OEI) who conducted the study include Margaret Himmelright and Jason Kwong. OEI staff who provided support include Althea Hosein, Christine Moritz, Adam Freeman, and Michael Novello. We would also like to acknowledge the contributions of other Office of Inspector General (OIG) staff, including Jason Lee, Scott Hutchinson, and Robert Gibbons.

This report was prepared under the direction of Jodi Nudelman, Regional Inspector General for Evaluation and Inspections in the New York regional office; Nancy Harrison, Deputy Regional Inspector General; and Meredith Seife, Deputy Regional Inspector General.

---

# ABOUT THE OFFICE OF INSPECTOR GENERAL

The mission of the Office of Inspector General (OIG), as mandated by Public Law 95-452, as amended, is to protect the integrity of the Department of Health and Human Services (HHS) programs, as well as the health and welfare of beneficiaries served by those programs. This statutory mission is carried out through a nationwide network of audits, investigations, and inspections conducted by the following operating components:

## **Office of Audit Services**

The Office of Audit Services (OAS) provides auditing services for HHS, either by conducting audits with its own audit resources or by overseeing audit work done by others. Audits examine the performance of HHS programs and/or its grantees and contractors in carrying out their respective responsibilities and are intended to provide independent assessments of HHS programs and operations. These assessments help reduce waste, abuse, and mismanagement and promote economy and efficiency throughout HHS.

## **Office of Evaluation and Inspections**

The Office of Evaluation and Inspections (OEI) conducts national evaluations to provide HHS, Congress, and the public with timely, useful, and reliable information on significant issues. These evaluations focus on preventing fraud, waste, or abuse and promoting economy, efficiency, and effectiveness of departmental programs. To promote impact, OEI reports also present practical recommendations for improving program operations.

## **Office of Investigations**

The Office of Investigations (OI) conducts criminal, civil, and administrative investigations of fraud and misconduct related to HHS programs, operations, and beneficiaries. With investigators working in all 50 States and the District of Columbia, OI utilizes its resources by actively coordinating with the Department of Justice and other Federal, State, and local law enforcement authorities. The investigative efforts of OI often lead to criminal convictions, administrative sanctions, and/or civil monetary penalties.

## **Office of Counsel to the Inspector General**

The Office of Counsel to the Inspector General (OCIG) provides general legal services to OIG, rendering advice and opinions on HHS programs and operations and providing all legal support for OIG's internal operations. OCIG represents OIG in all civil and administrative fraud and abuse cases involving HHS programs, including False Claims Act, program exclusion, and civil monetary penalty cases. In connection with these cases, OCIG also negotiates and monitors corporate integrity agreements. OCIG renders advisory opinions, issues compliance program guidance, publishes fraud alerts, and provides other guidance to the health care industry concerning the anti-kickback statute and other OIG enforcement authorities.