Trajectory analysis for postoperative pain using electronic health records: A nonparametric method with robust linear regression and K-medians cluster analysis

Yingjie Weng
Lu Tian
Dario Tedesco
Karishma Desai
Stanford University, USA

Steven M Asch
VA Palo Alto Health Care System, USA

Ian Carroll
Stanford University, USA

Catherine Curtin
Veterans Health Administration, USA

Kathryn M McDonald
Tina Hernandez-Boussard
Stanford University, USA

Corresponding author:
Tina Hernandez-Boussard, Department of Medicine, School of Medicine, Stanford University, Stanford, CA, 94305-6104, USA.
Email: boussard@stanford.edu
Abstract
Postoperative pain scores are widely monitored and collected in the electronic health record, yet current methods fail to fully leverage the data with fast implementation. A robust linear regression was fitted to describe the association between the log-scaled pain score and time from discharge after total knee replacement. The estimated trajectories were used for a subsequent K-medians cluster analysis to categorize the longitudinal pain score patterns into distinct clusters. For each cluster, a mixture regression model estimated the association between pain score and time to discharge adjusting for confounding. The fitted regression model generated the pain trajectory pattern for given cluster. Finally, regression analyses examined the association between pain trajectories and patient outcomes. A total of 3442 surgeries were identified with a median of 22 pain scores at an academic hospital during 2009–2016. Four pain trajectory patterns were identified and one was associated with higher rates of outcomes. In conclusion, we described a novel approach with fast implementation to model patients’ pain experience using electronic health records. In the era of big data science, clinical research should be learning from all available data regarding a patient’s episode of care instead of focusing on the “average” patient outcomes.

Keywords
electronic health records, K-medians cluster analysis, pain scores, robust linear regression

Background
Every year over 53 million Americans have surgery and pain is an expected treatment-related side effect.1,2 Appropriate postoperative pain management is critical, as poor management can lead to adverse events (e.g. deep vein thrombosis, pneumonia), compromise care of the underlying disease, and promote the transition into chronic pain.3–5 However, appropriate postoperative pain management remains a major challenge.6–8 Although patient-reported pain scores are routinely collected and widely monitored in electronic health records (EHRs),9 the appropriate utilization of these scores is not clear from a policy, clinician, or research point of view.10 Pain scores are typically reported on a 0–10 scale, where 0 indicates “no pain” and 10 indicates “worst pain.” These scores are generally used reduced to in a single moment, for example, mean or last pain score on discharge day.

Postoperative pain scores are often used as critical indicators for quality of care, providing information on patients’ recovery, guiding pain medications, including opioids, and assisting with clinical judgment regarding their postoperative care. However, there is a big gap between condensing these abundant amounts of data with plausible statistical assumptions and delivering evidence based on these statistical results to care providers. Most studies examining postoperative pain use a single time point or simple summary measures of pain scores (e.g. mean or maximum).11–13 Nevertheless, within the EHR, pain scores are captured at multiple time points and vary greatly throughout the inpatient stay.13,14 The reduction of pain scores to a single value leads to loss of information that could be critical to pain management and hence patient recovery.

Currently, there is no consensus on the best approaches for reducing the pain score data into a single value. One of the most commonly used methods is to select one summary score on the day of discharge, the mean, maximum, or last pain score before discharge, which is often then categorized into distinct groups (e.g. ‘no pain, pain score = 0,’ “mild pain, pain score 1–3,” “moderate pain, pain score 4–6”, and “severe pain, pain score 7–10”).15 These categories are then used to represent patients’ entire postoperative pain experience during inpatient stay, which can range from days to weeks depending on the patient’s diagnosis. One criticism of this naïve method is that the selection of a single summary pain score is subjective and sometimes controversial.16 Simply averaging pain scores across the entire admission might overemphasize irrelevant portions of the clinical course.
Statistical methods, such as latent class growth analysis (LCGA) implemented in PROC TRAJ (SAS),\textsuperscript{17} longitudinal LCGA, and growth mixture models (GMMs) implemented with Mplus\textsuperscript{18} and R,\textsuperscript{19} make it possible to cluster patients’ longitudinal pain path using a unified statistical model. However, most of these methods are sensitive to outliers and model assumptions, and are therefore not suitable for analyzing big data extracted from EHRs. Specifically, mixture models such as LCGA require the outcome of interest to follow a normal distribution.\textsuperscript{20} LCGA and GMM models put restrictive parametric assumptions on the structure of clusters, for example, the regression coefficients for individual trajectory need to follow normal distribution, whose mean and variance–covariance matrix depend on the corresponding cluster. Violation of these assumptions may lead to false discoveries and non-reproducible results. Another pitfall for the latent class growth model lies in the fact that the module of some analytical programs, such as PROC TRAJ in SAS, only allows pain scores to be measured at the same schedule for all patients (i.e. 2 days post operation). This is not a realistic case for application with EHR data, whereas pain scores are recorded randomly at any time point throughout the inpatient stay. Furthermore, models such as GMM, implemented in Mplus and R, were computationally intensive, because the model would estimate both the coefficients of the trajectories and the cluster parameters simultaneously, which would lead to exhausted computational memory. To make things worse, computational time could rise exponentially with each increase in the number of clusters. To ensure fast implementation, recent publications on pain trajectory analysis have focused on some semiparametric methods. For example, Kannampallil et al.\textsuperscript{21} proposed a method to identify pain trajectories by introducing K-means cluster analysis upon the empirical Bayes (EB) estimates generated from a single mixed-effect model of the entire data. This method is easy to implement, however, the employment of nonparametric K-means cluster analysis is not compatible with its key assumption that estimates generated from the mixed-effect model are from the distribution of one single target population, which also contradicted our original motivation for clustering patients into distinctive underlying subgroups. Therefore, a new set of methods should be proposed to fully leverage the rich EHR data with fast implementation and appropriate model assumptions that current methods failed to consider. Unlike most existing methods that focus on limited data with stringent assumptions not applicable to EHR data, we propose a new set of methods in consideration of the strength (large sample size) and weakness (big noise) of the EHR data and attempted to address two major challenges: (1) scalability for computation and (2) robustness to model misspecifications. The resulting proposal is a multistep process that requires minimum model assumptions, is robust to outliers, and fast to implement. Therefore, the methods we proposed are the most appropriate to leverage the vast amount of EHR data we currently face and generate evidence that can guide clinical care.

Methods

The method we propose here can be separated into three major steps (Figure 1). We propose to first use robust linear regression to get the individual trajectory. Second, K-medians cluster analysis are applied on these trajectories to identify clusters. The final step is to run generalized mixed models on each cluster to plot the corresponding trajectory patterns.

Construct individual trajectory for each inpatient stay

As opposed to longitudinal cohorts, which have a limited number of baseline and follow-up measures, in EHRs, each patient can have multiple pain scores recorded each day throughout their inpatient stay. This provides enough data to separately fit a regression model for each patients’ inpatient stay. Here, we propose to perform robust linear regression (M-estimator from the “rlm
function” in MASS package of R) to model the log transformed pain score as a function of the time. This accommodates the non-normal distribution of pain score measures and potential outliers. Coefficient estimates from each regression model are obtained via iterative weighted least square method and used for further cluster analysis.

Specifically, for each individual inpatient stay \( i = 1, 2, 3, \ldots, N \), we fit a robust linear regression as below

\[
\log(Y_{ij}) = \eta_i(t_{ij}) + \epsilon_{ij}
\]

where

\[
\eta_i(t_{ij}) = X_{ij}^T \beta_i
\]

\[X_{ij} = (1, f_1(t_{ij}), \ldots, f_m(t_{ij}))'\]

where \( Y_{ij} \) is the \( j \)th score for inpatient stay \( i \), \( t_{ij} \) is the \( j \)th time point for inpatient stay \( i \) when the \( j \)th time score was measured, \( f_k(t), k = 1, \ldots, m \) are the basis functions to expand the time variable and \( f_1(t_{ij}), f_2(t_{ij}), \ldots, f_m(t_{ij}) \) form an \( m \)-dimensional covariate in regression model.

The “basis function” can be any function that may characterize the trajectory pattern, for example, polynomial function, B-spline, S-spline, and so on. Taking the three-degree polynomial function as an example, the regression model can be written as
\[
\log(Y_{ij}) = \beta_{10} + \beta_{11}t_{ij} + \beta_{12}t_{ij}^2 + \beta_{13}t_{ij}^3 + \epsilon_{ij}
\]

**Nonparametric cluster analysis to identify the trajectory subgroups**

The aim is to cluster patient stays according to their estimated trajectory \( \hat{\gamma}_i(t) \). Specifically, patient stays \( i \) and \( j \) should be clustered together, if

\[
\int_0^\tau |\hat{\gamma}_i(t) - \hat{\gamma}_j(t)| \, dt \approx \text{small}, \quad \text{where } [0, \tau] \text{ is the time interval of interest.}
\]

In practice, we employ the following approximation in the clustering algorithm

\[
\int_0^\tau |\hat{\gamma}_i(t) - \hat{\gamma}_j(t)| \, dt \approx S^{-1} \sum_{s=1}^S |\hat{\gamma}_i(t_s) - \hat{\gamma}_j(t_s)| = S^{-1} \sum_{s=1}^S \sum_{k=0}^m \sum_{k=0}^m \beta_{ik}f_k(t_s) - \sum_{k=0}^m \beta_{jk}f_k(t_s)
\]

where \((t_1, \ldots, t_S)\) are \( S \) equally spaced points within the interval \([0, \tau]\), \( f_0(t) = 1 \), and \( \hat{\beta}_i = (\hat{\beta}_{10}, \ldots, \hat{\beta}_{1m})' \). Specifically, we have employed the following clustering algorithm:

1. For patient-stay \( i = 1, \ldots, N \), obtain \( \hat{\gamma}_i = (\hat{\gamma}_{i1}, \ldots, \hat{\gamma}_{is})' \), where

\[
\hat{\gamma}_{is} = \sum_{k=0}^m \beta_{ik}f_k(t_s)
\]

2. K-medians cluster analysis\(^{18}\) is applied to \( \{\hat{\gamma}_i, i = 1, \ldots, N\} \) to categorize inpatient stays into distinguishable groups. Specifically, the clusters are constructed by minimizing the \( L_1 \) loss function measuring the within-cluster variation

\[
\sum_{s=1}^S \sum_{i,j \in R_s} |\hat{\gamma}_{is} - \hat{\gamma}_{js}|
\]

where \( \{R_1, \ldots, R_L\} \) represent the index sets of \( L \) clusters. The \( L_1 \) metric instead of the commonly used squared Euclidean distance is used here for its robustness. Other unsupervised learning algorithms, such as K-means, K-medoids, can be implemented with different choices of the distance measure as well.

We will conduct the K-medians clustering analysis with increasing number of clusters \( L = 2, 3, \ldots \). The process will terminate if either one of the following criteria is violated:

1. The increase of between cluster variation (BCV) is above 5 percent;
2. The smallest cluster is over 5 percent of the overall population,

where the BCV is measured as

\[
\sum_{1 \leq k \leq L} \sum_{i \in R_k} \sum_{j \in R_k} |\hat{\gamma}_{is} - \hat{\gamma}_{js}|
\]

These criteria can be adjusted according to the sample size and cluster performance. The final clustering result is given based on the largest number of clusters prior to termination. The cluster
performance is also graphically examined by plotting the first two principal components (PCs) from the principal components analysis (PCA) for the trajectory parameters, $\gamma_k$. A “good” clustering result will typically demonstrate separable groups of observations projected onto the two-dimensional (2D) space spanned by the first two PCs. Other diagnostic methods can be used as well, such as plotting the distribution of distance from final centroid by cluster and bootstrapping Rand Index.\(^{22}\)

**Estimate the trajectory patterns for each subgroup**

For each cluster we identify in section “Nonparametric cluster analysis to identify the trajectory subgroups,” we further fit a generalized mixed-effects model using the log-scaled pain score as the outcome measure. We may incorporate patient demographics, clinical variables, and treatment variables as independent variables in the generalized mixed-effects model to estimate the adjusted trajectory. Specifically, for all inpatient stays in an identified cluster, we fit the following mixed-effects model

$$\log(Y_{ij}) = \beta_{i0} + \sum_{k=1}^{m} f_k(t_{ij})\beta_{ik} + \alpha'Z_i + \epsilon_{ij}$$

where $Z_i$ is the confounding factor to be adjusted

$$(\beta_0, \ldots, \beta_m)' \sim N((\beta_0, \ldots, \beta_m)', \Sigma_{\beta})$$

and $\epsilon_{ij} \sim N(0, \sigma^2_0)$. To display the cluster-specified pain score pattern, we predict the pain score using its estimated median based on the generalized mixed-effects model. In the prediction, all the confounding factors are set at their sample medians level of the entire population (thus are equal across different clusters). In sum, the predicted pain score at time $t$ is

$$\exp\left(\hat{\beta_0} + \sum_{x=1}^{m} f_x(t)\hat{\beta_x} + \hat{\alpha}'\bar{Z}\right)$$

where $\hat{\beta_0}, \ldots, \hat{\beta_m}, \hat{\alpha}$ are estimated regression coefficients and $\bar{Z}$ is the sample median of the adjusted confounding factors. We may plot time versus the predicted pain score along with the corresponding 95 percent confidence interval for each cluster.

The method implementation is coded using R software (version 3.2.4).

**Results**

We used data captured in our institution’s EHR database, CLARITY, which is a component of the Epic Systems software. We identified patients undergoing total knee arthroplasty (TKA), which is a common and often painful surgery, using ICD-9-CM, ICD-10-CM and CPT codes, 2009–2016. We captured patient demographics, inpatient/outpatient medications (down to the ingredient level), pain scores spanning the episode of care, type of insurance coverage, and follow-up visits/diagnoses/procedures up to 90 days after discharge. Patients were excluded if age at surgery was less than 18 years or death occurred during the hospitalization. This study was approved by our Institutional Review Board.
A total of 4453 encounters were identified. We excluded encounters that had a length of stay (LOS) less than 1 day or less than 10 pain scores recorded during their inpatient stay. Patients with less than 10 pain scores were excluded from analysis to prevent variability in regression coefficients estimates for individual trajectories. A total of 3442 encounters from 3025 patients were included in our final analytical dataset. There were 81,106 pain scores for the first surgery during the last 3 days of their inpatient stay. The median number of pain scores per inpatient stay was 22 (IQR: 17–28), which is sufficient to fit a cubic polynomial regression model for each inpatient stay. We focused on the last three postoperative days before discharge since the median of inpatient stay for TKA was 3.2 (IQR: 3.0 – 3.4). A three-degree polynomial function of time, which represented the linear, quadratic, and cubic terms of time from discharge, was used for building the regression model (section “Construct individual trajectory for each inpatient stay”). The pain scores were incremented by one to take into account the information corresponding to 0 pain score (indicating no pain) in the logarithmic response variable. Estimated coefficients (including the intercept) per inpatient stay were used to calculate the trajectory values at seven equally spaced time points \((t_0, t_1, \ldots, t_6)\) within the time interval [0, 3 days] referring to the last 3 postoperative days before the patient discharge described in section “Nonparametric cluster analysis to identify the trajectory subgroups”. The rationale behind choosing the last 3 postoperative days was that the median LOS was 3.2 days and this approach induces uniformity in the analysis. From 3442 estimated trajectories, four distinguished clusters were further established by K-medians algorithm with the percentage of BCV at 44 percent and the minimum size of cluster of 294 inpatient stays (8.5% of all trajectories). If the number of clusters is 5, the percentage of BCV would marginally increase to 51 percent, while the size of the smallest group fell to 127 (<5%) encounters. Therefore, four clusters were ultimately selected according to the criteria we described in section “Nonparametric cluster analysis to identify the trajectory subgroups”. To visualize the clustering results, trajectories were represented by their first two PCs, which are plotted and colored differently by clusters in Figure 2.

Next, we estimated the trajectory pattern by fitting the mixed-effects model for each cluster, adjusting for several patient and clinical-related covariates: patients’ age at admission, race-ethnicity,
gender, marital status, number of comorbidities at admission, body mass index (BMI), LOS, length of the procedure in hours, presurgery pain score, and pre- and postsurgery morphine equivalent value per day calculated using oral morphine conversion factors. Marital status was included because it is related to social isolation and associated with overall well-being. The predicted pain score as a function of time from discharge were plotted for each cluster as shown in Figure 3.

Patients’ characteristics and their clinical information are summarized by cluster in Table 1. Four unique patterns of postoperative pain experience were discovered in our patient cohort. Cluster 1 encounters had mild pain after surgery followed by a steady rise in pain scores before discharge ("Slightly Rise" Group). Their final pain levels at discharge were between three and four. This group of patients was more likely to be female (66%), living without a partner (21%), stayed

Figure 3. Patients’ patterns of pain score versus days from discharge.
in the hospital longer (3.5 days), had higher opioid usage after surgery (67.1 mg/day), and higher preoperative pain scores (2.3). Cluster 2 encounters represented a pain trajectory of patients undergoing TKA with moderate pain scores after surgery and fluctuated pain level during their stay but reported very low pain at discharge (“Completely Drop” Group). The patients of this group were older (69.4 years), took less opioids (55.0 mg/day), and received less complicated procedures (length of procedure: 1.7 h). Cluster 3 was a small group of unique patients that initially experienced very low pain immediately following surgery, but their pain rose sharply before discharge (“Sudden Rise” Group). These patients were younger (66.5 years old), more likely to be male (49%), more likely to be Hispanic and Black (20%), had lower preoperative pain scores (1.9), and

| Variable                              | Cluster 1 | Cluster 2 | Cluster 3 | Cluster 4 | p value |
|---------------------------------------|-----------|-----------|-----------|-----------|---------|
| n                                     | 1020      | 694       | 294       | 1434      |         |
| Age at admission, years, mean (SD)    | 67.5 (10.69) | 69.4 (10.12) | 66.5 (10.78) | 67.2 (10.93) | <0.001  |
| Gender, n (%)                         |           |           |           |           | <0.001  |
| Female                                | 670 (66)  | 426 (61)  | 150 (51)  | 894 (62)  |         |
| Male                                  | 350 (34)  | 268 (39)  | 144 (49)  | 540 (38)  |         |
| Race-ethnicity, n (%)                 |           |           |           |           | 0.015   |
| White                                 | 642 (63)  | 480 (69)  | 186 (63)  | 929 (65)  |         |
| Black                                 | 41 (4)    | 23 (3)    | 19 (6)    | 58 (4)    |         |
| Hispanic                              | 122 (12)  | 58 (8)    | 42 (14)   | 146 (10)  |         |
| Asian                                 | 100 (10)  | 64 (9)    | 29 (10)   | 172 (12)  |         |
| Other                                 | 97 (10)   | 61 (9)    | 15 (5)    | 107 (7)   |         |
| BMI, mean (SD)<sup>b</sup>            | 31.0 (6.96) | 30.4 (6.59) | 30.5 (6.32) | 30.5 (6.80) | 0.39    |
| Marital status, n (%)                 |           |           |           |           | 0.038   |
| Married/life partner                  | 651 (64)  | 476 (69)  | 202 (68)  | 916 (64)  |         |
| Single                                | 152 (15)  | 87 (13)   | 38 (13)   | 246 (17)  |         |
| Widowed/divorced/separated            | 213 (21)  | 126 (18)  | 50 (17)   | 265 (18)  |         |
| Comorbidities, n (%)                  |           |           |           |           | 0.23    |
| >2                                    | 399 (39)  | 240 (35)  | 103 (36)  | 544 (38)  |         |
| 2                                     | 261 (26)  | 170 (25)  | 75 (26)   | 386 (27)  |         |
| 1                                     | 232 (23)  | 183 (26)  | 69 (24)   | 301 (21)  |         |
| None                                  | 127 (12)  | 100 (14)  | 43 (15)   | 201 (14)  |         |
| Preoperative pain score, mean (SD)    | 2.3 (2.86) | 2.0 (2.65) | 1.9 (2.52) | 2.4 (2.91) | 0.016   |
| Length of procedure, h, mean (SD)     | 1.8 (0.79) | 1.7 (0.55) | 1.6 (0.54) | 1.9 (0.83) | <0.001  |
| Length of stay, days, mean (SD)       | 3.5 (1.29) | 3.3 (1.35) | 3.0 (1.32) | 3.2 (1.67) | <0.001  |
| Preoperative anxiety level, n (%)     |           |           |           |           | 0.71    |
| Severe/moderate                       | 53 (6)    | 32 (5)    | 9 (3)     | 70 (5)    |         |
| Mild                                  | 460 (54)  | 302 (52)  | 129 (54)  | 613 (51)  |         |
| None                                  | 346 (40)  | 244 (42)  | 102 (42)  | 519 (43)  |         |
| Preoperative MME<sup>a</sup> per day, mg/day, mean (SD) | 7.3 (35.65) | 7.6 (35.50) | 4.3 (24.18) | 9.0 (40.30) | 0.10    |
| Postoperative MME<sup>a</sup> per day, mg/day, mean (SD) | 87.4 (87.78) | 73.0 (83.82) | 79.0 (115.6) | 99.7 (113.4) | <0.001 |

BMI: body mass index; SD: standard deviation.

<sup>a</sup>Morphine equivalent value.

<sup>b</sup>One patient in cluster 2 with missing/invalid value for BMI was excluded.
had less complicated procedures (length of procedure: 1.6 h) compared to patients in the other clusters. Cluster 4 consisted of patients who reported moderate pain (pain score = 2–3) throughout their inpatient stay (“Steady” Group). The “steady” group tended to be younger (67.2 years), had higher preoperative pain scores (pain score = 2.4), experienced longer procedure time (1.9 h), and received more postoperative opioid drugs (69.3 mg/day) during the inpatient stay compared to other clusters.

In our cohort, patients’ inpatient pain experience was marginally associated with patients’ demographics, injury severity, treatment options, and opioid medications. We hypothesized that these trajectories could be used as surrogates of patients’ recovery or early indicators of post-discharge complications in addition to other important clinical factors. Among all the four groups, we also hypothesized that “Sudden Rise” would be one distinct group of patients who might be more susceptible to complications after discharge.

To test these hypotheses, we conducted a set of logistic regressions with the occurrence of 30, 60, and 90-day follow-up visits of all purposes, inpatient readmissions or subsequent emergency department visits, and post-discharge complications (surgery-pain-related revisits, wound infection and others, see Supplementary Table 1) as the binary outcomes, respectively. The pain trajectory pattern, patients’ demographics, and clinical covariates were included as covariates. The “Steady” cluster was set to be the reference group in our analysis because it was a group with the largest sample size and was considered a clinically typical “well-managed” group. Compared to the “Steady” group (Cluster 4), patients in the “Sudden Rise” group (Cluster 3) were associated with higher risk of follow-up revisits (odds ratio (OR): 2.37, 2.11, and 1.97 for 30, 60, and 90-day windows, respectively), any surgery-related pain (OR: 5.49, 3.41, 2.73), and surgery-related chronic pain (OR: 5.82, 2.96, 2.03). In addition, we noticed that the “Complete Drop” group had higher risk of any surgery-related pain (OR: 2.36), follow-up visits of any types (OR: 1.36), and inpatient readmissions/subsequent ED (emergency department) visits (OR: 1.93) 30 days after discharge, although we failed to observe statistically significant effects in 60 and 90 days for these outcomes. No statistical significance was detected for complications of any types across all the observation windows (Table 2). Since readmissions and complications were rare in our population, Poisson regressions and negative-binomial regressions were also performed using the number of post-discharge revisits with any specific outcome as the dependent variable. Consistent with the results from logistic regression, the “Sudden Rise” group had higher rates of follow-up visits at 30-day post-discharge as well as higher rates of any surgical-related pain, including chronic pain in 30, 60, or 90-day window (Supplementary Table 2).

The trajectory analyses were compared with other basic analytical approaches that are common in the literature, that is last recorded pain score, mean pain score on discharge day, and max pain score on discharge day (Supplementary Figure 1).27–30 In terms of predictions and model fits, our trajectory analyses outperform all the single score discharge pain methods in regard to their area under curve (AUC) and Akaike information criterion (AIC). (Supplementary Table 3).

**Discussion**

We have entered a new era in which the healthcare system has undergone dramatic changes. In 2017, over 90 percent of US hospitals had a functioning EHR system.31 With the improvement of health informatics technology, massive amounts of patient and clinical information are now captured and stored in EHRs. However, how to meaningfully extract and analyze these data becomes a new challenge in both the clinical and statistics world. Pain scores derived from EHRs, as an example here, are often not used efficiently and effectively in clinical research. The practical difficulty lies in that current analytical methods, such as mixed growth curve modeling or GMM, are
not scalable to cope with the amount of data in large EHR datasets and other methods of limiting assumptions cannot be used on the complex and abundant EHR data. Therefore, novel methods that are applicable to massive EHR data are critically needed in new areas in big data clinical research.

The method we proposed here, which combined robust linear regression and unsupervised K-medians cluster analysis, was able to compile all the pain score data recorded in the EHR and identify distinguished patterns of inpatient pain scores after TKA surgery. This is achieved by using a scalable approach for dimension reduction, which utilizes regression models to describe the relationship between score and time for each time series/patient stay. Predicted values at a set of pre-specific time points that converted varied numbers of scores into a fixed number of trajectories.

Table 2. Logistic regression for major post-discharge outcomes by cluster.a

| Outcome                  | Cluster 1          | Cluster 2          | Cluster 3          |
|--------------------------|--------------------|--------------------|--------------------|
| **n =**                  | 1020               | 694                | 294                |
| **30-day (n=)**b         |                    |                    |                    |
| All revisits             | 1.15 (0.97–1.38)   | 1.36 (1.11–1.66)** | 2.37 (1.77–3.17)** |
| Inpatient                | 1.10 (0.53–2.23)   | 1.84 (0.87–3.78)   | 1.26 (0.35–3.55)   |
| Emergency department     | 0.97 (0.52–1.78)   | 1.67 (0.89–3.09)   | 0.36 (0.06–3.54)   |
| Inpatient + ED           | 1.10 (0.68–1.77)   | 1.93 (1.19–3.12)** | 0.78 (0.29–1.74)   |
| Complications (any)      | 1.00 (0.67–1.50)   | 1.14 (0.72–1.79)   | 0.68 (0.29–1.40)   |
| Surgery-related pain     | 1.77 (1.02–3.10)*  | 2.36 (1.31–4.27)** | 5.49 (2.99–10.09)**|
| Surgery-related acute pain | 2.86 (1.06–8.54)*  | 2.60 (0.79–8.70)   | 2.44 (0.51–10.41)  |
| Surgery-related chronic pain | 2.13 (0.94–5.33) | 2.00 (0.78–5.23) | 5.82 (2.44–15.60)**|
| **60-day (n=)**b         |                    |                    |                    |
| All revisits             | 1.11 (0.92–1.35)   | 1.10 (0.89–1.37)   | 2.11 (1.50–3.01)** |
| Inpatient                | 0.91 (0.51–1.59)   | 1.31 (0.71–2.34)   | 1.08 (0.40–2.50)   |
| Emergency department     | 0.95 (0.56–1.57)   | 1.41 (0.81–2.39)   | 0.39 (0.09–1.12)   |
| Inpatient + ED           | 0.95 (0.64–1.41)   | 1.48 (0.97–2.22)   | 0.76 (0.34–1.50)   |
| Complications (any)      | 0.91 (0.65–1.25)   | 1.14 (0.79–1.62)   | 0.69 (0.36–1.21)   |
| Surgery-related pain     | 1.21 (0.80–1.84)   | 1.35 (0.84–2.14)   | 3.41 (2.09–5.52)** |
| Surgery-related acute pain | 1.96 (0.86–4.57)   | 1.64 (0.58–4.36)   | 1.95 (0.52–6.10)   |
| Surgery-related chronic pain | 1.11 (0.59–2.07) | 0.90 (0.42–1.84) | 2.96 (1.47–5.90)**|
| **90-day (n=)**b         |                    |                    |                    |
| All revisits             | 1.08 (0.89–1.31)   | 1.08 (0.87–1.34)   | 1.97 (1.39–2.85)** |
| Inpatient                | 0.96 (0.59–1.54)   | 1.01 (0.57–1.73)   | 1.07 (0.45–2.24)   |
| Emergency Department     | 0.93 (0.57–1.50)   | 1.28 (0.75–2.13)   | 0.36 (0.09–1.01)   |
| Inpatient + ED           | 0.95 (0.66–1.36)   | 1.25 (0.84–1.84)   | 0.78 (0.38–1.45)   |
| Complications (any)      | 0.92 (0.67–1.24)   | 1.21 (0.86–1.68)   | 0.74 (0.41–1.26)   |
| Surgery-related pain     | 1.10 (0.76–1.59)   | 1.25 (0.82–1.88)   | 2.73 (1.71–4.27)** |
| Surgery-related acute pain | 1.79 (0.83–3.92)   | 1.58 (0.61–3.93)   | 2.11 (0.64–5.99)   |
| Surgery-related chronic pain | 0.82 (0.46–1.42) | 0.86 (0.46–1.57) | 2.03 (1.06–3.80)*  |

aLogistic regression was fitted for each outcome, adjusting for age at admission, gender, race-ethnicity, marital status, length of stay, postoperative morphine-equivalent-values per day, number of comorbidities, preoperative pain score. Odds ratios (ORs) and their corresponding 95 percent confidence intervals (CI) were reported in the table. Cluster 4 was the control group for all analysis (n = 1407, 1392, 1373 at 30, 60, and 90 days).

bInpatient stays with patients whose admission date fell out of the corresponding observation window were not included in the analysis, despite that they were included in the original cluster analysis.

*p < 0.05; **p < 0.01; ***p < 0.001.
per patient stay were therefore constructed from the regression models for further unsupervised learning/cluster analysis. Our method is flexible with any time metrics (e.g. time before discharge and time after surgery) and any hypothesized shape of the pain scores (e.g. polynomial and S-pline). It is not limited to model pain scores and other numeric or ordinal values that were commonly recorded in the EHR data (e.g. lab values) can be modeled similarly with appropriate modifications. The method is scalable to large amounts of data and does not heavily rely on the restrictive underlying distributional assumptions as the other statistical methods, for example, the GMM. In addition, K-medians clustering was proposed for clustering, which minimized the effects of outliers upon the clustering results. Although we proposed to implement the K-medians clustering and robust linear regression to address the negative effect of the outliers, other techniques, such as weighted dynamic time warping or longest common subsequence distance measure, could be incorporated in our proposed 3-step method with minor modifications as well.

The method developed here was robust and superior in terms of prediction compared to other commonly used pain score analyses, that is mean, last, and maximum pain scores on discharge day.32,33 The single-value analyses were not able to distinguish between patient in the steady pain trajectory and those in the cluster with a sudden rise in pain scores at discharge. This is an important distinction, as patients in the sudden rise trajectory were more likely to have adverse pain-related events following discharge. When analyses focus on a single pain score, it is important understand the ubiquity of pain score recordings in the EHRs. For example, the last pain score recorded can occur minutes or hours prior to discharge, making this number extremely susceptible to inpatient pain medication. Our method that leverages all pain scores captured during the inpatient stay is a clear step toward patient-centered care, enabling clinicians to treat a patient’s pain experience rather than a single pain score.

From a clinical perspective, our method was able to identify subpopulations of patients whose distinguishable inpatient pain trajectories were associated with adverse outcomes, in particular pain-related readmissions. Pain-related readmissions following surgery are not uncommon and are costly to the healthcare system.34,35 The methods developed in this study could be used to identify patients needing additional pain management resources upon discharge. Such pain trajectories could also be incorporated into clinical decision tools at the point of care, providing evidence to guide pain management—a clear need given our nation’s current opioid epidemic.36–38

There are several limitations in our method. First, the method was developed under the EHR setting in which pain scores are attempted to be recorded at varying intervals over the entire inpatient stay. Furthermore, many question the utilization of pain scores to represent a patient’s pain experience.39,40 However, to date, pain scores are the best representation of a patient’s pain experience at a population level and outside of controlled, qualitative studies. For a single robust linear regression with three-degree polynomial expansion of time, we need to have at least 10 pain scores per stay to obtain an acceptable fit. Therefore, our methods exclude stays with too few values, which may still contain valuable information. Second, we employed several criteria to select the optimal number of clusters, such as the size of the smallest cluster. These criteria can be subjective based on both clinical and statistical judgments and often are determined in an ad hoc manner. Selection of different criteria will therefore result in different results. In addition, for patients that lie close to the boundary between two clusters, the distance to two cluster centers can be similar. Although we choose to assign them into the cluster with the nearest distance numerically, interpretation of these patients is difficult. Such assignment criteria ignored the uncertainty in cluster membership for those patients and tend to dilute the association between the trajectory pattern and clinical outcome. On the other hand, a large number of patients or inpatient stays are needed to identify a trajectory pattern, which is not commonly observed but may be clinically important. For cohorts with small numbers of observations, it will be difficult
to distinguish genuine trajectory patterns from “artificial” clusters formed by a random chance. Future application of the method in a nationwide EHR system should be implemented to validate our findings. If validated, the approach could be applied to many other longitudinal clinical data to predict meaningful patient outcomes. Finally, our method could be criticized by the fact that it fails to incorporate the sampling variability of the coefficient estimates in its first step of the individual regressions. In our sample implementation, we specifically restricted the inpatient stays into the analysis with those that had at least 10 pain scores to ensure stability of estimates for the coefficients across all individual regressions. However, further modification that enables the incorporation of the sampling variability could be promising on the premises of our current algorithm.

This work provides several avenues for future research. First, the replication of this study in other healthcare systems could provide meaningful evidence on health setting pain management practices and its effect on postoperative pain outcomes. Here, we look at the associations of pain trajectories with adverse events; however, future work could use pain trajectories to predict patients at risk for adverse events prior to discharge—a clinical tool that would be useful for quality improvement and resource allocation. Finally, the methodology developed in pain can be applied to other clinical domains, such as prostate-specific antigen (PSA) trajectory analyses, and their association with mortality in prostate cancer patients.

Conclusion

In summary, we have described a novel approach with fast implementation to model patients’ pain experience using EHRs. Higher rates of surgery-related pain after discharge were observed in one empirically distinguishable inpatient pain trajectory with EHR data. This approach could be applied to many other longitudinal clinical data to predict meaningful patient outcomes. Moving toward a learning healthcare system, clinical care should be learning from all available data regarding a patient’s episode of care instead of focusing on an “average” patient or score. We now have ever-advancing analytic capacity and our method provides a rigorous and statistically sound approach to leverage longitudinal clinical data from EHRs for personalized treatment plans.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This project was supported by Grant No. R01HS024096 from the Agency for Healthcare Research and Quality. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Agency for Healthcare Research and Quality.

ORCID iD

Tina Hernandez-Boussard https://orcid.org/0000-0001-6553-3455

Supplemental material

Supplemental material for this article is available online.
References

1. DeFrances CJ and Podgornik MN. 2004 National Hospital Discharge Survey. *Adv Data* 2006(371): 1–19.
2. Kozak LJ, DeFrances CJ and Hall MJ. National hospital discharge survey: 2004 annual summary with detailed diagnosis and procedure data. *Vital Health Stat* 2006; 162: 1–209.
3. Adamson SJ, Deering DE, Sellman JD, et al. An estimation of the prevalence of opioid dependence in New Zealand. *Int J Drug Policy* 2012; 23(1): 87–89.
4. Katz JN, Losina E, Barrett J, et al. Association between hospital and surgeon procedure volume and outcomes of total hip replacement in the United States medicare population. *J Bone Joint Surg Am* 2001; 83(11): 1622–1629.
5. Goldstein DH, Ellis J, Brown R, et al. Recommendations for improved acute pain services: Canadian collaborative acute pain initiative. *Pain Res Manag* 2004; 9(3): 123–130.
6. Apfelbaum JL, Chen C, Mehta SS, et al. Postoperative pain experience: results from a national survey suggest postoperative pain continues to be undermanaged. *Anesth Analg* 2003; 97(2): 534–540.
7. Edlund MJ, Martin BC, Fan MY, et al. Risks for opioid abuse and dependence among recipients of chronic opioid therapy: results from the TROUP study. *Drug Alcohol Depend* 2010; 112(1–2): 90–98.
8. Waljee JF, Li L, Brummett CM, et al. Iatrogenic opioid dependence in the United States: are surgeons the gatekeepers. *Ann Surg* 2017; 265(4): 728–730.
9. Kerns R, Wasse L, Ryan B, et al. *Pain as the 5th vital sign toolkit*. Washington, DC: *Veterans Health Administration*, 2000.
10. Ruau D, Liu LY, Clark JD, et al. Sex differences in reported pain across 11,000 patients captured in electronic medical records. *J Pain* 2012; 13(3): 228–234.
11. Kalkman CJ, Visser K, Moen J, et al. Preoperative prediction of severe postoperative pain. *Pain* 2003; 105(3): 415–423.
12. Rice DA, Kluger MT, McNair PJ, et al. Persistent postoperative pain after total knee arthroplasty: a prospective cohort study of potential risk factors. *Br J Anaesth* 2018; 121(4): 804–812.
13. Desai K, Carroll I, Asch SM, et al. Utilization and effectiveness of multimodal discharge analgesia for postoperative pain management. *J Surg Res* 2018; 228: 160–169.
14. Yim WW, Wheeler AJ, Curtin C, et al. Secondary use of electronic medical records for clinical research: challenges and Opportunities. *Converg Sci Phys Oncol* 2018; 4(1).
15. Hughes R. *Patient safety and quality: An evidence-based handbook for nurses*, vol. 3. Rockville, MD: Agency for Healthcare Research and Quality, 2008.
16. Lund I, Lundeborg T, Sandberg L, et al. Lack of interchangeability between visual analogue and verbal rating pain scales: a cross sectional description of pain etiology groups. *BMC Med Res Methodol* 2005; 5: 31.
17. Jones BL, Nagin DS and Roeder K. A SAS procedure based on mixture models for estimating developmental trajectories. *Sociol Method Res* 2001; 29(3): 374–393.
18. Jung T and Wickrama K. An introduction to latent class growth analysis and growth mixture modeling. *Soc Personal Psychol Compass* 2008; 2(1): 302–317.
19. Proust-Lima C, Dartigues JF and Jacqmin-Gadda H. Joint modeling of repeated multivariate cognitive measures and competing risks of dementia and death: a latent process and latent class approach. *Stat Med* 2016; 35(3): 382–398.
20. Ram N and Grimm KJ. Growth mixture modeling: a method for identifying differences in longitudinal change among unobserved groups. *Int J Behav Dev* 2009; 33(6): 565–576.
21. Kannamapillil T, Galanter WL, Falck S, et al. Characterizing the pain score trajectories of hospitalized adult medical and surgical patients: a retrospective cohort study. *Pain* 2016; 157(12): 2739–2746.
22. Gentleman RC, Carey VJ, Bates DM, et al. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* 2004; 5(10): R80.
23. Wasey J. icd: tools for working with ICD-9 and ICD-10 codes, and finding comorbidities (*R Package Version 2.1; 2016*), 2017, https://www.rdocumentation.org/packages/icd/versions/2.2
24. Nielsen S, Degenhardt L, Hoban B, et al. A synthesis of oral morphine equivalents (OME) for opioid utilisation studies. *Pharmacoepidemiol Drug Saf* 2016; 25(6): 733–737.
25. Alcantara Montero A, Sanchez Carnerero CI, Ibor Vidal PJ, et al. [CDC guidelines for prescribing opioids for chronic pain]. *Semergen* 2017; 43(4): e53–e54.
26. Sorlie PD, Backlund E and Keller JB. US mortality by economic, demographic, and social characteristics: the National Longitudinal Mortality Study. *Am J Public Health* 1995; 85(7): 949–956.
27. Griffioen MA, Greenspan JD, Johantgen M, et al. Acute pain characteristics in patients with and without chronic pain following lower extremity injury. *Pain Manag Nurs* 2017; 18(1): 33–41.
28. Prasad A, Mukherjee KA, Kaul S, et al. Postoperative pain after cholecystectomy: conventional laparoscopy versus single-incision laparoscopic surgery. *J Minim Access Surg* 2011; 7(1): 24–27.
29. Trzcinski S, Rosenberg RE, Vasquez Montes D, et al. Use of gabapentin in posterior spinal fusion is associated with decreased postoperative pain and opioid use in children and adolescents. *Clin Spine Surg* 2019; 32(5): 210–214.
30. Lingren T, Sadhasivam S, Zhang X, et al. Electronic medical records as a replacement for prospective research data collection in postoperative pain and opioid response studies. *Int J Med Inform* 2018; 111: 45–50.
31. Adler-Milstein J and Jha AK. HITECH act drove large gains in hospital electronic health record adoption. *Health Aff* 2017; 36(8): 1416–1422.
32. Huang N, Cunningham F, Laurito CE, et al. Can we do better with postoperative pain management? *Am J Surg* 2001; 182(5): 440–448.
33. Curtin CM and Hernandez-Boussard T. Readmissions after treatment of distal radius fractures. *J Hand Surg Am* 2014; 39(10): 1926–1932.
34. Finnegan MA, Shaffer R, Remington A, et al. Emergency department visits following elective total hip and knee replacement surgery: identifying gaps in continuity of care. *J Bone Joint Surg Am* 2017; 99(12): 1005–1012.
35. Shaffer R, Backhus L, Finnegan MA, et al. Thirty-day unplanned postoperative inpatient and emergency department visits following thoracotomy. *J Surg Res* 2018; 230: 117–124.
36. Gawande AA. It’s time to adopt electronic prescriptions for opioids. *Ann Surg* 2017; 265(4): 693–694.
37. Murthy VH. Ending the opioid epidemic: a call to action. *N Engl J Med* 2016; 375(25): 2413–2415.
38. Tedesco D, Asch SM, Curtin C, et al. Opioid abuse and poisoning: trends in inpatient and emergency department discharges. *Health Aff* 2017; 36(10): 1748–1753.
39. Lorenz KA, Sherbourne CD, Shugarman LR, et al. How reliable is pain as the fifth vital sign. *J Am Board Fam Med* 2009; 22(3): 291–298.
40. Krebs EE, Lorenz KA, Bair MJ, et al. Development and initial validation of the PEG, a three-item scale assessing pain intensity and interference. *J Gen Intern Med* 2009; 24(6): 733–738.