Comparative effectiveness of extended release naltrexone and sublingual buprenorphine for treatment of opioid use disorder among Medicaid patients

Rachael K. Ross, Edward V. Nunes, Mark Olfson, Matisyahu Shulman,
Noa Krawczyk, Elizabeth A. Stuart, Kara E. Rudolph

Affiliations: Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY (Ross, Olfson, Rudolph); Department of Psychiatry, Columbia University Irving Medical Center (Nunes, Olfson, Shulman); New York State Psychiatric Institute, New York, NY (Nunes, Olfson, Shulman); Department of Population Health, New York University, New York, NY (Krawczyk); Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD (Stuart)

Corresponding author: Dr. Rachael Ross, Department of Epidemiology, Mailman School of Public Health, Columbia University, 722 W 168th St, New York, NY 10032; rachael.k.ross@columbia.edu

Running title: Naltrexone versus buprenorphine

Word count: 3501

Data availability statement: Research data are not shared. The data that support the findings of this study are available from the Research Data Assistance Center. Restrictions apply to the availability of these data, which were used under license for this study. All code and data definitions are available at
`https://github.com/CI-NYC`

Abstract

Aims: To compare the real-world effectiveness of extended release naltrexone (XR-NTX) and sublingual buprenorphine (SL-BUP) for the treatment of opioid use disorder (OUD)

Design: An observational active comparator, new user cohort study

Setting: Medicaid claims records for patients in New Jersey and California, 2016-2019

Participants/Cases: Adult Medicaid patients aged 18-64 years who initiated XR-NTX or SL-BUP for maintenance treatment of OUD and did not use medications for OUD in the 90-days before initiation

Comparators: New initiation with XR-NTX versus SL-BUP for the treatment of OUD

Measurements: We examined two outcomes up to 180 days after medication initiation, 1) composite of medication discontinuation and death, and 2) composite of overdose and death

Findings: Our cohort included 1,755 XR-NTX and 9,886 SL-BUP patients. In adjusted analyses, treatment with XR-NTX was more likely to result in discontinuation or death by the end of follow-up than treatment with SL-BUP: cumulative risk 76% (95% confidence interval [CI] 75%, 78%) versus 62% (95% CI 61%, 63%), respectively (risk difference 14 percentage points, 95% CI 13, 16). There was minimal difference in the cumulative risk of overdose or death by the end of follow-up: XR-NTX 3.8% (95% CI 2.9%, 4.7%) versus SL-BUP 3.3% (95% 2.9%, 3.7%); risk difference 0.5 percentage points, 95%CI -0.5, 1.5. Results were consistent across sensitivity analyses.

Conclusions: Longer medication retention is important because risks of negative outcomes are elevated after discontinuation. Our results support selection of SL-BUP over XR-NTX. However, most patients discontinued medication by 6 months indicating that more effective tools are needed to improve medication retention, particularly after initiation with XR-NTX, and to identify which patients do best on which medication.

Key words: extended release naltrexone, sublingual buprenorphine, opioid use disorder, Medicaid, treatment retention, overdose

Introduction

In the US, there are three Food and Drug Administration approved medications for opioid use disorder (OUD) treatment. Methadone, a full opioid agonist, can only be provided for the treatment of OUD at federally accredited Opioid Treatment Programs (OTPs). In contrast, buprenorphine (BUP), a partial opioid agonist, and naltrexone (NTX), an opioid antagonist, can be provided for the treatment of OUD in any medical setting. Although NTX is available as a daily oral medication and an extended-release (XR) monthly injection, current OUD guidelines only recommend XR-NTX.[1, 2] BUP also has oral (sublingual, SL) and XR formulations, though SL-BUP is by far the most commonly prescribed BUP formulation for OUD.[3, 4] Therefore, outside of OTPs, patients and their providers may have a choice between XR-NTX and SL-BUP for the treatment of OUD and more evidence, particularly evidence from usual care settings, is needed on their relative effectiveness.[2]

XR-NTX and SL-BUP for treatment of OUD have previously been compared in randomized clinical trials.[5, 6, 7] In these trials there was large variability (range 47%-89%) in the proportion of participants assigned XR-NTX who successfully initiated treatment. In contrast, most participants (73%-94%) successfully initiated SL-BUP. Treatment retention was similar between XR-NTX and SL-BUP among those who successfully initiated medication across the trials. Only one trial had sufficient overdose events to compare the medications for this outcome.[7] Although the original analysis did not report a significant difference in the proportion of overdoses by medication, a reanalysis found that XR-NTX had a higher hazard of overdose than SL-BUP.[8] Thus, questions remain about the relative protection against overdose provided by these medications.[9, 10, 11]

Few studies have compared these medications in usual care settings using real-world data. Although randomized clinical trials are the gold-standard in producing evidence of comparative efficacy, there are important reasons to compare these medications using real-world data. Patients enrolled in clinical trials may not be representative of the real-world population of patients with OUD.[12] How treatment is delivered (e.g., intensity of follow-up, monetary incentives) also differs between trials and real-world practice. These differences in patient characteristics and treatment delivery are likely to modify treatment effectiveness.[13, 14] Additionally, real-world data sources tend to provide larger samples sizes, which enables analysis of rare outcomes like overdose. We identified just two studies that compared XR-NTX and SL-BUP initiation using real-world data; one examined treatment retention and the other examined overdose.[15, 16] However these studies used data (US commercial claims) from 2010-2014 and 2010-2016, respectively; periods that do not capture the rise in fentanyl.[17] Other studies did not distinguish between XR and oral NTX.[18,

19] Therefore, we sought to update real-world evidence on the comparative effectiveness of XR-NTX and SL-BUP. Specifically, we compared the risk of early treatment discontinuation (before 6 months) and the risk of overdose after initiation of XR-NTX or SL-BUP for the treatment of OUD among Medicaid patients in New Jersey (NJ) and California (CA).

Methods

Study design and data source

We implemented an active comparator, new user cohort study[20] using 2016-2019 Medicaid claims data from CA and NJ (Transformed Medicaid Statistical Information System [T-MSIS] Analytic Files). Both CA and NJ implemented Medicaid expansion in 2014. These data include inpatient and other services claims and outpatient pharmacy dispensing claims for Medicaid enrolled patients. The Columbia University Institutional Review Board approved this study (AAAU3366).

We used the target trial emulation framework to inform the study design and analytic decisions.[21] Appendix Table 1 outlines the elements of the trial we attempted to emulate using observational data. The new user cohort study emulates a trial that enrolls individuals with OUD who are not currently taking XR-NTX or SL-BUP or any other medication for OUD (MOUD).

Study cohort

Appendix Figure 1 depicts the study cohort design including the time windows used to assess inclusion and exclusion criteria. Figure 1 is the flow diagram illustrating cohort assembly. As described in detail below, the cohort included adult Medicaid patients aged 18-64 years who initiated XR-NTX or SL-BUP for maintenance treatment of OUD, did not have MOUD use in the 90-days before initiation, and had 6-months of continuous enrollment before initiation. We excluded patients who were pregnant, had a cancer diagnosis, or received long-term, palliative, or hospice in the baseline period. We included patients at their first eligible medication initiation and the date of initiation was the date of cohort entry (day 0).

Cohort creation details

We first identified patients aged 18-64 years with a claim for XR-NTX or SL-BUP using 1) a National Drug Code (NDC) on an outpatient pharmacy claim, inpatient claim, or other services claim; or 2) a procedure code (Healthcare Common Procedure Coding System or Current Procedural Terminology) on an inpatient or

other services claim (Appendix Table 2). Codes were compiled from multiple sources and the literature.[22, 23, 24, 25] We refer to this medication claim as the index claim and the date is day 0. A single patient may have multiple claims for XR-NTX or SL-BUP. The exclusion criteria below were applied to each index claim (as a potentially eligible treatment initiation).

First, to establish a baseline period, we excluded patients who were not continuously enrolled with full benefits and without dual Medicare coverage in the calendar month of index claim or in the 6 calendar months prior. Second, to restrict to new users, we excluded patients with other MOUD claims (for BUP (any route), XR-NTX, or methadone) on the same day as the index claim or in the 90 days prior. Third, to restrict to new users of maintenance medication, as opposed to medication for medically managed withdrawal, we excluded SL-BUP treatment that was less than 7 days in duration (assessed using claims during days 0 to +14). Because this criterion necessarily required us to use data after day 0, we also excluded patients who were censored (because of loss to follow-up or study period ended) or who had the outcome in the first 14 days to avoid immortal time bias.[26] Fourth, to restrict to OUD treatment, we excluded patients who did not have an ICD-10 diagnosis code indicating opioid dependence, abuse, or overdose in the 6-month baseline period (defined above). Fifth, we excluded patients who were pregnant on day 0, or who had a cancer diagnosis or received long-term, palliative, or hospice care in the baseline period. Finally, as patients may be eligible for our cohort multiple times in the study period, we restricted to the first eligible medication initiation.

Treatments

We compared initiation of XR-NTX to initiation of SL-BUP for the treatment of OUD.

Outcomes

We assessed two composite outcomes: 1) medication discontinuation or death, and 2) overdose or death. We divided the period after treatment initiation into weeks, where week 1 was days 1-7, week 2 was days 8-14, etc. and we continued follow-up to 180 days (week 26). Outcome measurement started in week 3 (day 15).

Medication discontinuation

We examined two definitions for medication discontinuation: 1) discontinued all MOUD (i.e., patients were allowed to switch treatments, including between XR-NTX and SL-BUP or to methadone or XR-BUP) and 2) discontinued initial medication. We considered time during which the patient had access to dispensed

or administered medication as continued use. We assumed use began on the day of the medication claim regardless of overlap with prior claims (i.e., we did not allow stockpiling). For XR-NTX, we assumed 30 days of use for each claim. For SL-BUP, we assumed use equaled the days supply for pharmacy fills or 1 day for all other codes. We summed days for multiple SL-BUP claims on the same day. In our first definition of discontinuation, we allowed switching to XR-BUP or methadone. For XR-BUP, we assumed 30 days for injection and 180 days for implant. For methadone, we assumed 1 day except for weekly bundle codes. We summed methadone claims on the same day.

We used a 31-day grace period: once medication for a claim was finished, we looked for a subsequent claim for up to 31 days; if there was no claim, the date of discontinuation was the end of the grace period. Prior research[27] has used a 30-day grace period, however we observed that some claims occurred exactly monthly.

Overdose

Given our data source, we only captured overdose that resulted in an insurance claim. We used ICD-10 diagnosis codes on inpatient or other services claims for opioid or unspecified narcotic poisoning (Appendix Table 2).[28]

Death

We identified the occurrence and date of death from the demographic and enrollment file. These data captured death during Medicaid enrollment only.

Covariates

Demographic factors were assessed during the calendar month of cohort entry: age, gender, race/ethnicity, state, and Temporary Assistance for Needy Families (TANF) benefits. Other baseline covariates (Table 1) were assessed in the baseline period (prior to cohort entry). They included OUD characteristics, other substance use disorders, psychiatric comorbid conditions, other comorbid conditions, medicationsand healthcare utilization. Code lists are available at `https://github.com/CI-NYC`.

Primary analysis

We aimed to estimate the average treatment effect (ATE), which is the difference in the marginal cumulative risk of the outcome throughout follow-up under two counterfactual scenarios: 1) had everyone in the study cohort initiated XR-NTX, versus 2) had everyone in the study cohort initiated SL-BUP. In the analysis

of each composite outcome, we censored patients at the earliest occurrence of 1) the outcome, 2) loss to follow-up, 3) study period end on 12/31/2019, or 4) 180 days (26 weeks). A patient was considered lost to follow-up when the patient disenrolled, no longer had full benefits, or acquired dual coverage.

First, we conducted a crude analysis. We used the Kaplan-Meier estimator to estimate the cumulative risk of each outcome by week for each medication and the risk difference between medication groups. Next, we conducted an adjusted analysis that accounted for confounding and potentially informative censoring by standardizing by the covariates described above and calendar year at cohort entry. For the adjusted analysis, we used a sequentially doubly robust estimator (R package lmtp).[29] This estimator relies on estimating models for treatment conditional on covariates; time-varying censoring conditional on treatment and covariates; and time-varying outcome conditional on treatment and covariates. These models were estimated flexibly using an ensemble of machine learning algorithms that included an intercept-only model, a main-effects logistic model, gradient boosting machines, and multivariate adaptive regression splines.[30] We estimated the variance using the sample variance of the efficient influence function and constructed Wald-type 95% confidence intervals (CI).[31] Such an analysis is accurate (i.e., our estimated risk difference can be interpreted as the ATE) when there is no residual bias from confounding or informative censoring (i.e., conditional exchangeability) and all patients (defined by the measured covariates) have the possibility of receiving either medication and of remaining uncensored (i.e., positivity). We used R (v4.3.1; R Core Team 2021).

Additional analyses

In our cohort, we observed that patients who initiated XR-NTX were more likely to have alcohol use disorder (AUD) and less likely to have chronic pain than patients who initiated SL-BUP. Given that XR-NTX is also indicated for AUD and BUP is indicated for pain, we suspected that providers were preferentially choosing specific medications if other indications were present. Therefore, we conducted two stratified analyses (by AUD and by chronic pain) to assess whether results varied by subgroups. We also conducted multiple sensitivity analyses. These are detailed in the appendix.

Results

Cohort

After applying our exclusion criteria (Figure 1), the study cohort included 11,641 patients: 1,755 initiated XR-NTX and 9,886 initiated SL-BUP. Table 1 includes the cohort characteristics overall and stratified by medication (Appendix Table 3 includes standardized mean differences). In general, patients who initiated XR-NTX had a higher prevalence of many health conditions, higher medication use (particularly antidepressants and antipsychotics), and higher healthcare utilization than patients who initiated SL-BUP. Disability and chronic pain were the only conditions that were notably higher among patients who initiated SL-BUP. The median days supply for SL-BUP was 15 (interquartile range [IQR] 11, 30).

Outcomes

Appendix Table 4 includes the number of outcomes identified, person-time of follow-up and the crude event rates. For the composite outcome of discontinuation all MOUD or death, 6,752 (58%) patients discontinued, 4 (0.03%) died, and 474 (4%) were lost to follow-up. For the composite outcome of overdose or death, we observed 349 overdoses (3%), 4 deaths (0.03%), and 928 (8%) patients were lost to follow-up.

Primary analysis

Figure 2 presents the crude and adjusted risk curves (in %) and risk difference (RD) curves (as percentage points) for the composite discontinuation of all MOUD and death. The numeric results at specific time points are presented in Table 2. Because XR-NTX treatment duration was 30 days and we used a 31-day grace period, patients who initiated XR-NTX were not at risk of discontinuation until day 61 (week 9). Thus, we see a jump in the risk of discontinuation at week 9 for the XR-NTX group. For SL-BUP, initial medication supply varied so patients were at risk of discontinuation earlier (end of supply + 31 days). Therefore, before week 9, the risk of discontinuation is lower for XR-NTX (RD<0), but after week 9, discontinuation is higher for XR-NTX (RD>0). The adjusted risk of discontinuation by week 26 (end of follow-up) after XR-NTX initiation was 76% (95% CI 75%, 78%); the risk was 62% (95% CI 61%, 63%) after SL-BUP initiation. The adjusted risk difference was 14 percentage points (95% CI 13, 16), i.e., among 100 patients, 14 additional patients are expected to discontinue MOUD or die by 6 months if everyone were treated with XR-NTX compared to if everyone were treated with SL-BUP. Results were similar for the outcome of discontinuation of initial medication (Table 2 and Appendix Figure 2).

8

Figure 3 and Table 2 present results for the composite of overdose and death. In the crude analysis, the risk among patients who initiated XR-NTX began to separate from the risk among patients who initiated SL-BUP around week 13. By the end of follow-up, the crude risk after XR-NTX initiation was 5.1% (95% 4.1%, 6.3%); the crude risk after SL-BUP initiation was 3.1% (95% 2.7%, 3.4%). The crude risk difference was 2.1 percentage points (95% CI 0.06, 3.5) indicating a higher observed risk of overdose after XR-NTX. In the adjusted analysis, however, the risk curves are more similar to each other. The adjusted risk was 3.8% (95% CI 2.9%, 4.7%) after XR-NTX initiation and 3.3% (95% 2.9%, 3.7%) after SL-BUP initiation; the adjusted risk difference was 0.5 percentage points, 95% CI -0.5, 1.5).

Additional analyses

Table 3 presents the adjusted risk differences by 26 weeks for the additional analyses. We did not observe heterogeneity by AUD or chronic pain in the comparative effect of medication on discontinuation. Results did not vary meaningfully across sensitivity analyses or compared to the primary analysis. For overdose, the largest risk difference observed was 1.1 percentage points (95% CI -0.7, 2.8).

Discussion

Our objective was to compare the real-world effectiveness of XR-NTX and SL-BUP. Such evidence can help inform Medicaid patients and their providers deciding between XR-NTX or SL-BUP for OUD. We found that, consistent with national data,[32] a majority of patients discontinued medication before 6 months (a minimally adequate duration of treatment[33]), though the proportion that discontinued was notably higher for XR-NTX than for SL-BUP. Despite this greater early medication discontinuation, there was only a small elevated risk of overdose (<1 percentage point) for XR-NTX after adjustment.

Randomized trials comparing these medications have not observed a notable difference in medication retention.[5, 6, 7] In contrast, and consistent with our results, a prior analysis of commercially insured patients found the hazard of discontinuation was 2 times higher for XR-NTX compared to SL-BUP.[15] There are several potential reasons for this discrepancy between trial and real-world data results. First, there are likely differences in the distribution of patient characteristics between the trial and real-world population[12] and these characteristics may increase or decrease the risk of medication discontinuation. For example, the proportion of people experiencing unstable housing may differ between trial and real-world populations and unstable housing has been shown to effect medication effectiveness.[13, 14] Second, there are

also differences in how treatment is delivered, including care setting, provider type, and retention incentives, and these differences may impact retention differently for each medication. Trials often have highly structured follow-up with research teams motivated to engage and keep in touch with patients, including reaching out to patients who miss appointments. This level of follow-up is often not available in community-based treatment. Third, the results from trials may be biased by selection. Some trial analyses were restricted to individuals who initiated medication without accounting for bias induced by conditioning on successful initiation.[34] Such selection bias may favor XR-NTX since the difficult induction (i.e., the need to be opioid abstinent) may select for participants more likely to have better outcomes.

Discontinuation of medication for OUD is associated with illicit opioid use, overdose, and other negative outcomes.[35] Thus, given the observed difference in medication discontinuation, we expected a larger difference in overdose risk. In our study, the crude risk of overdose was higher for XR-NTX than SL-BUP, but this difference was attenuated after adjustment. This attenuation reflects that patients with greater severity of illness and greater prevalence of co-morbidities were more likely to receive XR-NTX than SL-BUP (Table 1).

Results from one trial, X:BOT, have sparked debate regarding the comparative risk of overdose between XR-NTX and SL-BUP.[7, 8, 9, 10, 11] The primary analysis found no significant difference in the proportion of patients with a MEDRA defined overdose.[7] Follow-up analyses identified additional probable or possible overdose events, in part highlighting the need for more precise measurement of overdose in future trials.[8, 9] One analysis found that XR-NTX had two times the hazard of overdose than SL-BUP.[8] Further analysis suggested that the number of overdoses was similar regardless of medication, but overdoses tended to occur sooner after medication discontinuation for XR-NTX than SL-BUP.[9] It appears that the overdose risk after XR-NTX treatment was higher in the X:BOT trial (range 4-6%) than in our study (adjusted 3.8%). In our study, we were only able to capture overdose events that resulted in a Medicaid claim and this likely misses overdoses that occur in the community not resulting in a visit to a healthcare facility; in contrast, the trial had systematic follow-up that has superior capture of overdose events. It is unclear how differences in capture may contribute to these different results. The risk difference was larger in our sensitivity analysis using a more sensitive overdose definition (0.9 versus 0.5 percentage points in our primary analysis) potentially indicating bias towards the null.

We found only one study that examined overdose in real-world data and specifically assessed XR-NTX.[16] This study did not compare XR-NTX and SL-BUP directly but instead compared each medication to no medication. Further, this study assessed MOUD as a time-varying treatment and thus was not focused on

comparing medication choice at initiation (as done in our analysis and X:BOT). The authors found that in a population of commercially insured adults both XR-NTX and SL-BUP had hazard ratios less than 1 compared to no medication though SL-BUP hazard ratio was smaller, potentially indicating greater protection against overdose for SL-BUP than for XR-NTX. Both medications have high affinity for opioid receptors and protect against overdose by blocking opioid receptor, however, after discontinuation this protection is lost and patients on XR-NTX may be at higher risk of overdose because of lack of tolerance.

To interpret our estimates as valid causal effects, we relied on assumptions (see methods), most notably that there was no residual confounding. We have taken steps, including using a large number of covariates and a doubly robust estimator with machine learning, to make those assumptions more likely to hold. However, there may still be bias, particularly from residual confounding. We used insurance claims records and there may be measurement error. For example, we did not have access to data on the medication dosing instructions and adherence potentially leading to measurement of medication discontinuation. Additionally, these data may not capture all deaths during enrollment and, as mentioned above, these data only capture a subset of overdoses that generate a claim. Finally, we only had access to data from CA and NJ so these results may not be generalizable to Medicaid patients in other States or uninsured or privately insured patients.

In addition to our use of a doubly robust estimator with machine learning, our study has other strengths. We estimated adjusted risk curves and differences over time, in contrast to prior work that estimated hazard ratios. Hazard ratios are averages and thus change with the length of follow-up. They do not provide specific information about how risk of the outcome varies over time.[36] In contrast, adjusted risk curves do not ignore the distribution of outcomes during follow-up. This is particularly important for MOUD as incrementally longer duration of treatment may have benefits even though the majority of patients discontinue by 6 months. Additionally, providers and patients prefer measures of absolute risk and differences for decision making as opposed to relative measures.[37, 38]. Finally, our study used recent Medicaid data. There has been limited research comparing XR-NTX and SL-BUP among Medicaid patients and Medicaid is the largest payer of OUD treatment in the US.[39]

Conclusion

Among Medicaid patients in California and New Jersey, medication retention was longer on SL-BUP than XR-NTX though the risk of overdose was similar. On average, these results support the use of SL-BUP over XR-NTX, however, given the diversity of the population of OUD patients, XR-NTX may the best

choice for some patients, particularly those who do not want agonist treatment. Further, a minority of patients stayed on medication for 6-months, highlighting the need for other interventions to improve retention, particularly for patients initiating XR-NTX. More research is needed to identify which patients do best on which medication.

References

1. TIP 63: Medications for Opioid Use Disorde. Tech. rep. Subtance Abuse and Mental Health Services Administration, 2021 Jul. Available from: `https://store.samhsa.gov/product/TIP-63-Medications-for-Opioid-Use-Disorder-Full-Document/PEP21-02-01-002` [Accessed on: 2023 Oct 2]

2. The ASAM National Practice Guideline for the Treatment of Opioid Use Disorder: 2020 Focused Update. en. Journal of Addiction Medicine 2020 Mar; 14:1–91. DOI: `10.1097/ADM.0000000000000633`. Available from: `https://journals.lww.com/10.1097/ADM.0000000000000633` [Accessed on: 2023 Oct 15]

3. Shover CL. Commentary on Larance *et al* . (2020): Priorities and concerns of people who use opioids are key to scaling up XR-buprenorphine. en. Addiction 2020 Jul; 115:1306–7. DOI: `10.1111/add.15048`. Available from: `https://onlinelibrary.wiley.com/doi/10.1111/add.15048` [Accessed on: 2023 Oct 4]

4. Ross RK, Rudolph KE, and Shover C. Availability and prescribing of extended release buprenorphine injection for Medicaid beneficiaries, 2018-2022. en. preprint. 2024 Jan. DOI: `10.1101/2024.01.05.24300895`. Available from: `http://medrxiv.org/lookup/doi/10.1101/2024.01.05.24300895` [Accessed on: 2024 Jan 15]

5. Tanum L, Solli KK, Latif ZeH, Benth JŠ, Opheim A, Sharma-Haase K, Krajci P, and Kunøe N. Effectiveness of Injectable Extended-Release Naltrexone vs Daily Buprenorphine-Naloxone for Opioid Dependence: A Randomized Clinical Noninferiority Trial. en. JAMA Psychiatry 2017 Dec; 74:1197. DOI: `10.1001/jamapsychiatry.2017.3206`. Available from: `http://archpsyc.jamanetwork.com/article.aspx?doi=10.1001/jamapsychiatry.2017.3206` [Accessed on: 2023 Oct 1]

6. Korthuis PT, Cook RR, Lum PJ, Waddell EN, Tookes H, Vergara-Rodriguez P, Kunkel LE, Lucas GM, Rodriguez AE, Bielavitz S, Fanucchi LC, Hoffman KA, Bachrach K, Payne EH, Collins JA, Matthews A, Oden N, Jacobs P, Jelstrom E, Sorensen JL, and McCarty D. HIV clinic-based extended-release naltrexone versus treatment as usual for people with HIV and opioid use disorder: a non-blinded, randomized non-inferiority trial. en. Addiction 2022 Jul; 117:1961–71. DOI: `10.1111/add.15836`. Available from: `https://onlinelibrary.wiley.com/doi/10.1111/add.15836` [Accessed on: 2023 Apr 20]

7. Lee JD, Nunes EV, Novo P, Bachrach K, Bailey GL, Bhatt S, Farkas S, Fishman M, Gauthier P, Hodgkins CC, King J, Lindblad R, Liu D, Matthews AG, May J, Peavy KM, Ross S, Salazar D, Schkolnik P, Shmueli-Blumberg D, Stablein D, Subramaniam G, and Rotrosen J. Comparative effectiveness of extended-release naltrexone versus buprenorphine-naloxone for opioid relapse prevention (X:BOT): a multicentre, open-label, randomised controlled trial. en. The Lancet 2018 Jan; 391:309–18. DOI: `10.1016/S0140-6736(17)32812-X`. Available from: `https://linkinghub.elsevier.com/retrieve/pii/S014067361732812X` [Accessed on: 2023 Apr 17]

8. Ajazi EM, Dasgupta N, Marshall SW, Monaco J, Howard AG, Preisser JS, and Schwartz TA. Revisiting the X:BOT Naltrexone Clinical Trial Using a Comprehensive Survival Analysis. en. Journal of Addiction Medicine 2022 Jul; 16:440–6. DOI: `10.1097/ADM.0000000000000931`. Available from: `https://journals.lww.com/10.1097/ADM.0000000000000931` [Accessed on: 2023 Apr 18]

9. Lee JD, Nunes EV, Novo P, May J, Matthews A, Van Veldhuisen P, Lindblad R, Liu D, and Rotrosen J. Commentary on Ajazi et al (2021) Re-analysis of the X:BOT Trial. en. Journal of Addiction Medicine 2022 Jul; 16:382–5. DOI: `10.1097/ADM.0000000000000930`. Available from: `https://journals.lww.com/10.1097/ADM.0000000000000930` [Accessed on: 2023 Apr 24]

10. Dasgupta N, Ajazi EM, Schwartz TA, and Marshall SW. Misclassification of overdose events in the X:BOT study. en. The Lancet 2023 Aug; 402:526–7. DOI: `10.1016/S0140-6736(23)00113-7`. Available from: `https://linkinghub.elsevier.com/retrieve/pii/S0140673623001137` [Accessed on: 2023 Sep 18]

11. Lee JD, Nunes EV, Van Veldhuisen P, Lindblad R, and Rotrosen J. Misclassification of overdose events in the X:BOT study – Authors' reply. en. The Lancet 2023 Aug; 402:527–8. DOI: `10.1016/S0140-6736(23)00049-1`. Available from: `https://linkinghub.elsevier.com/retrieve/pii/S0140673623000491` [Accessed on: 2023 Sep 18]

12. Rudolph KE, Russell M, Luo SX, Rotrosen J, and Nunes EV. Under-representation of key demographic groups in opioid use disorder trials. en. Drug and Alcohol Dependence Reports 2022 Sep; 4:100084. DOI: `10.1016/j.dadr.2022.100084`. Available from: `https://linkinghub.elsevier.com/retrieve/pii/S2772724622000592` [Accessed on: 2022 Aug 18]

13. Rudolph KE, Díaz I, Luo SX, Rotrosen J, and Nunes EV. Optimizing opioid use disorder treatment with naltrexone or buprenorphine. en. Drug and Alcohol Dependence 2021 Nov; 228:109031. DOI: `10.1016/j.drugalcdep.2021.109031`. Available from: `https://linkinghub.elsevier.com/retrieve/pii/S0376871621005263` [Accessed on: 2022 Nov 20]

14. Nunes EV, Scodes JM, Pavlicova M, Lee JD, Novo P, Campbell AN, and Rotrosen J. Sublingual Buprenorphine-Naloxone Compared With Injection Naltrexone for Opioid Use Disorder: Potential Utility of Patient Characteristics in Guiding Choice of Treatment. en. American Journal of Psychiatry 2021 Jul; 178:660–71. DOI: `10.1176/appi.ajp.2020.20060816`. Available from: `http://ajp.psychiatryonline.org/doi/10.1176/appi.ajp.2020.20060816` [Accessed on: 2023 Dec 6]

15. Morgan JR, Schackman BR, Leff JA, Linas BP, and Walley AY. Injectable naltrexone, oral naltrexone, and buprenorphine utilization and discontinuation among individuals treated for opioid use disorder in a United States commercially insured population. en. Journal of Substance Abuse Treatment 2018 Feb; 85:90–6. DOI: `10.1016/j.jsat.2017.07.001`. Available from: `https://linkinghub.elsevier.com/retrieve/pii/S0740547216304135` [Accessed on: 2023 Jul 27]

16. Morgan JR, Schackman BR, Weinstein ZM, Walley AY, and Linas BP. Overdose following initiation of naltrexone and buprenorphine medication treatment for opioid use disorder in a United States commercially insured cohort. en. Drug and Alcohol Dependence 2019 Jul; 200:34–9. DOI: `10.1016/j.drugalcdep.2019.02.031`. Available from: `https://linkinghub.elsevier.com/retrieve/pii/S0376871619301310` [Accessed on: 2023 Sep 18]

17. Friedman J and Shover CL. Charting the fourth wave: Geographic, temporal, race/ethnicity and demographic trends in polysubstance fentanyl overdose deaths in the United States, 2010–2021. en. Addiction 2023 Dec; 118:2477–85. DOI: `10.1111/add.16318`. Available from: `https://onlinelibrary.wiley.com/doi/10.1111/add.16318` [Accessed on: 2024 Jan 3]

18. Larochelle MR, Bernson D, Land T, Stopka TJ, Wang N, Xuan Z, Bagley SM, Liebschutz JM, and Walley AY. Medication for Opioid Use Disorder After Nonfatal Opioid Overdose and Association With Mortality: A Cohort Study. en. Annals of Internal Medicine 2018 Aug; 169:137. DOI: `10.7326/M17-3107`. Available from: `http://annals.org/article.aspx?doi=10.7326/M17-3107` [Accessed on: 2023 Sep 18]

19. Zhang P, Tossone K, Ashmead R, Bickert T, Bailey E, Doogan NJ, Mack A, Schmidt S, and Bonny AE. Examining differences in retention on medication for opioid use disorder: An analysis of Ohio Medicaid data. en. Journal of Substance Abuse Treatment 2022 May; 136:108686. DOI: `10.1016/j.jsat.2021.108686`. Available from: `https://linkinghub.elsevier.com/retrieve/pii/S0740547221004128` [Accessed on: 2023 Jun 26]

20. Lund JL, Richardson DB, and Stürmer T. The Active Comparator, New User Study Design in Pharmacoepidemiology: Historical Foundations and Contemporary Application. Pharmacoepidemiology and Drug Safety 2015; 2:221–8. DOI: `10.1007/s40471-015-0053-5`

21. Hernán MA, Wang W, and Leaf DE. Target Trial Emulation: A Framework for Causal Inference From Observational Data. en. JAMA 2022 Dec; 328:2446. DOI: `10.1001/jama.2022.21383`. Available from: `https://jamanetwork.com/journals/jama/fullarticle/2799678` [Accessed on: 2022 Dec 27]

22. Opioid Prescription Data Resources. Available from: `https://www.cdc.gov/opioids/data-resources/index.html` [Accessed on: 2023 Jun 27]

23. Opioid Use Disorder (OUD) 1: Overarching Opioid Use Disorder. Available from: `https://www2.ccwdata.org/documents/10280/19140001/oth-cond-algo-oud.pdf` [Accessed on: 2023 Jul 7]

24. Athena Vocabularies Repository. OHDSI: Observational Health Data Sciences and Informatics. Available from: `https://athena.ohdsi.org/search-terms/start` [Accessed on: 2023 Jun 15]

25. Busch AB, Kennedy-Hendricks A, Schilling C, Stuart EA, Hollander M, Meiselbach MK, Barry CL, Huskamp HA, and Eisenberg MD. Measurement Approaches to Estimating Methadone Continuity in Opioid Use Disorder Care. en. Medical Care 2023 May; 61:314–20. DOI: `10.1097/MLR.0000000000001838`. Available from: `https://journals.lww.com/10.1097/MLR.0000000000001838` [Accessed on: 2023 Jul 14]

26. Suissa S. Immortal Time Bias in Pharmacoepidemiology. en. American Journal of Epidemiology 2008 May; 167:492–9. DOI: `10.1093/aje/kwm324`. Available from: `https://academic.oup.com/aje/article-lookup/doi/10.1093/aje/kwm324`

27. Samples H, Williams AR, Olfson M, and Crystal S. Risk factors for discontinuation of buprenorphine treatment for opioid use disorders in a multi-state sample of Medicaid enrollees. en. Journal of Substance Abuse Treatment 2018 Dec; 95:9–17. DOI: `10.1016/j.jsat.2018.09.001`. Available from: `https://linkinghub.elsevier.com/retrieve/pii/S074054721830134X` [Accessed on: 2023 Jun 20]

28. Vivolo-Kantor A, Pasalic E, Liu S, Martinez PD, and Gladden RM. Defining indicators for drug overdose emergency department visits and hospitalisations in ICD-10-CM coded discharge data. en. Injury Prevention 2021 Mar; 27:i56–i61. DOI: `10.1136/injuryprev-2019-043521`. Available from: `https://injuryprevention.bmj.com/lookup/doi/10.1136/injuryprev-2019-043521` [Accessed on: 2023 Oct 17]

29. Williams N and Díaz I. lmtp: An R Package for Estimating the Causal Effects of Modified Treatment Policies. en. Observational Studies 2023; 9:103–22. DOI: `10.1353/obs.2023.0019`. Available from: `https://muse.jhu.edu/article/883479` [Accessed on: 2023 Sep 26]

30. Van Der Laan MJ, Polley EC, and Hubbard AE. Super Learner. en. Statistical Applications in Genetics and Molecular Biology 2007; 6:Article 25. DOI: `10.2202/1544-6115.1309`

31. Daniel RM. Double Robustness. en. *Wiley StatsRef: Statistics Reference Online*. Ed. by Balakrishnan N, Colton T, Everitt B, Piegorsch W, Ruggeri F, and Teugels JL. Chichester, UK: John Wiley & Sons, Ltd, 2018 May :1–14. Available from: `http://doi.wiley.com/10.1002/9781118445112.stat08068`

32. Pharmacotherapy for Opioid Use Disorder (POD). Available from: `https://www.ncqa.org/hedis/measures/pharmacotherapy-for-opioid-use-disorder/` [Accessed on: 2024 Jan 10]

33. Quality ID #468: Continuity of Pharmacotherapy for Opioid Use Disorder (OUD). Available from: `https://qpp.cms.gov/docs/QPP_quality_measure_specifications/CQM-Measures/2023_Measure_468_MIPSCQM.pdf` [Accessed on: 2023 Oct 17]

34. Hernán MA and Hernández-Díaz S. Beyond the intention-to-treat in comparative effectiveness research. en. Clinical Trials: Journal of the Society for Clinical Trials 2012 May; 9:48–55. DOI: `10.1177/1740774511420743`. Available from: `http://journals.sagepub.com/doi/10.1177/1740774511420743`

35. Krawczyk N, Mojtabai R, Stuart EA, Fingerhood M, Agus D, Lyons BC, Weiner JP, and Saloner B. Opioid agonist treatment and fatal overdose risk in a state-wide US population receiving opioid use disorder services. en. Addiction 2020 Sep; 115:1683–94. DOI: `10.1111/add.14991`. Available from: `https://onlinelibrary.wiley.com/doi/10.1111/add.14991` [Accessed on: 2023 Nov 30]

36. Hernán MA. The Hazards of Hazard Ratios. en. Epidemiology 2010 May; 21:13–5. DOI: `10.1097/EDE.0b013e3181c1ea43`. Available from: `https://insights.ovid.com/crossref?an=00001648-201001000-00004`

37. Murray EJ, Caniglia EC, Swanson SA, Hernández-Díaz S, and Hernán MA. Patients and investigators prefer measures of absolute risk in subgroups for pragmatic randomized trials. Journal of Clinical Epidemiology 2018; 103:10–21. DOI: `10.1016/j.jclinepi.2018.06.009`

38. Freeman AL and Spiegelhalter DJ. Communicating health risks in science publications: Time for everyone to take responsibility. BMC Medicine 2018; 16. Publisher: BMC Medicine:5–8. DOI: `10.1186/s12916-018-1194-4`

39. Medicaid's Role in Addressing the Opioid Epidemic. Available from: `https://www.kff.org/infographic/medicaids-role-in-addressing-opioid-epidemic/` [Accessed on: 2023 Dec 12]

Tables

Table 1. Baseline characteristics of study cohort, overall and stratified by medication

Characteristic	N(%)		
	Medication		
	XR-NTX N = 1755	**SL-BUP** N = 9886	**Overall** N = 11641
Demographics			
Age[a]	33 (28, 42)	35 (29, 46)	34 (29, 45)
Male	1,015 (58%)	5,460 (55%)	6,475 (56%)
Race/ethnicity			
White, non-Hispanic	1,130 (69%)	6,561 (71%)	7,691 (70%)
Black, non-Hispanic	139 (8.5%)	835 (9.0%)	974 (8.9%)
Hispanic, all races	332 (20%)	1,587 (17%)	1,919 (18%)
Asian, non-Hispanic	24 (1.5%)	130 (1.4%)	154 (1.4%)
AIAN, non-Hispanic	16 (1.0%)	161 (1.7%)	177 (1.6%)
Hawaiian/Pacific Islander	2 (0.1%)	18 (0.2%)	20 (0.2%)
Unknown	112	594	706
TANF benefits	28 (1.6%)	334 (3.4%)	362 (3.1%)
State			
California	908 (52%)	7,767 (79%)	8,675 (75%)
New Jersey	847 (48%)	2,119 (21%)	2,966 (25%)
OUD characteristics			
Inpatient hospitalization[b]	273 (16%)	465 (4.7%)	738 (6.3%)
Psychosocial/behavioral therapy	942 (54%)	3,352 (34%)	4,294 (37%)
Overdose	197 (11%)	742 (7.5%)	939 (8.1%)
Other substance use disorders			
Alcohol	686 (39%)	1,030 (10%)	1,716 (15%)
Cannabis	194 (11%)	507 (5.1%)	701 (6.0%)
Sedatives	181 (10%)	491 (5.0%)	672 (5.8%)
Cocaine	183 (10%)	360 (3.6%)	543 (4.7%)
Amphetamine	297 (17%)	1,403 (14%)	1,700 (15%)
Other[c]	502 (29%)	2,013 (20%)	2,515 (22%)
Psychiatric comorbid conditions			
Depression	634 (36%)	2,284 (23%)	2,918 (25%)
Schizophrenia/psychosis	191 (11%)	613 (6.2%)	804 (6.9%)
Anxiety	611 (35%)	2,635 (27%)	3,246 (28%)
Bipolar	269 (15%)	967 (9.8%)	1,236 (11%)
ADD/ADHD	67 (3.8%)	254 (2.6%)	321 (2.8%)
PTSD	162 (9.2%)	643 (6.5%)	805 (6.9%)
Other	301 (17%)	1,234 (12%)	1,535 (13%)
Other comorbid conditions			
Disability[d]	102 (5.8%)	1,450 (15%)	1,552 (13%)
Chronic pain	183 (10%)	2,228 (23%)	2,411 (21%)

Medications			
Benzodiazepines	518 (30%)	3,064 (31%)	3,582 (31%)
Antidepressants	1,018 (58%)	4,152 (42%)	5,170 (44%)
Antipsychotics	550 (31%)	1,903 (19%)	2,453 (21%)
Stimulants	64 (3.6%)	351 (3.6%)	415 (3.6%)
Gabapentinoids	575 (33%)	2,563 (26%)	3,138 (27%)
Alpha-2 agonists	304 (17%)	1,762 (18%)	2,066 (18%)
Healthcare utilization			
Inpatient hospitalization	590 (34%)	1,916 (19%)	2,506 (22%)
Emergency department encounter	1,168 (67%)	5,830 (59%)	6,998 (60%)
Other services encounters, count[a]	23 (7, 56)	11 (3, 28)	12 (3, 31)

Abbreviations: XR-NTX, extended release naltrexone; SL-BUP, oral buprenorphine; AIAN, American Indian or Alaskan Native; TANF, Temporary Assistance for Needy Families; OUD, opioid use disorder; ADD/ADHD, attention-deficit disorder/attention deficit hyperactivity disorder; PTSD, post-traumatic stress disorder

[a]Median (interquartile range)

[b]With primary diagnosis code of opioid dependence or overdose

[c]Excluding nicotine

[d]Reason for eligibility for Medicaid insurance

Table 2. Primary analysis results: crude and adjusted risks (as %) and risk differences (in percentage points)

| | Crude (95% CI) | | | Adjusted (95% CI) | | |
| | Risks | | | Risks | | |
Week	XR-NTX	SL-BUP	Difference	XR-NTX	SL-BUP	Difference
Discontinuation of all MOUD or death						
9	36.1 (33.8, 38.4)	31.0 (30.1, 32)	5.1 (1.9, 8.3)	33.7 (31.9, 35.4)	30.7 (29.7, 31.7)	3.0 (1.0, 5.0)
13	45.1 (42.7, 47.5)	42.7 (41.6, 43.7)	2.5 (-1.0, 5.9)	43.0 (40.5, 45.5)	42.4 (41.4, 43.5)	0.6 (-2.2, 3.3)
18	60.0 (57.6, 62.4)	53.0 (51.9, 54.0)	7.0 (3.6, 10.4)	59.3 (56.9, 61.7)	53.0 (52.0, 54.1)	6.3 (3.6, 8.9)
22	69.5 (67.3, 71.8)	58.0 (56.9, 59.0)	11.6 (8.3, 14.9)	69.7 (67.5, 71.9)	58.1 (57.1, 59.2)	11.5 (9.1, 14.0)
26	75.1 (72.9, 77.2)	61.8 (60.8, 62.8)	13.3 (10.1, 16.5)	76.4 (75.1, 77.6)	62.2 (61.2, 63.3)	14.1 (12.5, 15.7)
Discontinuation of initial medication or death						
9	37.6 (35.3, 40.0)	32.9 (31.9, 33.8)	4.7 (1.4, 8.0)	34.4 (31.9, 36.9)	32.7 (31.7, 33.6)	1.8 (-0.9, 4.4)
13	47.0 (44.6, 49.4)	44.7 (43.7, 45.7)	2.3 (-1.2, 5.7)	44.7 (42.2, 47.2)	44.6 (43.6, 45.7)	0.1 (-2.6, 2.8)
18	62.4 (60.0, 64.8)	55.0 (54.0, 56.1)	7.4 (4.0, 10.8)	62.0 (60.5, 63.6)	55.2 (54.1, 56.2)	6.9 (5.0, 8.7)
22	71.7 (69.4, 73.9)	60.1 (59.1, 61.2)	11.5 (8.3, 14.8)	72.8 (71.5, 74.2)	60.4 (59.4, 61.4)	12.4 (10.7, 14.1)
26	77.3 (75.2, 79.4)	63.8 (62.8, 64.9)	13.5 (10.4, 16.6)	79.4 (78.2, 80.6)	64.4 (63.4, 65.4)	15.0 (13.5, 16.6)
Overdose or death						
9	1.2 (0.8, 1.8)	1.1 (0.9, 1.3)	0.1 (-0.6, 0.8)	1.1 (0.8, 1.4)	1.2 (1.0, 1.4)	-0.1 (-0.5, 0.3)
13	2.3 (1.7, 3.2)	1.7 (1.5, 2.0)	0.6 (-0.4, 1.6)	1.9 (1.5, 2.3)	1.9 (1.6, 2.2)	0 (-0.5, 0.5)
18	3.6 (2.8, 4.6)	2.3 (2.0, 2.7)	1.2 (0.0, 2.4)	2.7 (2.2, 3.2)	2.5 (2.2, 2.9)	0.2 (-0.4, 0.8)
22	4.9 (3.9, 6.1)	2.8 (2.4, 3.1)	2.1 (0.7, 3.5)	3.7 (3.2, 4.3)	3.0 (2.6, 3.4)	0.7 (0.1, 1.4)
26	5.1 (4.1, 6.3)	3.1 (2.7, 3.4)	2.1 (0.6, 3.5)	3.8 (2.9, 4.7)	3.3 (2.9, 3.7)	0.5 (-0.5, 1.5)

Abbreviations: XR-NTX, extended release naltrexone; SL-BUP, oral buprenorphine; MOUD, medications for opioid use disorder
[a]Week 9: days 57-63; week 13: days 85-91; week 18: days 120-126; week 22: days 148-154; week 26: days 176-180

Table 3. Additional analyses results: risk difference (in percentage points, with 95% confidence interval) at the end of follow-up comparing XR-NTX and SL-BUP

Stratified analyses, outcome: discontinuation of all MOUD or death[a]	
Alcohol use disorder, yes	14.9 (10.3, 19.6)
Alcohol use disorder, no	15.1 (13.5, 16.7)
Chronic pain, yes	11.2 (6.8, 15.6)
Chronic pain, no	13.7 (11.9, 15.4)

Sensitivity analyses[b]	
Outcome: discontinuation of all MOUD or death	
ATE	12.8 (9.8, 15.8)
Among XR-NTX treated	12.9 (10.2, 15.7)
Among SL-BUP treated	13.0 (9.6, 16.3)
In optimally trimmed cohort[c]	12.9 (10.3, 15.6)
Include LTFU in outcome	12.2 (9.2, 15.2)
14-day grace period	10.4 (7.7, 13.0)
Outcome: overdose or death	
ATE	0.5 (-0.8, 1.9)
Among XR-NTX treated	1.1 (-0.7, 2.8)
Among SL-BUP treated	0.5 (-0.1, 2.0)
In optimally trimmed cohort[c]	1.1 (-0.6, 2.7)
More sensitive definition[b,d]	0.9 (-0.6, 2.4)

[a] Using survival analysis methods

[b] Analysis at the end of follow-up (not using survival analysis methods)

[c] After trimming, n=6,924 (XR-NTX n=1725, SL-BUP n=5199)

[d] After exclusion of early overdose events with new definition, n=11,630 (XR-NTX n=1752, SL-BUP n=9878)

Figures

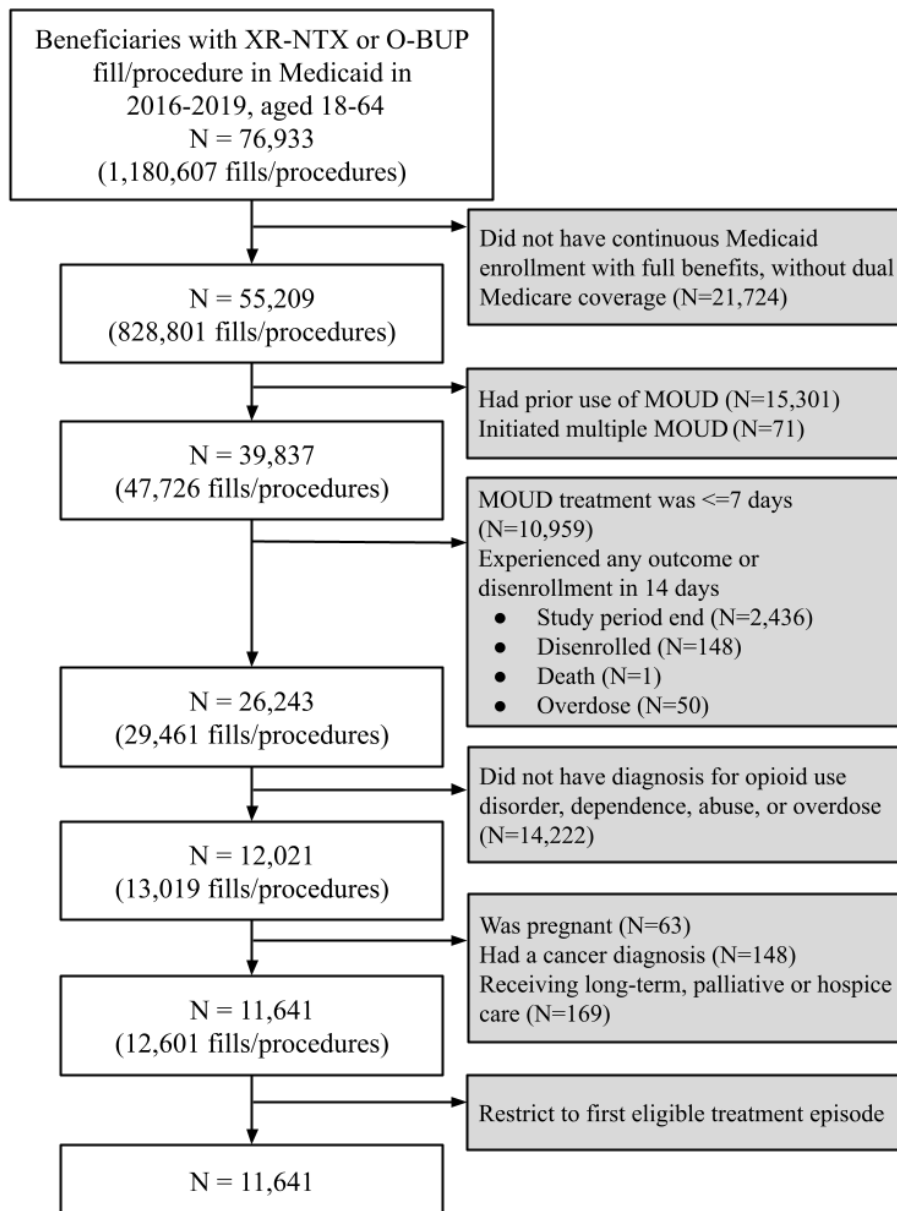**Figure 1**. Flowchart of cohort creation

Figure 2. Analysis results for composite of discontinuation of all MOUD and death: A. Crude risk. B. Adjusted risk. C. Adjusted risk difference. Shaded areas are 95% confidence intervals. Abbreviations: MOUD, medications for opioid use disorder; XR-NTX, extended release naltrexone; SL-BUP, sublingual buprenorphine
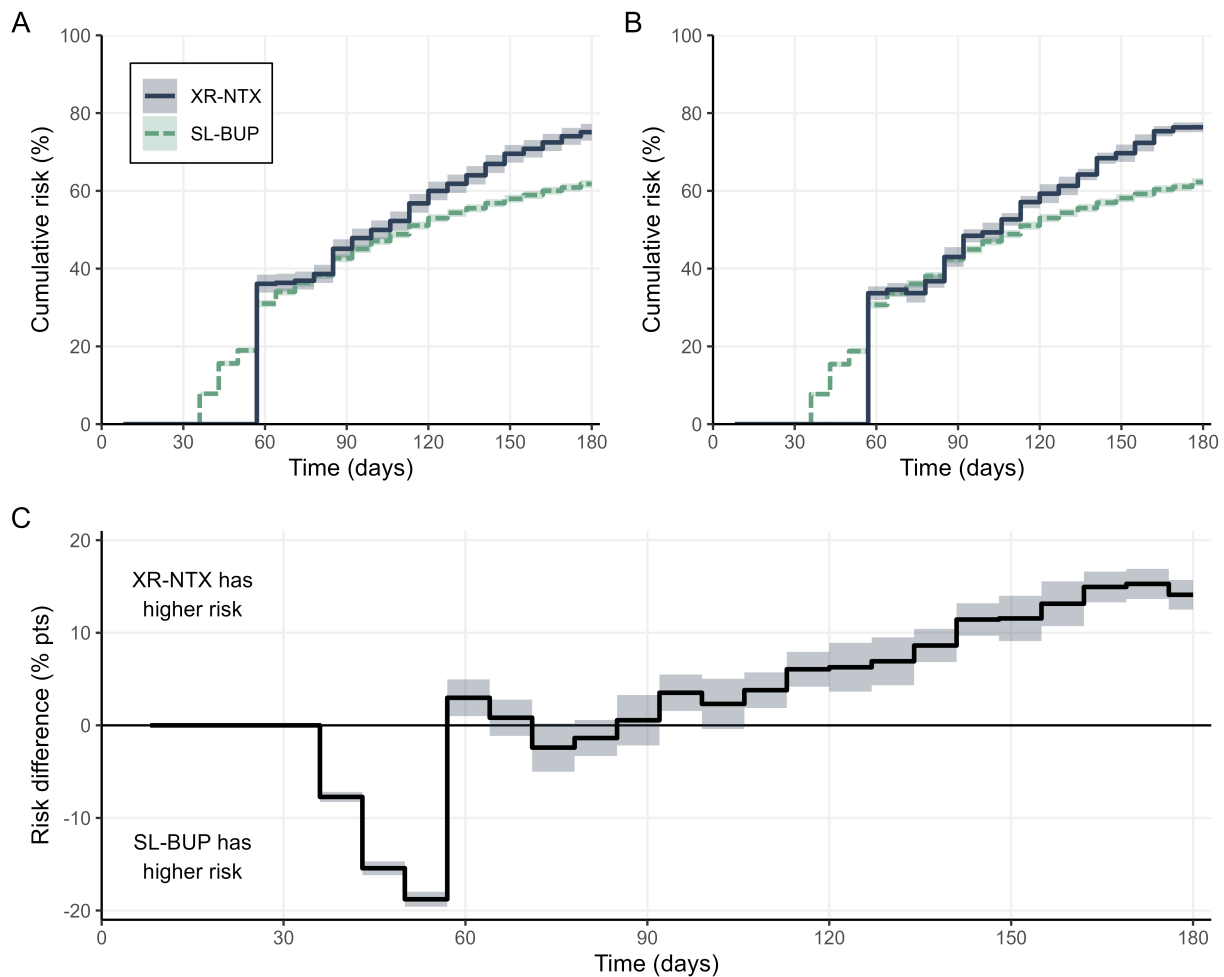
Figure 3. Analysis results for composite of overdose and death: A. Crude risk. B. Adjusted risk. C. Adjusted risk difference. Shaded areas are 95% confidence intervals. Abbreviations: XR-NTX, extended release naltrexone; SL-BUP, sublingual buprenorphine