RADARS® System
777 Bannock Street, MC0180
Denver, Colorado 80204
(303) 389-1240
www.radars.org

11 September 2017

Response to Docket No. FDA-2017-N-2903
Data and Methods for Evaluating the Impact of Opioid Formulations with Properties Designed to Deter Abuse in the Postmarket Setting

The intent of this document is to offer comment on the United States Food and Drug Administration's (FDA) Issues Paper as well as the discussion at the public workshop held July 10-11, 2017.

1. Key Points Summarized

- The RADARS System recognizes that monitoring prescription opioid abuse and misuse is a challenge and that a mosaic approach is required to address the complexities of multiple outcome measures, the study of aberrant behaviors, and monitoring a population beyond the intended patient.
- Just as no randomized controlled trial is perfect, no surveillance data source is perfect. All current data sources come with strengths and limitations. These limitations should be minimized and further studied to understand potential bias and effect of confounders, while recognizing the overall value of the data.
- Importantly, the value of understanding the relation between various data sources, the *mosaic* design, should be better defined with establishment of a systematic and consistent approach to interpretation of the results.
- Choice of denominator when calculating rates is critical and should be aligned with the research question and types of comparator products. Thus no one denominator may be ideal for every analysis.

In relation to specific data sources:

RADARS System Poison Center Program

- The volume of adult exposures to pharmaceutical products reported to poison centers nationwide has been stable for the years during which opioids with abuse deterrent properties have come to market. Concerns have been expressed about declining volume, but this decline is primarily in Information calls and does not impact the Exposure call volume for adults.
- External validation studies indicate good correlation when studying trends over time with the Drug Abuse Warning Network (DAWN) and the National Vital Statistics System (NVSS).
- Additional work is needed to understand who utilizes poison centers, if that user base is changing over time, and if changing, the potential impact on the data.
- Product ascertainment is important in the assessment of products with abuse deterrent properties and is regularly addressed in the RADARS System Poison Center Program.

RADARS System Treatment Center Programs

- Product ascertainment is important in the evaluation of products with abuse deterrent properties and is regularly addressed in RADARS System Treatment Center Programs.
- Statistical approaches to address known product-specific reporting limitations are described, as well as additional studies to assess accuracy of product selection and ways to enhance reporting.

2. On the capture of opioid abuse-related calls in the RADARS System Poison Center Program

In section 3.1 of the Issues Paper, FDA stated: "First, an unknown, and likely small, fraction of abuse and overdose events result in a call to a PCC. It is unclear what factors might influence whether an opioid abuse-related event generates a call, or how these factors might vary over time or across drugs."

RADARS System Comment: We agree with the FDA in that no in-depth studies have been published which describe callers' motivations for contacting poison centers in general, or specifically for opioid abuse. While actual exposures are undoubtedly underreported to poison centers nationwide, below we present data supporting that reporting biases specific to adult abuse exposures involving opioids have not likely changed during the period of study related to opioids with abuse deterrent properties. Hence, trends over time offer a valuable measure of the acute medical consequences of prescription opioid abuse.

As with nearly all other surveillance systems, complete case capture does not preclude the ability to understand trends over time, route, or population. By comparison, national mortality data are an exception to the surveillance paradigm in their completeness of overdose deaths, however route and product specificity are greatly limited, as well as timeliness of reporting. The other prominent examples of complete case counting are the National Program of Cancer Registries and the National Notifiable Disease Surveillance System (NNDSS). In all three of these instances, complete case reporting is underpinned by legal requirements to collect and report cases regardless of where they arose in the healthcare system (for a review see McLaughlin 2011). No similar national legal imperative exists nationwide for reporting overdose deaths. Without such a national directive, it would be nearly impossible to create a complete case overdose surveillance program with the required route and product specificity, and would also represent a considerable burden to medical examiners.

In light of these limitations, poison center data are the unique national data source that can be used to inform evaluations on the effectiveness of opioids with abuse deterrent properties. There are four major national data collection paradigms to identify acute poisoning cases for opioid analgesics: poison centers, death certificate data, emergency department (ED) visits in health insurance claims or syndromic surveillance systems (Slavova et al. 2014), and adverse event reports submitted to FDA, but the latter three all have route and product specificity limitations. We have undertaken multiple efforts over the last decade to compare poison center intentional exposure calls to emergency department visits and mortality data, as a measure of external validation.

While poison centers may not capture every overdose death, our approach to evaluation of opioid products with abuse deterrent properties relies on statistical modeling where complete case counts are not necessary for causal inference, similar to extrapolations of surveillance data

for infectious disease and other areas of epidemiology. We are undertaking further studies to mathematically adjust for *possible* changes in acute medical (overdose) reporting. Yet, there is little evidence to support (or refute) that mechanisms for overdose deaths to appear in poison center calls have changed over time. This is an area where further research may be useful.

External validation offers a possible path for understanding poison center data. FDA noted analyses by Davis et al. 2014 and Bau et al. 2016 that showed correlation between poison center abuse calls for opioid analgesics and emergency department visits. Trends over time have also been studied by comparing methadone exposures reported to the RADARS System Poison Center Program to methadone deaths reported to the National Center for Health Statistics (Dasgupta et al. 2012). The results indicate good correlation between these independent data sources. A more recent evaluation of opioid deaths reported to the RADARS System Poison Center Program and those reported to National Vital Statistics System (NVSS) suggest a similar correlation (Ronk et al. 2014).

- RADARS System Poison Center Program has been evaluated by comparing trends to those reported by the ED surveillance in the Drug Abuse Warning Network (DAWN) (Bau et al. 2016). Active pharmaceutical ingredient-specific intentional exposure population rates from the RADARS System Poison Center Program are highly correlated with non-medical use population rates from DAWN (R^2=0.84 to 0.99 for oxycodone, fentanyl, hydrocodone, hydromorphone, morphine, methadone, and buprenorphine), Figure 1.
- Dasgupta et al. 2012 found that compared to mortality data from the National Center of Health Statistics, RADARS System Poison Center Program methadone related exposures tended to involve younger patients, more often female, in the home and were less likely to require medical attention. A strong association was found with exposures in RADARS System Poison Center Program and methadone mortality (ß = 0.88, se = 0.42, t = 9.5, df = 1, p<0.0001, R^2 = 0.77). These findings were robust to changes in a sensitivity analysis assessing the impact of underreporting of methadone overdose deaths, Figure 2.
- Ronk et al. 2014 found that the RADARS System Poison Center Program mortality rates trended well with the CDC mortality rates (p=0.01). The Pearson correlation coefficient for this association was 0.80 and the adjusted R-squared value was 0.58, Figure 3.

FIGURE 1. CORRELATION BETWEEN RADARS SYSTEM POISON CENTER PROGRAM INTENTIONAL EXPOSURES AND DAWN EMERGENCY DEPARTMENT VISITS (BAU ET AL. 2016)

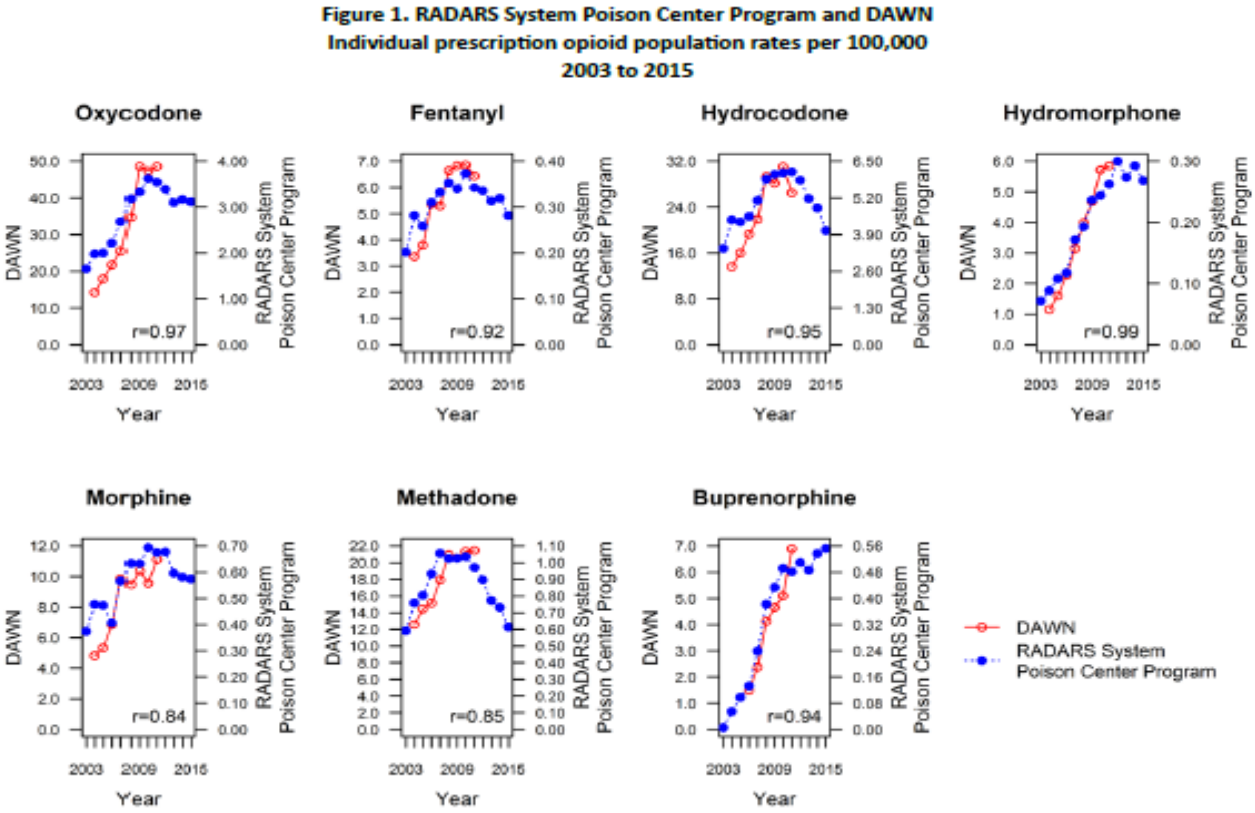


FIGURE 2. EXCERPT FROM DASGUPTA ET AL. 2012 ILLUSTRATING CORRELATION BETWEEN METHADONE-RELATED DEATHS AND METHADONE-RELATED INTENTIONAL AND UNINTENTIONAL EXPOSURES REPORTED TO THE RADARS SYSTEM POISON CENTER PROGRAM

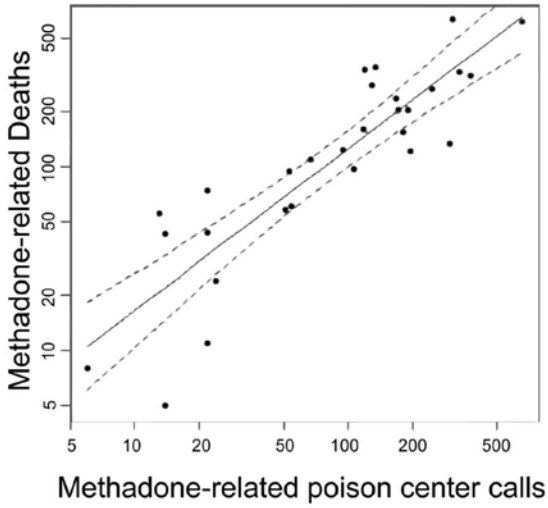
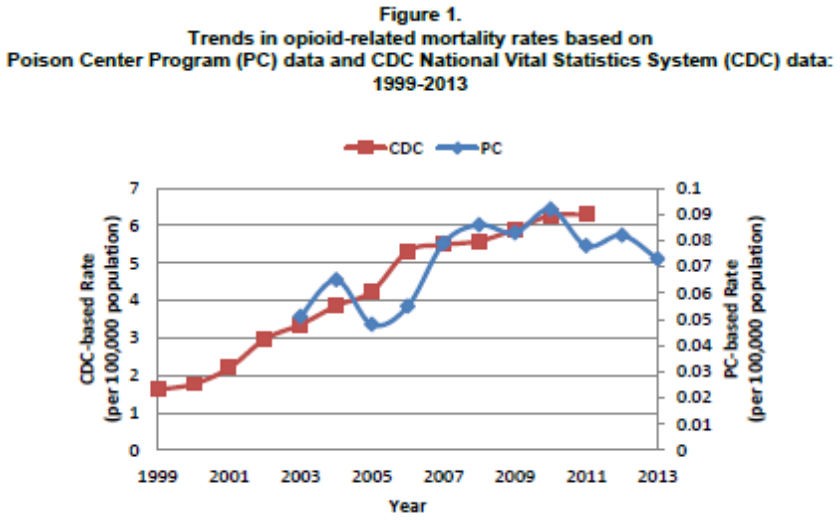


Figure 3. Poison center calls predict 77% of the variation in death counts from vital statistics data, 29 states, 2006–7. The plot shows the association between log-methadone death counts and log-methadone poison center calls; log scales are used for the sake of including all data on one graph since the state populations (and counts) vary widely. A strong association was found between poison center exposure calls and methadone-related mortality reported to vital statistics ($b = 0.88$, $SE = 0.42$, $t = 9.5$, $df = 1$, $p < 0.0001$, $R^2 = 0.77$), where a one unit increase of poison center calls corresponded to a 0.88 unit increase in methadone related deaths.

FIGURE 3. EXCERPT FROM RONK ET AL. 2014 ILLUSTRATING CORRELATION BETWEEN OPIOID-RELATED FATALITIES REPORTED TO RADARS SYSTEM POISON CENTER PROGRAM AND OVERDOSE DEATHS FROM NATIONAL VITAL STATISTICS SYSTEM (NVSS)



While these studies provide external validation, additional work to identify the factors that influence whether an opioid abuse-related event generates a call or how these factors might vary over time or across drugs is warranted.

Bau G, Bucher Bartelson B, Severtson SG, Green JL, Dart RC. (2016). Comparison of Population Rates between the Drug Abuse Warning Network (DAWN) and the RADARS® System Poison Center Program. RADARS System Technical Report, 2016-Q4. Available at: http://www.radars.org/Portals/1/TechReports/2016%204Q%20RADARS%20System%20QTR.pdf?ver=2017-01-06-123313-053. Accessed 11 September 2017.

Dasgupta N, Davis J, Jonsson Funk M, Dart R. Using poison center exposure calls to predict methadone poisoning deaths. *PLoS One*. 2012;7(7):e41181. doi: 10.1371/journal.pone.0041181. Epub 2012 Jul 19. PMID: 22829925

McLaughlin RH. Are cancer registries unconstitutional? *Soc Sci Med*. 2010 PMID: 20199835

Ronk C, Bucher Bartelson B, Severtson SG, Green JL, Dart RC (2014). Opioid-Related Mortality: National Vital Statistics System versus RADARS® Poison Center Program. RADARS® System Technical Report, 2014Q3. Available at: http://radars.org/Portals/1/TechReports/Third%20Quarter%202014%20Technical%20Report.pdf. Accessed 11 September 2017.

Slavova S, Bunn TL, Talbert J. Drug Overdose Surveillance Using Hospital Discharge Data. *Public Health Rep*. 2014 Sep-Oct; 129(5): 437–445. doi: 10.1177/003335491412900507 PMCID: PMC4116371

3. On Fluctuations in Call Volume to Poison Centers

- *In Section 3.1 of the Issues Paper, FDA stated: "Although there is some evidence that trends in call rates are correlated with trends in rates of emergency department visits involving misuse and abuse of prescription opioids (Davis et al., 2014, Bau et al., 2016), there is also evidence suggesting that patterns of PCC use have been changing in recent years (Mowry et al., 2015), further complicating the interpretation of analyses using PCC data with regard to making inferences about abuse trends in the population."*
- *In Section 4.5 of the Issues Paper, FDA stated: "comparison of the pre- and post-reformulation abuse rates for a product based on treatment center data may not be valid unless it can be assumed that the likelihood of an individual abusing this product interacting with treatment centers in the surveillance network (or in the case of PCC data an abuse or overdose event generating a call) would be the same in both time periods."*

RADARS System Comment: Poison center calls of interest are divided into two major categories, human Exposure calls and Information calls. In the former, an actual substance was ingested or a human was otherwise exposed (injection, inhalation/intranasal, dermal exposure, etc.). In the latter category, the caller only wanted information about a substance. From 2014 to 2015, the number of human exposure calls to poison centers increased slightly from 2,165,142 to 2,168,371 (see Mowry et al. 2016; Figure 4 provided below), after the absolute volume declined from 2009 to 2014.

FIGURE 4. EXCERPT FROM MOWRY ET AL. 2016 ILLUSTRATING POISON CENTER EXPOSURE AND INFORMATION CALL VOLUME OVER TIME

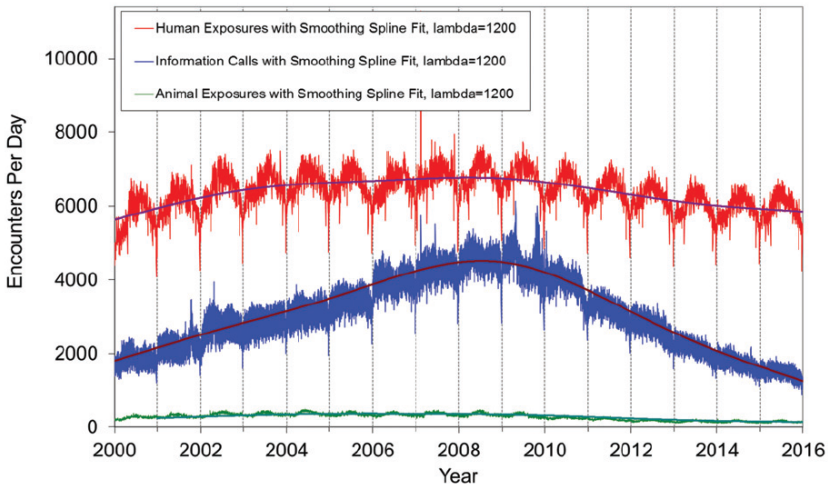


Figure 1. Human Exposure Cases, Information Calls and Animal Exposure Cases by Day since 1 January 2000
Smoothing Spline Fits using lambda =1200 for Human Exposures had associated RSqr =0.410, Information Calls RSqr =0.874 and Animal Exposures RSqr =0.841.

The decline between 2009 and 2014 can be explained by the following factors:

- The decline in call volume since 2009 is largely attributable to fewer information calls, as depicted in Figure 4 above; the number of human exposure calls has seen minimal change. Also in that paper, the specific type of Information call that has declined is drug identification calls, as people have turned to the Internet for this function.
- Human exposures for the sedative/hypnotics/antipsychotics and analgesics categories appear to have leveled off from 2010 to 2015 corresponding to prescription drug abuse trends in the National Survey on Drug Use and Health, while

calls involving other drugs (antidepressants, cardiovascular drugs) have continued to increase (Figure 5).

While the overall volume of adult exposures remains stable, further information about poison center utilization would be useful. Further work is warranted to understand who is likely to contact a poison center and if the geographic or demographic of these callers is changing.

Poison centers collect exposure calls for both pharmaceutical and non-pharmaceutical substances. Figure 6 illustrates the call volume from 2009 through 2015 published in the AAPCC NPDS Annual Reports stratified by pharmaceutical (e.g. prescription medications, over-the-counter medications) and non-pharmaceutical substances (e.g. household cleaners, pesticides) as well as by age (pediatric <12 years of age and adult ≥12 years of age). The most dramatic change has been in the pediatric population. The number of adult pharmaceutical exposures (the primary population studied for opioid abuse and misuse) has remained stable during this time, suggesting poison center utilization is not a significant confounder when studying opioid trends.

We propose three approaches to further explore the issues raised above. First, Wolowich et al. (2001) proposed a weighting scheme with age and penetrance (a metric of poison center utilization using call volume and total US population) to adjust for poison center call variations. This weighting method could address some of the concerns with fluctuating call volumes over time, and we are exploring it now. Second, a study of motivations for people who call poison centers could be envisioned, possibly via interviews with people who utilize poison centers. We are also exploring this type of study at the moment. Third, continued characterization and external validity to mortality data will inform how poison center data differ from overdose death data. This too is an ongoing area of research for us.

Mowry JB, Spyker DA, Brooks DE, McMillan N, Schauben JL. 2014 Annual Report of the American Association of Poison Control Centers' National Poison Data System (NPDS): 32nd Annual Report. *Clin Toxicol* (Phila). 2015;53(10):962-1147. doi: 10.3109/15563650.2015.1102927. PMID: 26624241

Mowry JB, Spyker DA, Brooks DE, Zimmerman A, Schauben JL. 2015 Annual Report of the American Association of Poison Control Centers' National Poison Data System (NPDS): 33rd Annual Report. *Clin Toxicol* (Phila). 2016 Dec;54(10):924-1109. doi: 10.1080/15563650.2016.1245421. PMID: 28004588

Wolowich WR, Casavant MJ, Fisher CA. Age-specific and age-adjusted penetrance as poison center outcome measures. *J Toxicol Clin Toxicol*. 2001;39(4):367-70. PMID: 11527231

FIGURE 5. EXCERPT FROM MOWRY ET AL. 2016 ILLUSTRATING INCREASING EXPOSURES BY DRUG CATEGORIES OVER TIME

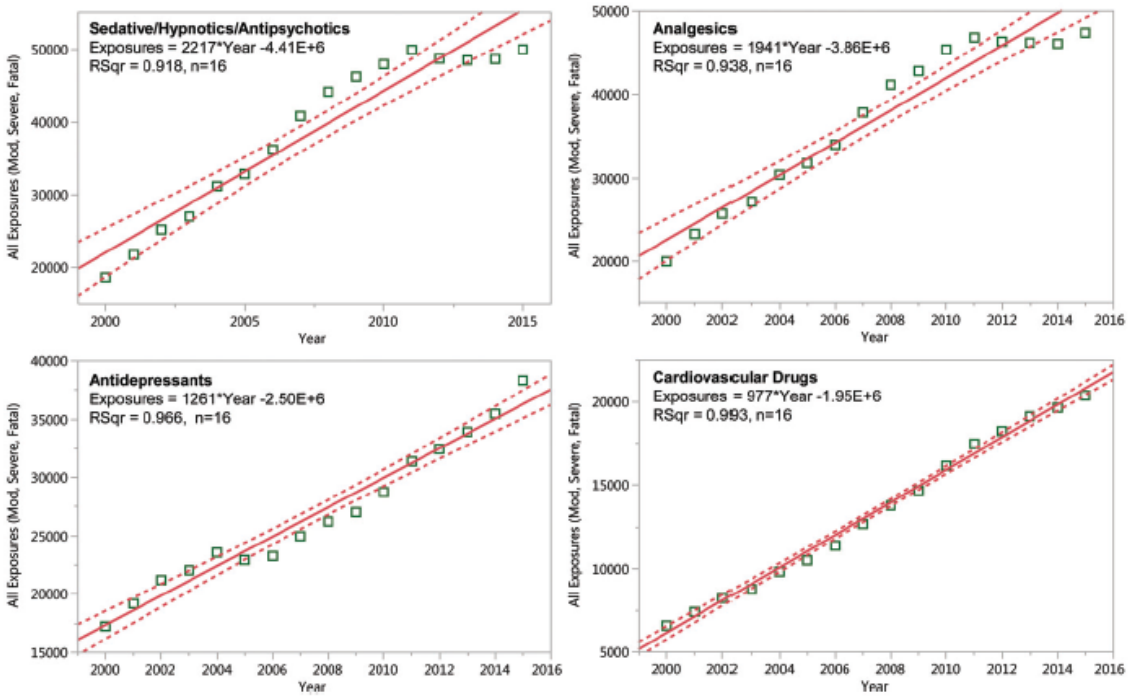
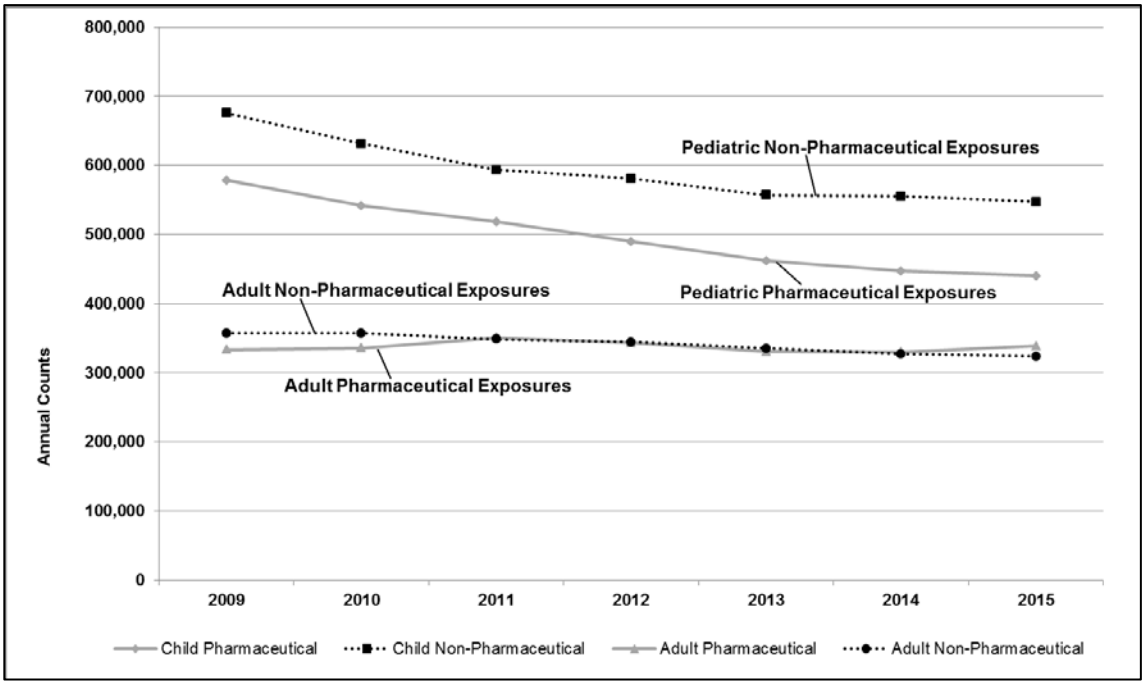


Figure 4. Substance Categories with the Greatest Rate of Exposure Increase since 1 January 2000 for More Severe Outcomes (Top 4)
Solid lines show least-squares linear regressions for the Human Exposure Cases per year for that category (□).
Broken lines show 95% confidence interval on the regression.

FIGURE 6. POISON CENTER EXPOSURE CALL VOLUME FOR ADULT/PEDIATRIC AND PHARMACEUTICAL/NON-PHARMACEUTICAL EXPOSURES 2009 THROUGH 2015



4. Deaths in Poison Center Calls

In Section 3.1 of the Issues Paper, FDA stated: "overdoses resulting in rapid, unattended death are unlikely to generate a call. Therefore, PCC may disproportionately fail to capture cases involving drugs with the highest risk of such fatal overdoses."

RADARS System Comment: It is true that exposures rapidly causing an unattended death are unlikely to result in contact with a poison center. It has been well known since inception of the program that all surveillance systems measuring deaths outside the health care system are by definition incomplete. Thus, incomplete ascertainment of deaths is a concern.

However, the reasonable correlation of deaths in the RADARS System Poison Center Program and the National Vital Statistics System (NVSS) suggests that the incomplete ascertainment is relatively consistent. In contrast to the coroner system, poison center data also include the specific products involved in the episode. For example, one can tell if a fentanyl exposure was to a prescription product or an illicit substance. Poison centers also detect serious medical outcomes that do not culminate in death (e.g., near misses). The trend in serious outcomes shows no diminution over the past 15 years (Figure 7).

Less serious human exposures declined between 2008 and 2014, while more serious exposures increased (Figure 7). This is in contrast to the statement in Section 3.1 of the Issue Paper that "PCC may disproportionately fail to capture cases involving drugs with the highest risk of such fatal overdoses." More than 70% of human exposures reported to poison centers appear to involve serious cases, including (but not exclusively) overdoses. While not all *fatal* overdoses are represented in poison center data, as previously stated, poison center data have considerable value for monitoring trends over time related to serious medical outcomes associated with opioid analgesics.

FIGURE 7. EXCERPT FROM MOWRY ET AL. 2016 ILLUSTRATING PROPORTION OF EXPOSURES RESULTING IN MORE SERIOUS VERSUS LESS SERIOUS MEDICAL OUTCOME

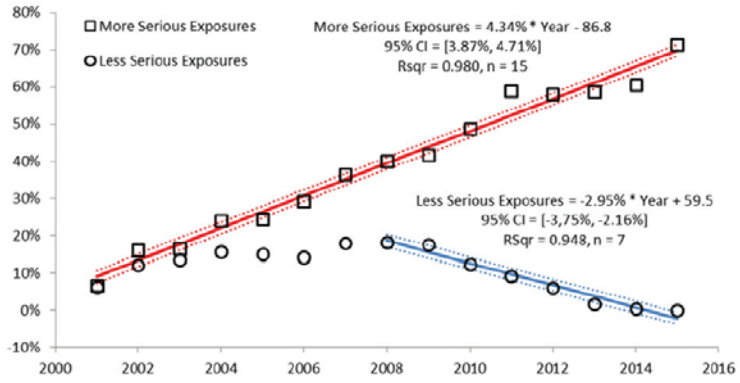


Figure 5. Change in Encounters by Outcome from Year 2000
The figure shows the percent change from baseline (year 2000) for Human Exposure Cases divided among the 10 Medical Outcomes. The More Serious Exposures (major, moderate and death) increased. The Less Serious Exposures (no effect, minor effect, not followed (non-toxic), not followed (minimal toxicity possible), unable to follow (potentially toxic) and unrelated effect) decreased after 2008. Solid lines show least-squares linear regressions for the change in More Serious Exposures per year (□) and Less Serious Exposures (○). Broken lines show 95% confidence intervals on the regression.

Mowry JB, Spyker DA, Brooks DE, Zimmerman A, Schauben JL. 2015 Annual Report of the American Association of Poison Control Centers' National Poison Data System (NPDS): 33rd Annual Report. *Clin Toxicol* (Phila). 2016 Dec;54(10):924-1109. doi: 10.1080/15563650.2016.1245421. PMID: 28004588

5. On Misclassification of Exposure

In Section 3.1 of the Issues Paper, FDA stated: "RADARS personnel then conduct additional quality checks on the call data, based on review of case narratives."

In Section 3.2 of the Issues Paper, FDA stated: "An important limitation of data collected from people entering or being assessed for substance use disorder treatment, as well as those calling PCCs, is the potential for various types of misclassification, including the specific product(s) being abused."

RADARS System Comment: We appreciate this point and the opportunity to elaborate on information available on poison center data accuracy as well as efforts to enhance product ascertainment.

5.1 RADARS System Poison Center Program

- Product identification is continuously addressed through ongoing education & trainings of Specialists in Poison Information (SPIs). Regular communications are sent highlighting products new to the market, changes to existing products, removal of products, and other topics illustrating the importance of accurate product ascertainment. Product identification is also an agenda item at all RADARS System Poison Center Program investigator meetings.
- As far back as 1999, efforts have been made to assess poison center data accuracy. For example, Hoyt et al. 1999 compared poison center calls with hospital chart reviews and found good concordance.
- A more recent study by Krenzelok et al. 2014 describes an evaluation of the accuracy of poison center data for acetaminophen-containing analgesics. While opioid products were not included in this study there are similarities in product ascertainment in that there are numerous acetaminophen-containing products available (both single-ingredient and combination products) and a brand effect in that Tylenol® is the most recognized name and often generics are misidentified. This study found that baseline accuracy for substance data fields was 89% (CI 88.20, 89.67), indicating strong accuracy even prior to product training of SPIs. After training this score rose to 93% (CI 92.24, 93.71), see Figure 8. While these data are comforting this study should be repeated for opioid-containing products (both pharmaceutical and illicitly manufactured products).

Hoyt B, Rasmussen R, Giffin S, Smiklstein MJ (1999). Poison Center Data Accuracy: A Comparison of Rural Hospital Chart Data with the TESS Database. Academic Emergency Medicine, 6(8), p. 851-5.

Krenzelok EP, Reynolds KM, Dart RC, Green JL. A model to improve the accuracy of US Poison Center data collection. *Clinical Toxicology*. 2014; 52:889-896.

FIGURE 8. EXCERPT FROM KRENZELOK ET AL. 2014 ILLUSTRATING STRONG DATA ACCURACY IN POISON CENTER RECORDS FOR ACETAMINOPHEN-CONTAINING PRODUCT EXPOSURES

Table 3. Mean quality component and total scores*, pre- and post-training.

Component	Pre-training score (95% CI)	Post-training score (95% CI)	p-value
Demographic fields	96.6 (96.09,97.08)	97.0 (96.54,97.53)	0.204
Exposure characteristic fields	95.3 (94.80,95.82)	95.6 (95.07,96.09)	0.462
Outcome fields	97.3 (96.78,97.76)	97.0 (96.56,97.54)	0.520
Substance fields	88.9 (88.20,89.67)	93.0 (92.24,93.71)	<0.001
Total score	94.0 (93.69,94.35)	95.3 (95.01,95.66)	<0.001

*Percentage of data correct for all expected data fields.

5.2 RADARS System Treatment Center Programs (Survey of Key Informants' Patients Program [SKIP] and Opioid Treatment Program [OTP])

We have taken a multi-step approach to enhance product ascertainment and evaluation specific to these paper-based survey programs. Based on our understanding, exposure misclassification in this setting arises primarily from **omission** (e.g., patient not recalling what was taken), **confusion** (e.g., patient mistakenly attributes what was taken, including synecdoche), and **careless responses** (e.g., on paper surveys, patients tick boxes haphazardly or endorse improbable numbers of products consumed in the past 30 days).

5.2.1 Omission

To address misclassification from **omission**, there is limited recourse. Visual identification cues have been suggested, but these too suffer from limitations. For example, because of limitations of resolution and screen size, a strong order-of-presentation effect can occur. Unless images are randomized, selection of the earlier products with pictures may be differentially induced. Given the dozens of generic formulations of differing strengths and appearances available (e.g., immediate-release [IR] hydrocodone), it is unrealistic to expect accurate visual differentiation between them. Still, the question of how important visual cues are remains unresolved. We are interested in exploring this issue further with a validation study.

5.2.2 Confusion

We have observed misclassification from **confusion** or lack of knowledge in treatment center survey data as well as in poison center exposures. In particular, endorsements of highly genericized opioid analgesics, such as extended-release (ER) morphine and IR hydrocodone, are often not formulation specific. The survey tools allow for a respondent to endorse the active ingredient even if they are unsure of the formulation. For example, a large percentage of morphine reports to poison centers are not classified as IR or ER resulting in a significant amount of missing data when studying rates of ER morphine, particularly in relation to the effectiveness of ER products with abuse deterrent properties. We have conducted extensive analysis into this phenomenon and feel that traditional missing data methods, such as multiple imputation, may offer a workable solution (described below).

In the RADARS System Poison Center and Treatment Center Programs, the proportion of cases for an active pharmaceutical ingredient (API) where the product and/or formulation is unknown exceeds traditionally ignorable amounts. These data are assumed to be missing at random or associated with observed variables such as age, gender, and time. Excluding these cases from product- or class-specific rate calculations will result in an underestimate of the true number of cases within each category. The inadequacy of single imputation methods and a need for a strategy to address missing values has been expressed extensively in epidemiology. Clinical trial designs have included multiple imputation in primary and secondary endpoints (see citations at end of this section). Therefore, we developed a method using multiple imputation (MI) to address missing product and formulation values, and submitted as part of proposed post-marketing requirement (PMR) studies. Missing values will be replaced with imputed values based on observed variables (e.g. age, gender, year-quarter). Imputations will be done using the PROC MI procedure in SAS 9.4 or later. The fully conditional specification method will be used to impute the missing product values and correlated variables. Discriminant function methods will be used to impute missing product values and other categorical covariates. Data will be imputed for 100 iterations and summarized across each of the imputation datasets. Adjusted rate estimates and confidence intervals will be calculated in each data set and the estimates pooled to provide robust rate estimates. More information is available upon request.

5.2.3 Careless Responses

To address **careless responses**, we have recently undertaken an extensive review of SKIP/OTP surveys, including visual inspection of completed forms. This is an active area of interest, where increasing numbers of products can lead to lengthy survey forms. We have constructed an empirical algorithm to identify improbable responses, especially in situations where massive numbers of products are endorsed or there is a clear pattern of careless responses, in order to apply appropriate exclusion criteria (described below).

SKIP and OTP use multi-page paper surveys completed by patients entering substance abuse treatment programs to identify self-reported past 30-day use of controlled substances. It was observed that a small proportion of respondents endorsed improbably high numbers of drugs used in the 30 days preceding treatment enrollment. For example, 0.1% of all SKIP and OTP respondents between 2008Q1 and 2016Q2 (N=77,555) endorsed more than 100 drugs. Approaches to identifying improbable responses in social science and consumer choice survey literature were considered. The identification and exclusion of improbable surveys could differentially affect rates of abuse of low volume prescription opioids, necessitating a careful approach that would not improperly influence drugs of interest, such as new products with low volume of sales.

Approach

Our general approach was to analyze 5+ years of SKIP and OTP data (separately) to identify parameters which could be used to exclude improbable survey responses. Visual inspection of surveys with nearly all drugs endorsed revealed patterns of misrepresentation such as extended lines connecting dozens of boxes in a column. After consideration of alternative approaches, a two-factor solution is being considered, encompassing the proportion of all drugs endorsed and the maximum number of consecutive drugs endorsed. Criteria were separate for SKIP and OTP due to page layout differences. We were also limited by the lack of availability of a gold standard reference dataset, and had to use empirically derived criteria.

Alternate Strategies Considered
Alternative strategies for identifying improbable response patterns were evaluated empirically, but ultimately rejected in favor of the solution presented below due to better performance: exclusion cut points based on the highest percentiles of all drugs endorsed, consecutive endorsements within active ingredient blocks, single run threshold regression, using the survey with the highest number of endorsements as a benchmark, page number/position influences, geographic stratification by state or treatment center, time-based change models, independent models for low versus high volume drugs, and principal components analysis using demographic and other descriptive model variables.

Maximum Consecutive Endorsements
Based on the LongString index created by Johnson (2005) to identify carelessness and misrepresentation in survey data, we calculated the maximum number of consecutive drugs endorsed on each survey. Maximum available data for OTP (N=37,564, 2011Q1 through 2016Q2) and SKIP (N=14,346 assessed from 2011Q1 through 2016Q4) were used. For OTP, 72% of respondents did not endorse any consecutive products within a column, with an average of 1.7 consecutive runs (standard deviation [SD] 2.5), whereas in SKIP 52.9% did not endorse consecutive products, averaging 2.3 consecutive endorsements (SD 2.8). This is likely explained by the difference in page length (2 versus 4) of the survey instrument between studies. Natural breakpoints emerged, with greater than 5 endorsements as potential candidates for both OTP and SKIP (data available upon request).

Proportion of All Drugs Endorsed
The number of drugs in the survey instruments fluctuated over time as new pharmaceutical products came on or were dropped from the market. Therefore, the proportion of all possible drugs endorsed was established as a denominator for each survey instrument version, and the numerator was a sum of all drugs endorsed. Visual inspection of univariate histogram distributions did not provide clear breakpoints. Threshold regression, as reviewed by Hansen 2000 and 2011, was implemented in Stata version 15, College Station, Texas, USA, with robust variance. For each program individually, a full-sample single-threshold regression was run, followed by a two-solution model. These yielded inflection points below 20%, but since it was plausible to have used 20% of drugs, especially in earlier survey versions, the single-threshold solution was rerun, limiting ranges to greater than 20%, 50% and 80%, to derive a spectrum of thresholds for further exploration.

Venn diagrams (provisional presented for consideration) were generated for each threshold level of proportion endorsed, conditional on endorsing >5 consecutive drugs, by program. We visually identified the threshold that had the least overlap between but similar total volumes (e.g., areas) of the circles. For OTP 16% of endorsements appeared to be a possible solution (Figure 9 as example), and 28% for SKIP (Figure 10 as example). The two putative criteria, when applied in conjunction resulted in the removal of 148 per 10,000 OTP surveys and 381 per 10,000 SKIP surveys. This method is being further developed and these results should be considered provisional at the current time.

FIGURE 9: OTP EXCLUSION CRITERIA OVERLAP, VARYING BY PERCENT OF ALL DRUGS ENDORSED

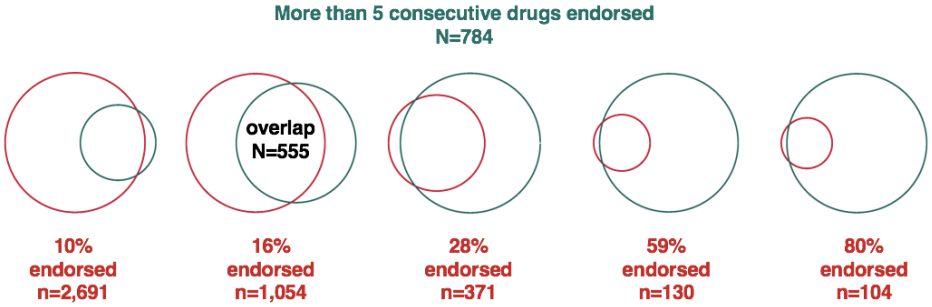
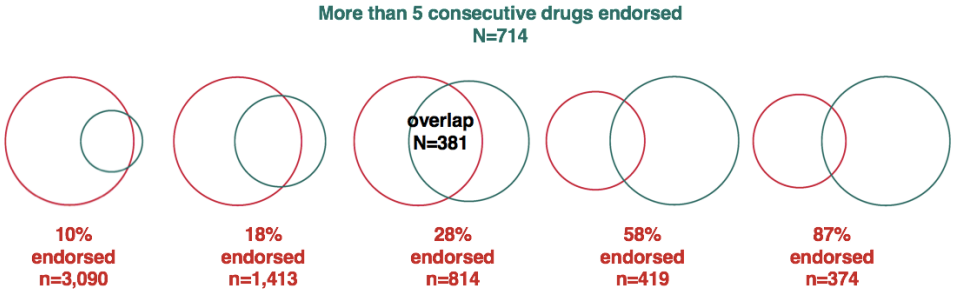


FIGURE 10: SKIP EXCLUSION CRITERIA OVERLAP, VARYING BY PERCENT OF ALL DRUGS ENDORSED



Hansen, BE (2000). Sample splitting and threshold estimation. *Econometrica,* 68: 575–603.

Hansen, BE (2011). Threshold autoregression in economics. *Statistics and Its Interface,* 4: 123–127.

Johnson JA (2005). Ascertaining the validity of individual protocols from Web-based personality interviews. *Journal of Research in Personality*, 39(1), 103-129. doi:10.1016/j.jrp.2004.09.009

Efficacy, Safety and Tolerability of Everolimus in Preventing End-stage Renal Disease in Patients With Autosomal Dominant Polycystic Kidney Disease. https://clinicaltrials.gov/ct2/show/NCT00414440?term=%22multiple+imputation%22&rank=1 Accessed 13 July 2017.

Study on Psychoeducation Enhancing Results of Adherence in Schizophrenia (SPERA-S). https://clinicaltrials.gov/ct2/show/NCT01433094?term=%22multiple+imputation%22&rank=5 Accessed 13 July 2017.

Pilot Study of Minocycline in Huntington's Disease. https://clinicaltrials.gov/ct2/show/NCT00277355?term=%22multiple+imputation%22&rank=3 Accessed 13 July 2017.

6. On Sampling of Treatment Centers

In Section 3.2 of the Issues paper, FDA stated: "Shifts in the geographic distribution of the sample, as well as in the distribution of the types of assessment sites contributing data to the system (e.g., public versus private treatment program, inpatient center versus probation office) have the potential to create bias when estimating trends in abuse rates over time."

RADARS System Comment: While generally in agreement, we welcome the opportunity to clarify the opportunity for bias arising from treatment center data.

- Research on criminal justice referrals to treatment, such as Drug Courts and Law Enforcement Assisted Diversion (LEAD) programs, has largely focused on retention and outcomes after entry into such programs; much less is known about who is offered enrollment in them. As these programs evolve, further research could elucidate the mechanisms by which people end up in treatment in general, and criminal justice referrals

more specifically. We will continue to monitor the scientific literature for these types of studies.

- One possible method to address this is to collect metadata from treatment program administrators. Effect modifiers (e.g., center effects) can be theorized, such as those listed by FDA. A relatively straightforward solution would be to survey treatment program administrators about the types of their service setting and stratify by these conditions in models of evaluation of the effectiveness of products with abuse deterrent properties to understand the extent to which, if any, that treatment setting influences findings. We are interested in pursuing this line of research in forthcoming proposals to FDA.

In Section 3.2 of the Issues Paper, FDA stated: "data from sources such as treatment centers are not based on a probability sample from a well-defined sampling frame or population. They are only captured when individuals interact with these surveillance systems. It is therefore difficult to characterize the underlying population about which statements regarding abuse and abuse-related outcomes are to be made."

RADARS System Comment: We thank the FDA for pointing out this perennial issue with interpreting treatment center information. While considerable research has focused on barriers to treatment entry, less is known about who accesses treatment and why. A recent publication on cannabis treatment seeking in Australia offers a direction for further research on opioids in the United States. The Papinczak et al. 2017 used the Social Cognitive Theory framework to assess enrollment in a cohort of people who were referred to a treatment program. They found: "Treatment seekers had significantly higher levels of negative cannabis outcome expectancies and significantly lower levels of emotional relief refusal self-efficacy (belief in ability to resist using cannabis when experiencing negative affect). Treatment seekers had significantly higher levels of psychological distress and self-perceived cannabis dependence compared to non-treatment seekers." It remains to be seen if similar patterns hold for opioids in the United States. However, a parallel can be drawn to the aforementioned criticism of poison center calls not capturing overdose deaths – treatment center admissions may represent more severe drug use disorders and those experiencing the worst consequences, compared to the general population of people with addictive disorders. To be consistent, if we are interested in the most severe cases of overdose, we should also be most concerned with the most severe addiction cases, to the extent that this observation can be supported. As such, data from treatment centers, while not perfect, offer a crucial glimpse into an important population, especially given the product and route specificity lacking in many other data sources.

Papinczak ZE, Connor JP, Feeney GF, Young RM, Gullo MJ. Treatment seeking in cannabis dependence: The role of social cognition. *Drug Alcohol Depend*. 2017 Jan 1;170:142-146. doi: 10.1016/j.drugalcdep.2016.11.005. Epub 2016 Nov 14. PMID: 27894043

7. On Direct Adjustment of Treatment Center Data

In Section 4.5 of the Issues Paper, FDA stated: "Geographic heterogeneity in abuse patterns can complicate the analysis and interpretation of the postmarket data, particularly in data sources that use nonprobability samples. When an estimate based on a coarse unit of analysis (e.g. national survey) shows a change in one direction while estimates based on a more granular unit of analysis (e.g., an analysis stratified by geographic region) show a change in a different direction, a paradox can result. A similar issue can arise in the context of multicenter clinical trials when patient treatment assignment varies from center to center (Rosenbaum, 2002). With treatment center data, for example, the number of individuals contributing data at the state level can decrease or increase simply by new treatment centers joining or existing treatment centers dropping out of the surveillance network over the course of the study period.

This can lead to conflicting results between the direction of change between state-specific estimates and estimates aggregated over all states. Restricting the analysis to sites that contribute information throughout the study period may alleviate this issue to some extent, but not eliminate it. Note, for multicenter clinical trials, Rosenbaum discussed a standardization approach using direct adjustment that may be potentially useful."

RADARS System Comment: The issue of fluctuations in treatment center enrollment over time is analogous to the discussion above for poison centers. For the specific example cited above by FDA, subset analyses of consistently reporting treatment centers could address the issue; consistency of effect in these subset programs would provide confidence in the findings.

It also appears that geographic heterogeneity of opioid use and consequences is a consistent finding from other independent data sources, regardless of probability sampling (hospital data) or complete case enumeration (mortality data). We point out the recent publication by Unick and Ciccarone 2017 showing geographical differences in inpatient hospitalizations for prescription opioid disorder (POD) and heroin overdoses using the National Inpatient Sample, with weighting for national estimates. The authors note: "Between 2012 and 2014 POD rates decreased in eight of the nine census divisions, with only New England showing an increase." For an example with complete case enumeration, Rudd et al. 2016 stated: "The drug overdose death rate increased significantly from 12.3 per 100,000 population in 2010 to 16.3 in 2015. Death rates increased in 30 states and DC and remained stable in 19 states." (The CDC report did not use direct adjustment when reporting the national estimate presumably because mortality data are expected to be complete case counts and not sampled.)

- We were unable to locate the specific text on multi-center clinical trials in Rosenbaum et al. 2002, and assume that the reference to Rosenbaum is Section 2.7.1 "Unbiased Estimates of the Average Effect," on pages 46-7. We thank the FDA for pointing out this approach. We are interested in exploring direct adjustment, pursuant to FDA's suggestion. Rosenbaum et al. 2002 provides the method in the context of efficacy and effectiveness studies, structured as cohorts of exposed and unexposed individuals; we suppose that the concept could be explored for observational studies for evaluating effectiveness of products with abuse deterrent properties, but note that these studies do not constitute findings of efficacy as per the regulatory definition.
- A fundamental question arises, however, namely how to translate the Rosenbaum method to units of geographic or center-based stratification. As Rosenbaum et al. 2002 points out: "In effect, direct adjustment views the treated units and the control units as two stratified random samples from the N units in the experiment… On the other hand, the average effect is but a summary of the effects, and not a complete description, when the effect varies from one stratum to another." (page 47) The fundamental issues is of heterogeneity of effect over the source observation units. In the case of evaluating abuse deterrent properties using surveillance data, it is not possible to stratify by treatment and control groups on a geographic basis (or treatment center), because a multitude of opioid analgesics are dispensed in nearly every location. Stated another way, no single geographic unit will have exclusive utilization of these products, and no location will have a single comparator exclusively utilized. Within this context, it is unclear what direct adjustment of geographic location (weighting by number of respondents per treatment centers to extend the Rosenbaum example) would estimate. We are open to further exploration of this statistical weighting approach, but we believe it requires further clarification.

  An analogous concept would be to weight by market share across opioid analgesics in a given location, with the caveat that using dispensed drug would not represent which

diverted products are most available to people using opioids non-medically. Perhaps drug diversion data from law enforcement could serve as an alternative approach to adjustment.

We suggest a few points that bear clarification:

1. Is geographic homogeneity an expected outcome for products with abuse deterrent properties? Within the issue articulated by FDA above, it is unclear what geographic homogeneity of effect in RADARS System data would mean in the context that abuse indicators in other data sources also show marked heterogeneity. Stated another way, if the underlying phenomenon of overdose and abuse are heterogeneously distributed by geography, would an effect need to be observed uniformly over geography to be convincing of the effectiveness of abuse deterrence?
2. In multi-center clinical trials submitted for drug approvals, study site effects are evaluated with tests of heterogeneity (e.g., Mantel-Haenszel, Wilcoxon Rank Sum, etc.) and random effects models. It is unclear what advantage direct adjustment would have over these traditional methods.

Rudd RA, Seth P, David F, Scholl L. Increases in Drug and Opioid-Involved Overdose Deaths — United States, 2010–2015. *MMWR Morb Mortal Wkly Rep* 2016;65:1445–1452. DOI: http://dx.doi.org/10.15585/mmwr.mm655051e1

Unick GJ, Ciccarone D. US regional and demographic differences in prescription opioid and heroin-related overdose hospitalizations. *Int J Drug Policy*. 2017 Jul 5. pii: S0955-3959(17)30166-4. doi: 10.1016/j.drugpo.2017.06.003. [Epub ahead of print] PMID: 28688539

8. On Confounding and Secular Trends

In Section 4.6 of the Issues Paper, FDA stated: "Many factors can affect trends in opioid prescribing and abuse, for example, changes in formularies and insurance coverage policies; provider education initiatives and clinical practice guidelines; increasing use of state-level Prescription Drug Monitoring Programs (PDMPs); large-scale initiatives to reduce rogue opioid prescribing and dispensing (e.g., pill mill crackdowns); and the availability of alternative drugs, including heroin. These and other larger community forces can be described as secular trends. Isolating the effect of a specific product's reformulation from the effects of secular trends is a challenging endeavor."

RADARS System Comment: We commiserate with the FDA on this difficult problem. However, the OxyContin reformulation data displayed a marked decrease in abuse in the months after reformulation and before the Florida interventions were implemented, as depicted in Figure 11, from a forthcoming publication. At least for OxyContin, the secular effects appear to be independent in time to reductions in abuse after reformulation. For newer opioids with abuse deterrent properties, however, the problem remains.

A recent publication by Alexandridis et al. 2017 evaluated a multi-faceted intervention to reduce prescription opioid overdose in North Carolina. The seven interventions evaluated could be conceptualized as competing interventions, and the intent of the paper was to ascertain which intervention strategies were most impactful in reducing overdoses. In their modeling approach, the authors constructed dichotomous indicators for each of the seven intervention strategies by county-month, and analyzed relative contributions using interrupted time series statistics, allowing different interventions to start at different calendar times. Notably, prescription volume (number of prescriptions) was treated as a covariate instead of an offset (which was population), and baseline general community health characteristics were adjusted for to account for non-random intervention allocation. It might be feasible to replicate a similar analysis on a national scale to possibly isolate the impact of formulations with abuse deterrent properties from other

concurrent interventions. However, this requires assembly of an exhaustive list of national, state and local interventions that have occurred, with start and end dates. Thus far this type of data resource has not been created. FDA or NIH may want to consider spurring the development of an interventions database which could be used for this type of analysis and be made available to all researchers.

FIGURE 11. EXCERPT FROM DART ET AL. 2017 IN JOURNAL OF OPIOID MANAGEMENT, SHOWING COMPETING INTERVENTIONS AFTER THE INTRODUCTION OF REFORMULATED OXYCONTIN.

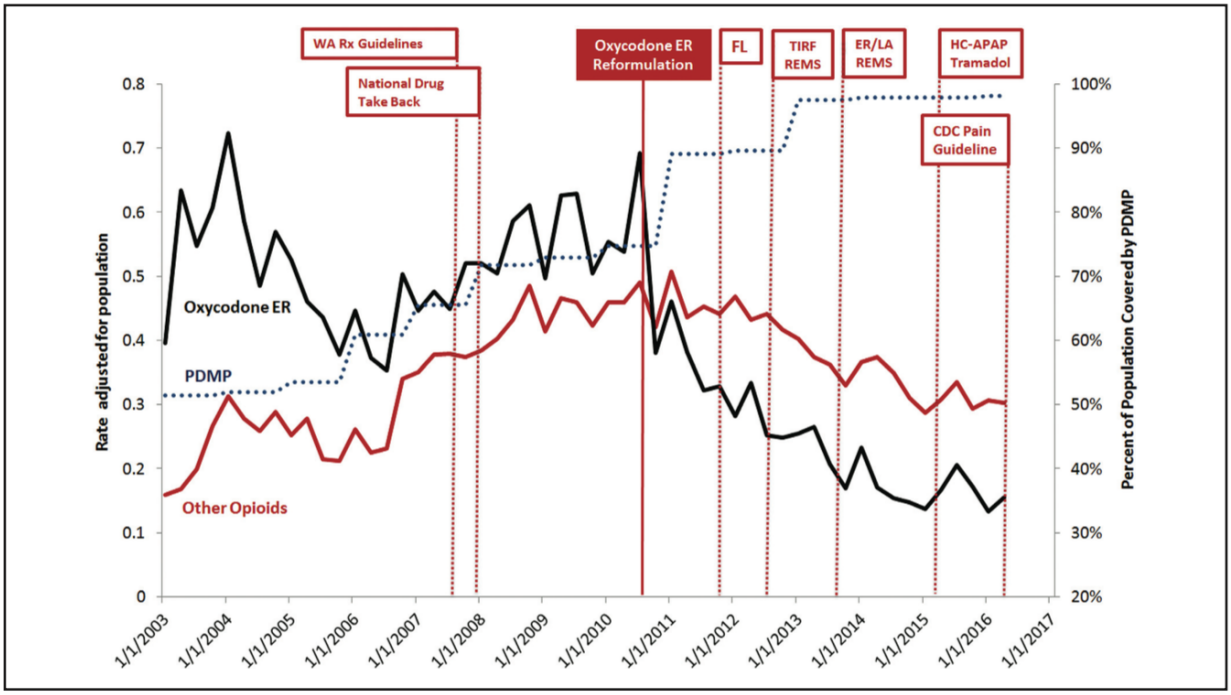


Figure 3. Temporal relation of RADARS poison Center Intentional Abuse case rate and interventions to reduce prescription drug abuse. Rates of intentional abuse in the RADARS System Poison Center Program adjusted for population are compared for oxycodone ER (black solid line) and all opioid analgesic oral dosage forms combined (hydrocodone, hydromorphone, morphine, oxymorphone, tramadol, tapentadol, and immediate release oxycodone). The proportion of states with an active Prescription Monitoring Program is provided in blue. The vertical intersects represent the initiation of a variety of interventions intended to reduce prescription opioid abuse. CDC, Centers for Disease Control and Prevention; WA, Washington; ER, extended release; FL, Florida; TIRF REMS, transmucosal immediate release fentanyl risk evaluation and mitigation strategy; ER/LA REMS, extended release/Long-acting risk evaluation and mitigation strategy; HC/APAP, rescheduling of hydrocodone-acetaminophen combination products from Schedule III to Schedule II; Tramadol, tramadol becomes a Schedule IV controlled substance.

Alexandridis AA, McCort A, Ringwalt CL, Sachdeva N, Sanford C, Marshall SW, Mack K, Dasgupta N. A statewide evaluation of seven strategies to reduce opioid overdose in North Carolina. *Inj Prev*. 2017 Aug 23. pii: injuryprev-2017-042396. doi: 10.1136/injuryprev-2017-042396. [Epub ahead of print] PMID: 28835443

Dart RC, Iwanicki JL, Dasgupta N, Cicero TJ, Schnoll SH (2017). Do abuse deterrent opioid formulations work? *Journal of Opioid Management*, 13(4), in press.

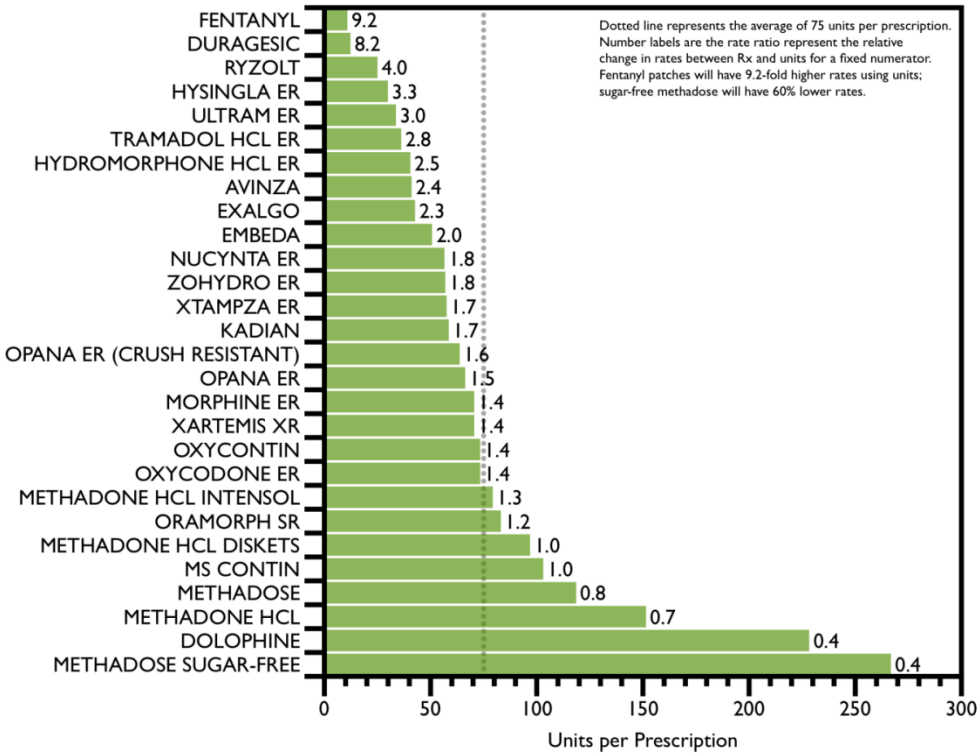
9. On Units Dispensed as a Denominator

In Section 4.3. of the Issues Paper, FDA stated: "FDA generally considers the number of dosage units dispensed to be superior to the number of prescriptions dispensed or the number of individuals receiving a prescription, because every dosage unit presents an opportunity for abuse, and the average number of dosage units per prescription may vary across opioids."

RADARS System Comment: While we welcome the FDA's efforts to understand denominators, we are concerned about considering units dispensed as the *ideal* denominator. We feel that multiple denominators can provide a more complete picture of the phenomenon.

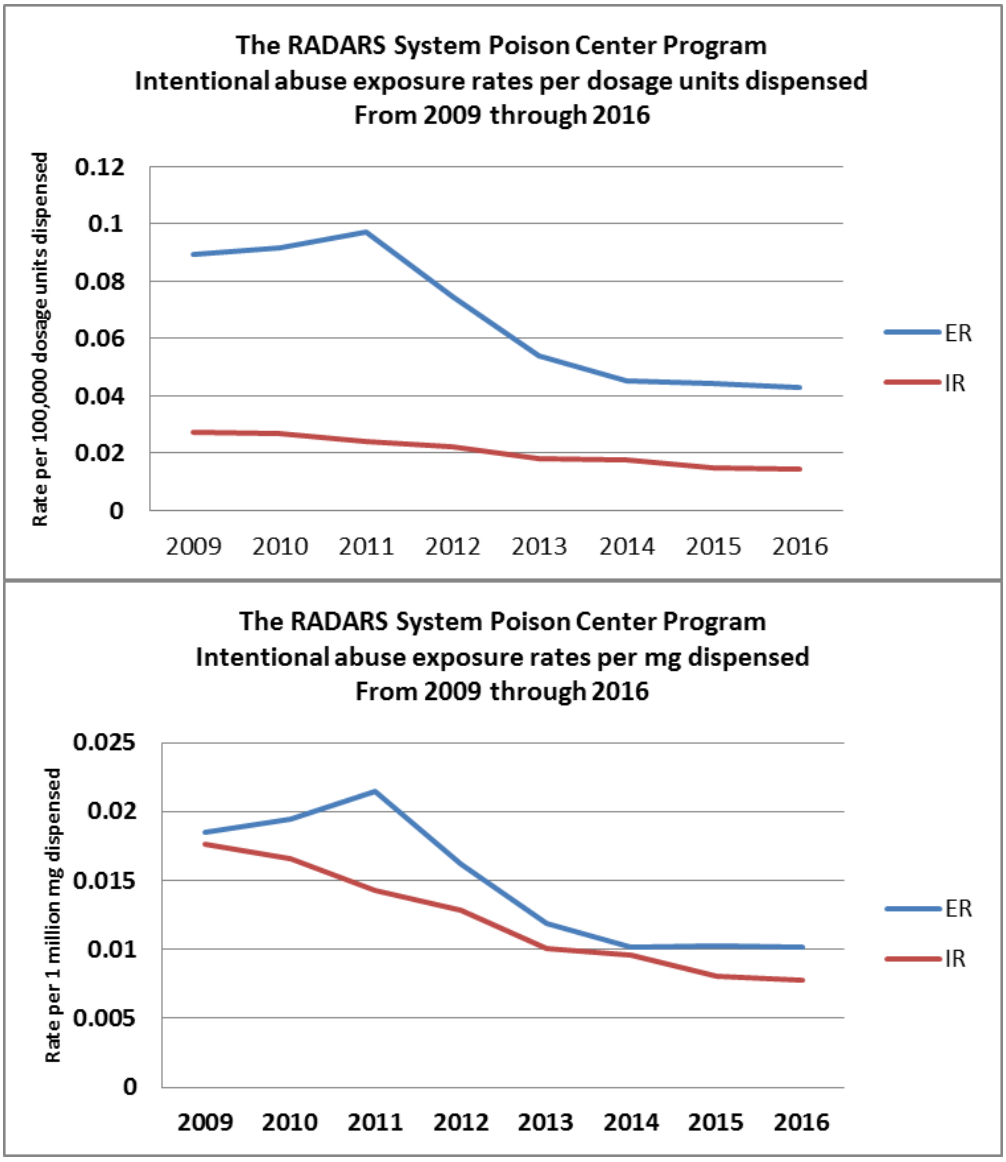
We recently conducted an analysis of 2016 QuintilesIMS national dispensing data. The (quarterly averaged) number of units varies greatly between ER opioid analgesics (Figure 12). There was an average of 75 units dispensed per prescription across ER opioids, with considerable variation by active ingredient. If every drug on this list had the same exact rate of abuse using the number of prescriptions as a denominator (from any given data system for the numerator), just by switching denominators to units dispensed, the abuse rate would jump 8-9x for fentanyl patches, Nucynta ER rates would go up by 77%, Xartemis ER goes up by 42%, etc. but the ER morphine products stay about the same. In effect, methadone abuse rates would be lower than if using number of prescriptions as a denominator because there are about 150 units per prescription; the denominator would be twice as large as the group average. This leads to a question of interpretation with specific consequences: Would this observation be interpreted as methadone is inherently less prone to abuse than ER morphine?

FIGURE 12. QUINTILESIMS NATIONAL DISPENSING DATA ILLUSTRATING VARIATION IN NUMBER OF UNITS DISPENSED PER PRESCRIPTION



We also conducted an analysis comparing all IR opioid products to all ER opioid products, using both units dispensed and milligrams dispensed denominators (Figure 13). By design, ER products often have significantly higher dosage strength than IR products. Hence, a unit dispensed approach would assume a 5 mg IR oxycodone product is equivalent to an 80 mg ER oxycodone product whereas a milligrams dispensed approach would account for the 16-fold difference. The interpretation of these rates would be different depending on the denominator used.

FIGURE 13. RATES OF INTENTIONAL ABUSE FOR IR AND ER OPIOIDS FROM THE RADARS SYSTEM POISON CENTER PROGRAM ILLUSTRATING IMPORTANCE OF DENOMINATOR SELECTION



We believe that it is important to select a denominator that addresses the question at hand, with consideration of the comparators being studied. The table below outlines the potential denominators, the assessment that they offer, and strengths and limitations of each. We suggest that the denominator be determined based upon the research question being asked in light of the special considerations noted.

TABLE 1. STRENGTHS AND LIMITATIONS OF AVAILABLE DENOMINATORS FOR RATE CALCULATION

Denominator	Can tell us…	Limitations	Considerations
Population	Risk per population (overall public health burden)	Does not account for drug availability which vary across API, drug products and formulations	National population does not change quickly but population sampled may so it is an important adjustment with changes in sampling frame.
Prescriptions Dispensed	Risk per prescription dispensed	Assumes equivalent risk of varying types of prescriptions (e.g. 5 day versus 30 day supply) AND of varying amount of active ingredient in each unit (e.g. 5 mg versus 80 mg tablet) AND of varying potency between active ingredients (e.g. tramadol versus hydromorphone)	Best use may be when studying like products (type, dosage strength, potency), particularly during times of changing prescribing behaviors.
Dosage Units Dispensed	Risk per dosage unit dispensed	Assumes equivalent risk of varying amount of active ingredient in each unit (e.g. 5 mg versus 80 mg tablet) AND of varying potency between active ingredients (e.g. tramadol versus hydromorphone)	Best use may be when studying like products (dosage strength and potency) or comparing products with large variation in prescription type (e.g. total amount dispensed per prescription).
Milligrams Dispensed	Risk per milligram dispensed	Assumes equivalent risk of varying potency between active ingredients (e.g. tramadol versus hydromorphone)	Best use may be when comparing products with large variation in prescription type (e.g. total amount dispensed per prescription) and different products (type, formulation, strength), such as IR versus ER products.
Morphine-equivalent Milligrams Dispensed	Risk per milligrams dispensed adjusted for potency	Assumes morphine-equivalence impacts abuse/misuse. Conversions are estimates and are not well understood for some active ingredients.	Best use may be when studying products with well-documented conversion factors. May not be appropriate for all active ingredients.

10. Catchment Areas for Substance Abuse Treatment Centers

In Section 4.3 of the Issues Paper, FDA stated: "Because prescription drugs may be used in areas remote from where they are dispensed, the best catchment area for utilization data can be unclear in studies that do not have nationally representative samples."

RADARS System Comment: Many efforts have been made to define catchment areas in healthcare (e.g., Dartmouth Atlas of Health Care), but has been limited in applications of substance abuse treatment centers. We make reference to a paper by Pang and Lee 2008 which articulated a method for calculating coverage of methadone maintenance in Hong Kong. In this paper, the authors used areal interpolation in a Geographic Information System (GIS) to layer data from a centralized heroin addiction registry and treatment center utilization patterns to generate catchment areas for each methadone clinic. They made logical assumptions for a predominately urban area (e.g., walking distance to clinics, etc.) which will need to be modified for the US. We can envision combining data from the SAMHSA Treatment Facility Locator and

NSSATS, RADARS System Treatment Center Programs, retail buprenorphine dispensing for addiction, and sub-state NSDUH estimates to replicate the approach in the United States. We also point out that even in “nationally representative” samples, catchment area is often poorly defined, or oversimplified to assume uniform coverage throughout the country. For substance abuse treatment programs, this areal uniformity assumption would obscure meaningful variation in state and local treatment program availability, and therefore believe that national representativeness alone is not as crucial as implied.

> Tak Ting P Pang, Shui Shan Lee. Measuring the geographic coverage of methadone maintenance programme in Hong Kong by using geographic information system (GIS). *Int J Health Geogr*. 2008; 7: 5. Published online 2008 Jan 30. doi: 10.1186/1476-072X-7-5. PMCID: PMC2268919

11. Final Comment

The RADARS System would like to acknowledge the commitment and thoughtfulness of the FDA in addressing these complex methodological issues and we look forward to working in partnership to enhance the currently available data sources as well as in the development of new potential data streams.