Original Investigation

Peer Comparison or Guideline-Based Feedback and Postsurgery Opioid Prescriptions
A Randomized Clinical Trial

Zachary Wagner, PhD; Allison Kirkegaard, MTS; Louis T. Mariano, PhD; Jason N. Doctor, PhD; Xiaowei Yan, PhD; Stephen D. Persell, MD, MPH; Noah J. Goldstein, PhD; Craig R. Fox, PhD; Chad M. Brummett, MD; Robert J. Romanelli, PhD, MPH; Kathryn Bouskill, PhD, MPH; Meghan Martinez, MPH; Kyle Zanocco, MD; Daniella Meeker, PhD; Satish Mudiganti, MS; Jennifer Waljee, MD; Katherine E. Watkins, MD, MSHS

Abstract

IMPORTANCE Excess opioid prescribing after surgery can result in prolonged use and diversion. Email feedback based on social norms may reduce the number of pills prescribed.

OBJECTIVE To assess the effectiveness of 2 social norm-based interventions on reducing guideline-discordant opioid prescribing after surgery.

DESIGN, SETTING, AND PARTICIPANTS This cluster randomized clinical trial conducted at a large health care delivery system in northern California between October 2021 and October 2022 included general, obstetric/gynecologic, and orthopedic surgeons with patients aged 18 years or older discharged to home with an oral opioid prescription.

INTERVENTIONS In 19 hospitals, 3 surgical specialties (general, orthopedic, and obstetric/gynecologic) were randomly assigned to a control group or 1 of 2 interventions. The guidelines intervention provided email feedback to surgeons on opioid prescribing relative to institutionally endorsed guidelines; the peer comparison intervention provided email feedback on opioid prescribing relative to that of peer surgeons. Emails were sent to surgeons with at least 2 guideline-discordant prescriptions in the previous month. The control group had no intervention.

MAIN OUTCOME AND MEASURES The probability that a discharged patient was prescribed a quantity of opioids above the guideline for the respective procedure during the 12 intervention months.

RESULTS There were 38,235 patients discharged from 640 surgeons during the 12-month intervention period. Control-group surgeons prescribed above guidelines 36.8% of the time during the intervention period compared with 27.5% and 25.4% among surgeons in the peer comparison and guidelines arms, respectively. In adjusted models, the peer comparison intervention reduced guideline-discordant prescribing by 5.8 percentage points (95% CI, −10.5 to −1.1; P = .03) and the guidelines intervention reduced it by 4.7 percentage points (95% CI, −9.4 to −0.1; P = .05). Effects were driven by surgeons who performed more surgeries and had more guideline-discordant prescribing at baseline. There was no significant difference between interventions.

CONCLUSIONS AND RELEVANCE In this cluster randomized clinical trial, email feedback based on either guidelines or peer comparison reduced opioid prescribing after surgery. Guideline-based feedback was as effective as peer comparison-based feedback. These interventions are simple, low-cost, and scalable, and may reduce downstream opioid misuse.

TRIAL REGISTRATION ClinicalTrials.gov NCT05070338

Key Points

Question Can email feedback on opioid prescribing informed by social norms reduce guideline-discordant postoperative opioid prescribing?

Findings In this 3-arm cluster-randomized clinical trial that included 640 surgeons, 2 forms of social norm-based email feedback sent to surgeons with postoperative opioid prescriptions above guideline-recommended amounts significantly reduced guideline-discordant prescribing.

Meaning Email feedback based on social norms is an effective and light-touch intervention for increasing guideline-concordant opioid prescribing after surgery.

+ Visual Abstract
+ Supplemental content

Author affiliations and article information are listed at the end of this article.

Open Access. This is an open access article distributed under the terms of the CC-BY License.
Introduction

Opioids prescribed after surgery are a significant driver of chronic opioid use and opioid use disorder.1-4 Postoperative opioid prescriptions also contribute to unused pills available for diversion to the community5,6; more than half of the opioids prescribed are never used by the patient and rarely disposed of safely.7,9 High-risk opioid prescribing is concentrated among surgeons who write the most prescriptions, with the top 5% of surgeons in terms of prescribing accounting for 40% of prescriptions written for large quantities.10 Despite widespread efforts in recent years to match opioid prescribing to patient need, excessive postoperative opioid prescribing has persisted.8,9

Providing surgeons with feedback on the quantity of opioids they prescribe relative to a social norm may help reduce opioid prescription quantities by challenging what is considered acceptable behavior.11,12 This could take the form of feedback that describes peer prescribing practices relative to an individual’s own (descriptive social norm) or feedback that provides information about opioid prescription quantities recommended by an authoritative source such as a guideline (injunctive social norm). Although previous randomized clinical trials have shown that social norms can be used to reduce inappropriate antibiotic prescribing among primary care clinicians13 and opioid prescribing in emergency departments,14 there is limited high-quality evidence on whether feedback based on social norms can influence opioid prescribing in the surgical setting, which has a different structure of care and has the second highest opioid prescribing rate of any specialty after pain medicine.15,16 Moreover, although recent research has examined the influence of either peer-based or guideline-based feedback (or both) on clinician behaviors,17-21 little is known regarding how these directly compare to one another.

To address these gaps, this study tests the effectiveness of 2 email feedback interventions relative to usual care across 3 surgical specialties in a multihospital cluster-randomized trial: (1) a descriptive social norm intervention that compares opioid prescription quantities of individual surgeons with that of their peers and (2) an injunctive social norm intervention that compares opioid prescription quantities to guideline recommendations.

Methods

Study Design and Oversight

This was a 3-arm cluster-randomized clinical trial conducted October 19, 2021, to October 18, 2022, in 3 surgical specialties (general, orthopedic, and obstetrics/gynecology) at 19 hospitals within Sutter Health, a large health care system in northern California. Details of the intervention and analysis plan were prespecified in a published protocol and are also available in Supplement 1.22 Activities were approved by the institutional review boards at Sutter Health and the RAND Corporation and monitored by an independent data safety and monitoring board. All data were obtained from the Sutter Health’s electronic health record (EHR) database. Gender and ethnic and racial categories were self-reported by patients; other racial category includes American Indian/Alaska Native, Native Hawaiian/Pacific Islander, other race, and multiple races. This study followed the Consolidated Standards of Reporting Trials (CONSORT) reporting guideline.

Inclusion and Enrollment

All hospitals within the Sutter Health system as of May 2021 were enrolled in the study. To avoid spillover effects, facilities identified as regularly sharing surgical staff were grouped together, resulting in 19 hospitals for the analysis. Surgeons working exclusively in 1 of the 3 surgical specialties were eligible to participate in the study if they performed any surgical procedures for which opioid prescribing guidelines were available (guidelines described herein) at any of the 19 study hospitals between June 1, 2020, and May 31, 2021. Surgeons were included in the study if they also performed at least 1 surgical procedure for which opioid prescribing guidelines were available during the intervention period.
Opioid Prescribing Guidelines
Both interventions as well as our primary outcome were based on prescribing guidelines developed by teams at the Mayo Clinic.23-26 These guidelines recommend procedure-specific ceilings for the total quantity of 5-mg oxycodone tablets needed over the postoperative recovery period (eTable 1 in Supplement 2). We converted these tablet quantities to morphine milligram equivalent (MME) quantities using the Centers for Disease Control and Prevention's conversion table27 so that we could evaluate a prescription's guideline compliance regardless of the type or strength of opioid prescribed.

Intervention and Control Conditions
Both interventions provided feedback to surgeons if at least 2 of their patients were discharged with opioid prescriptions larger than the guideline-recommended amount in the prior month. Feedback was sent via emails from and signed by the surgeon's department chair, chief medical executive, or chief of staff. Surgeons were not explicitly informed they were part of a study. We chose not to obtain informed consent from participants to avoid a Hawthorn effect and a waiver of informed consent was approved by the institutional review boards. However, hospital administrators could opt out of the intervention and surgeons could opt out of emails.

Although the content of the emails was distinct, the trigger for sending an email was identical: a surgeon having at least 2 patients with eligible discharge opioid prescriptions over guideline recommendations in the previous month. Eligible discharge prescriptions were those for adults discharged to home after a single surgical procedure with an applicable postoperative prescribing guideline, and where the prescription was for an oral medication. The email was sent to the operating surgeon even if the surgeon did not write the discharge prescriptions; 44% of discharge opioid prescriptions were written by someone other than the operating surgeon (usually a hospitalist) although the operating surgeon discharged the patient 84% of the time.

Peer Comparison (Descriptive Norm) Intervention
The subject line for the peer comparison email read “Your peers vs your opioid prescribing safety record.” The email content informed surgeons that their patients were prescribed opioid quantities exceeding the amounts prescribed by more than a particular percentage of their peers (eFigure 1 in Supplement 2). The peer percentage was calculated as the share of surgeons at Sutter Health in their specialty who had fewer than 2 patients discharged with opioid quantities exceeding guidelines in the past month, and ranged from 28% to 87% (mean, 61%) during the study period. The email also listed the procedures for which they prescribed higher quantities than their peers and the range prescribed by their peers for each procedure.

Guideline Feedback (Injunctive Norm) Intervention
The subject line of the guideline feedback email read “Best practice guidelines vs your opioid prescribing safety record.” The email content informed the surgeon that their patients were prescribed opioid quantities “exceeding the amounts recommended by safety guidelines for these procedures” (eFigure 2 in Supplement 2). The email also listed the procedures for which this occurred in the prior month and the guideline range for each procedure.

Control Condition
Surgeons in the control arm did not receive any emails and were not informed that they were being studied. We made this decision to avoid Hawthorn effects.

Implementation
Each month, intervention emails were automatically created and sent to eligible surgeons via an email management platform. Sutter Health staff were responsible ensuring emails were sent out correctly.
Randomization
Randomization occurred at the level of the surgical specialty. At each hospital, 1 specialty was randomized to each arm, stratified by hospital size and specialty. The allocation of specialties within hospitals to study arms was drawn using the sample function in R statistical software (version 4.3.2, R Foundation) applied to each block. We chose to randomize by specialty within hospital rather than at the hospital level to maximize the number of randomization clusters, which increases statistical power, and to avoid chance imbalance which is more likely with fewer randomization units. After randomization but before the start of the intervention, 1 hospital-specialty combination refused to participate in the peer comparison intervention owing to a hospital administrator’s concerns. To preserve the integrity of the randomized design, we treat the 43 discharges (<1% of sample) from this group as peer comparison although they did not receive the intervention.

Outcome Measures
The primary prespecified outcome was the probability that a discharged patient was prescribed a quantity of opioids above the guideline for the respective procedure during the 12 intervention months. We extracted all initial discharge prescriptions from the EHR and defined a prescription as being above guidelines if the total MME quantity of opioids prescribed was above the ceiling of the guideline-recommended range. Surgeries where no opioids were prescribed were considered within guidelines.

We analyzed 2 prespecified secondary opioid prescribing outcomes: (1) the MME quantity prescribed at discharge, coded as zero when no opioid was prescribed and truncated at 500 MMEs (the 99th percentile), and (2) the probability that a discharged patient received any opioid prescription.

We also analyzed 3 prespecified outcomes that assess potential harmful effects of the interventions: (1) whether the patient got an additional opioid prescription filled in the 30 days after the surgery, (2) whether the patient had an emergency department visit in the 30 days after the surgery, and (3) whether the patient was hospitalized in the 30 days after the surgery.

Statistical Analysis
We conducted 3 prespecified analyses. In each, we estimated intervention effects using the surgeon’s treatment assignment, regardless of whether the surgeon received an email during the study period. The first set of analyses modeled effects on primary and secondary outcomes at the level of the surgical specialty using a hierarchical linear model (HLM) with random effects for the surgeon and specialty within hospital and controlling for the baseline levels of the outcome (averaged by surgeon in the 24 months prior to the intervention). We included a version of the results controlling for a more comprehensive set of covariates in the eAppendix in Supplement 2. We used a logit-link for binary outcomes and reported absolute risk differences using recycled predictions. The second analysis estimated intervention effects over time by using separate HLMs for each of the 12 study months. The third set of analyses estimated effects for prespecified subgroups of surgeons based on 3 surgeon characteristics: (1) effects by surgical procedure volume during the intervention period, (2) effects by baseline guideline-discordant opioid prescribing in the 24 months before the intervention, and (3) effects by surgical specialty.

We also conducted 2 exploratory analyses. First, we estimated average effects at the level of the discharge using ordinary least squares (OLS) adjusted for the baseline level of the outcome and with standard errors clustered by randomization cluster. This discharge-level model weights each discharge equally and captures the average effect per discharged patient. Conversely, discharges from higher-volume surgeons tend to be downweighted in the HLM owing to their tendency toward higher variation in the outcome. Second, we used quantile regression to estimate the effect of the interventions on each decile of MMEs prescribed to determine whether findings were robust with a nonnormal distribution of MMEs prescribed.
The study was powered to detect a difference of 12 percentage points between each study arm and the control group. Baseline outcome measures were more strongly correlated with the outcome than the original conservative estimate, yielding additional power. All statistical analyses are described in more detail in the eAppendix in Supplement 2.

Multiple Hypothesis Testing
We used false discovery rate (FDR) corrections to account for simultaneously tested hypotheses of the effect of both treatments. All reported P values have been adjusted for an FDR of .05.

Results
Of the 778 surgeons eligible for inclusion in the study, 640 surgeons had 38,235 eligible discharges during the intervention period (Figure 1). Surgeon and patient characteristics were mostly balanced across arms (Table 1). The 3 arms were also balanced on the share of prescriptions above guideline quantities (primary outcome) and the share of discharges with any opioid prescribing in the 24 months leading up to the intervention. Although the guidelines arm had higher mean MMEs at baseline (100) relative to the other 2 arms (70 in control and 75 in peer comparison), median (IQR) MMEs were more balanced: 61 in guidelines (31-134), 57 in control (33-92), and 72 in peer comparison (35-99). More than 50% of total joint replacement procedures at baseline occurred in clusters assigned to guidelines, driving the guideline mean MME higher.

Intervention Effects on Above-Guideline Opioid Prescriptions
In the 12 months prior to the intervention, the share of discharges with an opioid prescription above guideline quantities was around 35% to 40% in all 3 arms (Figure 2A). During the intervention period, above guideline prescribing was 27.5% on average in the peer comparison arm and 25.4% in the guidelines arm compared with 36.8% in the control arm (Table 2). In adjusted hierarchical models (Table 2), the peer comparison intervention reduced the share of discharges with prescriptions above guideline quantities by 5.8 percentage points (95% CI, −10.5 to −1.1; P = .03) and the guidelines intervention reduced it by 4.7 percentage points (95% CI, −9.4 to −0.1; P = .05) compared with the control group. These constitute 17% and 14% relative reductions. Effect sizes

Figure 1. Consort Flow Diagram

![Consort Flow Diagram](image-url)
were not significantly different between the 2 intervention arms. Figure 2B shows that effects became more pronounced over time in both arms, with 9 to 13 percentage point reductions in each of the past 4 months. Results were similar when regressions controlled for a more comprehensive set of covariates (Table E3 in Supplement 2).

The effects among surgeons who performed above the median number of procedures during the intervention period were 8.8 and 9.5 percentage point reductions for peer comparison and guideline interventions, respectively (P < .001 for both; eTables 4 and 5 in Supplement 2; Figure 3), compared with 3.4 and 1.0 among surgeons with below-median numbers of procedures (P = .73 and P = .80, respectively. P values on difference in effect were .21 and .10, respectively). The effects among surgeons with above-median guideline-discordant prescribing at baseline were 10.6 and 10.9 percentage point reductions for peer comparison and guideline interventions, respectively (P = .005 for both), compared with 1.8 and 0.8 among surgeons with below-median guideline-discordant prescribing at baseline.

<table>
<thead>
<tr>
<th>Table 1. Sample Characteristicsa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristic</td>
</tr>
<tr>
<td>Total hospitals, No.</td>
</tr>
<tr>
<td>Total surgeons, No.</td>
</tr>
<tr>
<td>Total patient discharges, No.</td>
</tr>
<tr>
<td>Surgeon characteristics</td>
</tr>
<tr>
<td>General specialty</td>
</tr>
<tr>
<td>OB/GYN specialty</td>
</tr>
<tr>
<td>Orthopedic specialty</td>
</tr>
<tr>
<td>Years since receiving degree, mean (SD)</td>
</tr>
<tr>
<td>Gender</td>
</tr>
<tr>
<td>Female</td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Unknown</td>
</tr>
<tr>
<td>Patient characteristics</td>
</tr>
<tr>
<td>Age, mean (SD), y</td>
</tr>
<tr>
<td>Medicaid beneficiary</td>
</tr>
<tr>
<td>Length of stay, mean (SD), h</td>
</tr>
<tr>
<td>Body mass index, mean (SD)b</td>
</tr>
<tr>
<td>Pain condition</td>
</tr>
<tr>
<td>Chronic opioid use</td>
</tr>
<tr>
<td>Opioid use past year</td>
</tr>
<tr>
<td>Opioid use past month</td>
</tr>
<tr>
<td>Opioid use past 24 h</td>
</tr>
<tr>
<td>Genderc</td>
</tr>
<tr>
<td>Female</td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Nonbinary</td>
</tr>
<tr>
<td>Race and ethnicityc</td>
</tr>
<tr>
<td>Asian</td>
</tr>
<tr>
<td>Black</td>
</tr>
<tr>
<td>White</td>
</tr>
<tr>
<td>Other</td>
</tr>
<tr>
<td>Outcomes at baseline (over 24 mo prior to intervention)</td>
</tr>
<tr>
<td>Prescriptions above guidelinesc</td>
</tr>
<tr>
<td>MMEs, mean (SD)d</td>
</tr>
<tr>
<td>MMEs, median (range)</td>
</tr>
<tr>
<td>Any opioid prescription</td>
</tr>
</tbody>
</table>

Abbreviations: MME, morphine milligram equivalent; OB/GYN, obstetrics and gynecology.

aData from Sutter Health’s electronic health records.
bCalculated as weight in kilograms divided by height in meters squared.
cGender and ethnic and racial categories were self-reported by patients; other racial category includes American Indian/Alaska Native, Native Hawaiian/Pacific Islander, other race, and multiple races.
dThe quantity of MMEs was coded as zero when no opioids were prescribed. Continuous variables show mean (SD).
prescribing ($P = .64$ for both; $P$ value on difference in effect was $0.03$ for both). The effect was not significantly different across specialties (eTable 4 in Supplement 2).

In models analyzed at the discharge level, there was an absolute 9.7-percentage point reduction in guideline-discordant prescribing in the peer-comparison arm (95% CI, $-14.5$ to $-4.9$; $P < .001$) and a 10.1 percentage point reduction in the guidelines arm (95% CI, $-15.6$ to $-4.7$; $P = .001$) relative to the control group (eTable 6 in Supplement 2).

### Intervention Effects on MMEs Prescribed

During the intervention period, patients of surgeons in the control arm were prescribed 67.7 MMEs per discharge on average compared with 58.4 in the peer comparison arm and 78.7 in the guidelines arm (Table 2). Adjusted hierarchical models estimated nonsignificant reductions of 4.6 MMEs in the peer comparison arm (95% CI, $-18.8$ to $9.6$; $P = .94$) and 0.5 MMEs in the guidelines arm (95% CI, $-14.4$ to $13.4$; $P = .94$) compared with the control group. The discharge-level model estimated a 12.4 MME reduction in the peer comparison arm (95% CI, $-24.2$ to $-0.6$; $P = .07$) and a 12.2 MME reduction in the guidelines arm (95% CI, $-25.2$ to $0.80$; $P = .07$; eTable 6 in Supplement 2).

### Figure 2. Prescribing Above Guideline Quantities Over Time

- **A.** The share of encounters with an opioid prescription at discharge exceeding guideline-recommended amounts.
- **B.** The effect of the intervention in each study month estimated using 12 separate hierarchical logistic models with random effects for surgeon and specialty within hospital. Intervention effects are absolute percentage point differences. Regression results for panel B are reported in eTable 2 in Supplement 2.

### Table 2. Effect of Social Norms-Based Nudges on Primary and Secondary Outcomes*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean of outcome during 12 intervention months</th>
<th>Effect of peer comparison intervention, percentage points</th>
<th>Effect of guidelines intervention, percentage points</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control (95% CI)</td>
<td>Unadjusted $P$ value</td>
<td>Adjusted (95% CI) $P$ value</td>
</tr>
<tr>
<td>Primary outcome</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Opioid prescriptions above guideline quantities, %</td>
<td>36.8</td>
<td>$-9.3$ ($-20.5$ to $1.9$) $P = .11$</td>
<td>$-5.8$ ($-10.5$ to $-1.1$) $P = .03$</td>
</tr>
<tr>
<td>Secondary outcomes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morphine milligram equivalent, mean</td>
<td>67.7</td>
<td>$-9.3$ ($-12.4$ to $13.8$) $P = .53$</td>
<td>$-4.6$ ($-18.8$ to $9.6$) $P = .94$</td>
</tr>
<tr>
<td>Any opioid prescribed, %</td>
<td>52.0</td>
<td>$-4.6$ ($-16.7$ to $7.5$) $P = .55$</td>
<td>$0.4$ ($-4.1$ to $4.8$) $P = .90$</td>
</tr>
</tbody>
</table>

*Unadjusted models compare means between each intervention arm and the control arm with standard errors clustered by specialty within hospital. Adjusted models include random effects for surgeon and specialty within hospital and control for baseline level of the outcome at the surgeon level. $P$ values are adjusted for multiple testing using a false discovery rate of .05.
Quantile regressions estimated reductions at the median of 14.1 MMEs (95% CI, −21.0 to −7.3; \(P < .001\)) in the peer comparison arm and 13.8 in the guidelines arm (95% CI, −20.8 to −6.7; \(P < .001\)) compared with the control group, and reductions were larger at the higher end of the distribution (eFigure 3 in Supplement 2).

**Intervention Effects on the Probability of an Opioid Being Prescribed**

Overall, 52% of discharges during the intervention period in the control group had any opioid prescription compared with 48% in both intervention arms. Adjusted hierarchical models estimated no significant differences between either intervention arm relative to the control group on the probability of prescribing any opioid (Table 2). The discharge-level model estimated a 4.9–percentage point reduction in any opioid prescribed at discharge in each intervention arm (peer comparison: 95% CI, −8.8 to −1.1; \(P = .02\); guidelines: 95% CI, −9.0 to −0.8; \(P = .02\); eTable 6 in Supplement 2).

**Intervention Effects on Patient Outcomes**

We found no significant differences in the probability of getting an additional opioid prescription filled, having an emergency department visit, or being hospitalized in the 30 days after the discharge (eTable 7 in Supplement 2).

**Sensitivity Analyses**

We conducted 2 sensitivity analyses. First, because only 56% of surgeons received an email during the study period, accounting for 81% of discharges (see eFigure 4 in Supplement 2), we conducted a sensitivity analysis where we restricted to surgeons who received an email at least once during the intervention period. Results were more pronounced for this subgroup (eTable 8 in Supplement 2).

Second, we estimated the intervention effects restricting to only surgeons who wrote the opioid prescription, if an opioid was prescribed, or discharged the patient themselves, if an opioid was not prescribed. Effect sizes for this subgroup were similar to the main results (eTable 8 in Supplement 2).

**Figure 3. Heterogeneity in Intervention Effects by Procedure Volume and Baseline Guideline-Discordant Opioid Prescribing**

Intervention effects are shown in absolute percentage point differences. Each point estimate is the intervention effect from a separate hierarchal model including a logit link and random effects for surgeon and specialty within hospital and controlling for each surgeon’s share of discharges with guideline-discordant opioid prescriptions in the 24 months prior to the intervention. Subgroup categories are based on quartiles of each subgroup variable (described in more detail in the eAppendix in Supplement 2). Full regression results are reported in eTable 3 in Supplement 2. Effects sizes are significantly different between the top 2 and the bottom 2 categories for baseline prescribing but not for procedure volume (eTable 4 in Supplement 2). The whiskers indicate 95% CIs.
Discussion

Principal Findings
Despite federal and state efforts, opioid overprescribing remains common in our setting, with 37% of postsurgical discharge prescriptions being above guideline quantities in our cohort before the intervention. Email feedback informing surgeons their patients were prescribed opioid amounts that exceeded institutionally endorsed guidelines or their peers’ prescriptions reduced guideline-discordant prescribing by about 5 percentage points relative to usual care. Effect sizes grew stronger over time (approximately a 10–percentage point reduction in the final 4 months of the intervention). These interventions were light-touch and incurred minimal ongoing costs after implementation, yet were associated with clinically meaningful improvements in opioid prescribing. Both interventions could be scaled to any hospital system with an EHR that recorded opioid prescriptions at patient discharge.

Implications for Public Health
A 5–percentage point reduction in guideline-discordant opioid prescribing could be important for public health outcomes. However, our prespecified analyses did not find significant effects on MMEs prescribed at discharge, which suggest the public health effects could be limited. In follow-up analysis, we found larger intervention effects (10–percentage point reductions) when we modeled effects at the discharge level (eTable 6 in Supplement 2) compared with the prespecified analyses, which included effects at the surgical specialty level. We also found significant reductions in MMEs prescribed and any opioids prescribed using the discharge-level model. This is because intervention effects were larger for surgeons with more discharges and discharges from higher-volume surgeons tend to be downweighted in the prespecified analyses. Moreover, high-volume surgeons contribute disproportionately to overprescribing. Thus, discharge-level analyses may be more relevant for assessing public health effects.

With 26,193 discharges in the intervention arms over the study year, estimates from the discharge-level model of a 12 MME reduction in opioids prescribed per discharged patient corresponded to 319,554 fewer MMEs or the equivalent of about 42,000 fewer pills of 5-mg oxycodone in Sutter Health patients' communities over the study period.

Implications for Scale-Up
Importantly, we found the guideline intervention was as successful as the peer comparison intervention. Going forward, guideline-based interventions to reduce excessive postoperative opioid prescribing may be preferable to peer-based interventions for several reasons. First, peer-comparison interventions are more challenging to implement because they require regularly updating feedback with current peer behavior. Second, although research suggests clinicians find peer comparison feedback helpful, some administrators in this study expressed concerns about peer comparison feedback. Third, guideline-based feedback can be implemented in smaller hospital systems where there are too few peers for a rich comparison.

Comparison With Other Studies
To our knowledge, this is the first study to examine the effect of feedback interventions on postoperative opioid prescribing using a randomized design, and one of the first studies to compare the effectiveness of 2 types of social norms on clinician behavior—peer-comparison descriptive norms and guideline-based injunctive norms. Prior efforts to change postoperative prescribing with behavioral interventions have used pretest-posttest designs and have shown reductions in the number of tablets prescribed. The present work is consistent with trials testing clinician feedback interventions on opioid prescribing in emergency department, primary care, and urgent care sites, although 2 other trials examining clinician feedback on concurrent prescribing of opioids and benzodiazepines did not show reductions in concurrent prescribing. The finding
that both the peer-based descriptive norm and guideline-based injunctive norm treatment conditions were equally effective contributes to an ongoing debate from the broader norms literature about which type of norm—descriptive or injunctive—is more likely to exert an effect in a fast-paced environment in which there is a great deal of cognitive depletion.45,46

Limitations
This study has several limitations. First, we could not disentangle the effects of the social norms mechanisms from the effects of simply having one’s prescribing monitored. Of course, monitoring effects would also be present in any clinical version of the intervention (part of the intervention was informing surgeons that their opioid prescribing was being monitored), so this did not bias our estimates of overall intervention effects. Second, we did not assess patient outcomes such as self-reported pain levels, daily functioning, or quality of life. However, other work has not shown a relationship between reduced postoperative prescriptions and adverse patient outcomes,39,47-51 and patients in the intervention arms were not more likely to have an emergency department visit, an inpatient visit, or additional opioid prescription fill following surgery. Third, although we used a clustered design to avoid spillover, surgeons could talk across specialties within hospitals, which could have resulted in contamination and thus understated intervention effects. However, the lack of reduction in the control group suggests that the design prevented substantial spillover effects. Finally, the email was sent to the operating surgeon even when they did not make the discharge prescription. In some cases, surgeons might not have control over prescribing at discharge, and negative feedback for these surgeons could decrease receptivity to other institutional goals. This design choice was intended to avoid spillover (nonsurgeon prescribers are more likely to prescribe across specialties) and any scaled-up version of this intervention would likely send emails directly to the clinicians who make the discharge prescription.

Conclusions
This 3-arm cluster-randomized clinical trial found that, among surgeons specializing in orthopedic, general, and obstetric/gynecological surgery, both email feedback informing surgeons that they ordered opioid prescriptions above guideline-recommended amounts and email feedback comparing their prescribing to that of their peers reduced guideline-discordant opioid prescribing compared with receiving no feedback.
**Author Contributions:** Dr Wagner had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

**Concept and design:** Wagner, Kirkegaard, Mariano, Doctor, Yan, Goldstein, Fox, Brummett, Bouskill, Martinez, Zanocco, Meeker, Waljee, Watkins.

**Acquisition, analysis, or interpretation of data:** Wagner, Kirkegaard, Mariano, Doctor, Yan, Persell, Brummett, Romanelli, Martinez, Zanocco, Meeker, Mudiganti, Waljee, Watkins.

**Drafting of the manuscript:** Wagner, Kirkegaard, Mariano, Yan, Brummett, Romanelli, Watkins.

**Critical review of the manuscript for important intellectual content:** Wagner, Kirkegaard, Mariano, Doctor, Yan, Persell, Goldstein, Fox, Brummett, Bouskill, Martinez, Zanocco, Meeker, Mudiganti, Waljee, Watkins.

**Statistical analysis:** Wagner, Mariano, Yan.

**Obtained funding:** Fox, Bouskill, Meeker, Watkins.

**Administrative, technical, or material support:** Kirkegaard, Mariano, Goldstein, Brummett, Romanelli, Bouskill.

**Supervision:** Wagner, Doctor, Yan, Watkins.

**Conflict of Interest Disclosures:**
- Dr Kirkegaard reported grants from the National Institutes of Health/National Institute on Drug Abuse (NIH/NIDA) IROIDA046226 (PI: Katherine Watkins) during the conduct of the study.
- Dr Mariano reported grants from NIH/NIDA Grant # IROIDA046226 during the conduct of the study.
- Dr Doctor reported grants from NIH during the conduct of the study.
- Dr Persell reported personal fees from RAND during the conduct of the study; personal fees from Omron Healthcare Co., Ltd.; and personal fees from the National Committee for Quality Assurance outside the submitted work.
- Dr Brummett reported personal fees from Heron Therapeutics, Merck, Vertex Pharmaceuticals, and personal fees from expert testimony outside the submitted work.
- Dr Martinez reported grants from Sutter Health during the conduct of the study.
- Dr Zanocco reported grants from NIH, Health and Human Services/NIH/NIDA (Katherine E. Watkins IROIDA046226) during the conduct of the study.
- Dr Meeker reported grants from NIH during the conduct of the study.
- Dr Mudiganti reported grants from NIDA during the conduct of the study.
- Dr Watkins reported grants from NIDA (Grant # IROIDA046226) during the conduct of the study. No other disclosures were reported.

**Funding/Support:** This study was funded by NIDA (Grant Number: IROIDA046226; PI: Katherine E Watkins).

**Role of the Funder/Sponsor:** NIDA had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

**Data Sharing Statement:** See Supplement 3.

**Additional Contributions:** We thank the study Data and Safety Monitoring Board members: Ali Arab, PhD, Georgetown University; Liisa Ecola (Executive Secretary), MA, RAND Corporation; Stephen Rosenfeld (Chair), MD, Indian University; N. Parisa Sadoughi, MD, University of California, Los Angeles; and James Wu, MD, University of California, Los Angeles and the helpful contributions of Tiffany Hruby. None were compensated.

**REFERENCES**


SUPPLEMENT 1.
Trial Protocol

SUPPLEMENT 2.
eAppendix
eTable 1. Prescribing guidelines for procedures included in the present study
eTable 2. Intervention impacts over time (full output from Figure 2B in main text)
eTable 3. Intervention effects with full set of covariates included
eTable 4. Subgroup analyses of intervention impacts on guideline-discordant prescribing
eTable 5. Tests for differences in intervention impacts by subgroup
eTable 6. Effect of social norms-based nudges on primary and secondary outcomes (discharge level)
eTable 7. Intervention effects on patient outcomes
eTable 8. Sensitivity analyses using surgeons who were more exposed to the intervention
eFigure 1. Example of Peer Comparison Intervention Email
eFigure 2. Example of Guidelines Intervention Email
eFigure 3. Effect of social norms-based nudges on each quantile of MMEs prescribed at discharge
eFigure 4. Exposure to interventions

SUPPLEMENT 3.
Data Sharing Statement