

# What is a “normal” postoperative temperature? Group based trajectory modeling in postoperative knee arthroplasty patients in a large health system

Nirav Shah, MD, MPH<sup>1</sup>, Vivek Vegi<sup>2</sup>, Ankit Dhingra<sup>2</sup>, Rema Padman, PhD<sup>2</sup>, Daniel Nagin, PhD<sup>2</sup>, Ari Robicsek, MD<sup>1,3</sup>

<sup>1</sup>University of Chicago Medicine, Chicago, IL; <sup>2</sup>Carnegie Mellon University, Pittsburgh, PA; <sup>3</sup>NorthShore University HealthSystem, Chicago, IL

## Abstract

*Normal temperature values in the postoperative setting have not been characterized or defined. While fever is commonly associated with infection, a postoperative fever may be a normal physiologic response. We evaluated 5,793 patients who had elective knee arthroplasty between January 1, 2007 and December 31, 2013, and were discharged within 5 days. Inpatient temperature, gender, age and BMI measurements were extracted from the data warehouse. Group based trajectory modeling was performed to generate temperature curves of postoperative patients following similar progression of maximum temperature over four-hour time intervals. Preliminary results indicate four distinct trajectories of maximum temperature over time that enables risk stratification of patient populations for targeted interventions, with age and BMI emerging as significant covariates in all clusters. Ongoing work will determine the influence of these baseline covariates on probability of an individual following the four baseline trajectories, thus enabling prediction of an individual's group membership over time.*

## Introduction

In the middle of the 19th century, a German physician named Carl Reinhold August Wunderlich embarked on a project to characterize normal human body temperature<sup>1</sup>. This massive effort is thought to have included some 25,000 subjects and a million temperature measurements. It established 98.6°F (37.0°C) as ‘normal’ human body temperature and 100.4°F (38.0°C) as the upper limit of normal ‘fever’<sup>1,2</sup>. There has never been an equivalent large scale evaluation of ‘what is a fever’ in settings of physiological perturbation, such as the postoperative setting. Using a health record in which all temperature measurements are recorded electronically, the Wunderlich Project was initiated at NorthShore University HealthSystem to characterize temperature expectations in such non-normal settings. In this paper, we apply a semi-parametric, group-based developmental trajectory analysis method using data on a cohort of patients over a six year period who had elective knee arthroplasty and were discharged within 5 days, with the objective of delineating the distinct trajectories of temperature development in the population and profiling their membership to facilitate risk prediction of elevated temperatures and poor outcomes.

## Background

Fever is typically associated with infection, however, a postoperative fever, as defined by a body temperature  $\geq 100.4^{\circ}\text{F}$  ( $38^{\circ}\text{C}$ ) may be a normal physiologic response<sup>3,4</sup>. In fact, postoperative fever is a relatively common occurrence after surgery with an incidence between 10% - 40%<sup>5,6,7</sup>. While most of these elevated temperatures are thought to be non-infectious and related to the underlying physiologic perturbation of surgery, they often provoke healthcare providers to work patients up for infections or noninfectious complications. These workups are invasive, costly and have been shown to be low yield<sup>8</sup>. Determining normal postoperative temperature may therefore be able to assist providers in better distinguishing patients at risk of postoperative complications.

To address this question, investigators at NorthShore have previously derived a web-based tool entitled Wunderlich, which can be found at the following website: <http://fad.northshore.org/wunderlich/>. In this web tool, more than 17,000 patients who underwent 11 common surgeries (cardiovascular surgery, vascular surgery, hip and knee replacement, laparoscopic cholecystectomy, colectomy, hysterectomy, craniotomy, prostatectomy, multilevel spinal surgery and posterior spinal fusion) were identified. Next, the maximum oral temperature on each postoperative day,  $T_{\text{max}}$ , as well as pertinent surgical, demographic and clinical data, was identified for each patient. A multivariable model was constructed for each postoperative day after each procedure. These multivariable models were used to generate a  $T_{\text{max}}$  ‘prediction’ that is specific to a patient, a procedure, and a postoperative day. The website allows providers to enter basic clinical data into the web based tool and see how their patient's temperature compares to

that of similar patients undergoing the same procedure on the same postoperative day. For example, if a patient's  $T_{\max}$  percentile is 95, that patient's temperature is higher than those of 95% of clinically similar patients undergoing the same procedure on the same postoperative day.

In our current study, we would like to further optimize these initial models to improve our characterization of a normal postoperative temperature and identify subgroups of the patient population that follow similar patterns of temperature development. To the best of our knowledge, there are no similar studies that are modeling postoperative temperatures to determine 'normal temperature' in this setting. Hence, other than some data showing that a temperature in the first 48 hours most likely represents a benign phenomenon but results in many costly workups that are not effective, we believe that the approach described in this study is fairly novel.

## Methods

### Basic Trajectory Analysis

A developmental trajectory describes the course of an outcome over age or time<sup>9</sup>. In this research, we apply a method called group-based trajectory modeling (GBTM) to study the developmental course of patient temperatures in the postoperative setting. Longitudinal data—data with a time-based dimension—provide the empirical foundation for the analysis of developmental trajectories. In contrast to hierarchical modeling and latent curve analysis<sup>10</sup>, GBTM is designed to identify clusters of individuals following possibly distinctive trajectories of an outcome of interest. The two primary outputs of the model are the shape of the trajectory for each group which may differ not only in their level but also in the direction and rate of their movement (e.g., one may be rising quickly and another declining slowly) and the size of the group as measured by the proportion of population under study following the trajectory. The method has been widely applied not only in the social and behavioral sciences to study phenomenon such as technology adoption in healthcare<sup>11</sup> and crime<sup>9</sup> but also in clinical settings to study the developmental course of symptoms of conduct disorder, schizophrenia, obesity, levels of disability in the elderly<sup>12</sup> and disease progression<sup>13</sup>.

GBTM is an application of finite mixture modeling<sup>9</sup>. The fundamental concept of interest is the distribution of outcomes conditional on a time-related metric such as time from the onset of treatment; that is, the distribution of outcome trajectories is denoted by  $P(Y_i | Z_i)$ , where the random vector  $Y_i$  represents individual  $i$ 's longitudinal sequence of behavioral outcomes and the vector  $Z_i$  represents characteristics of  $i$  measured at baseline or are otherwise invariant. The group-based trajectory model assumes that the population distribution of trajectories arises from a finite mixture of unknown order  $J$ . The likelihood for each individual  $i$ , conditional on the number of groups  $J$ , may be written as  $P(Y_i | Z_i) = \sum_{j=1}^J \pi_j \cdot P(Y_i | Z_i, j; \beta_j)$ , where  $\pi_j$  is the probability of membership

in group  $j$ , and the conditional distribution of  $Y_i$  given membership in  $j$  is indexed by the unknown parameter vector  $\beta_j$  which among other things determines the shape of the group-specific trajectory. Typically, the trajectory is modeled using a polynomial function of the time-related metric. For given  $j$ , conditional independence is assumed for the sequential realizations of the elements of  $Y_i$ ,  $y_{it}$ , over the  $T$  periods of measurement. Thus, we may write

$P(Y_i | Z_i, j; \beta_j) = \prod_{t=1}^T p(y_{it} | z_{it}, j; \beta_j)$ , where  $p(\cdot)$  is the distribution of  $y_{it}$  conditional on membership in

group  $j$  and the  $z_{it}$ .

Canned software for the estimation of GBTM allows for the specification of  $p(\cdot)$  as the normal or censored normal distribution, the zero-inflated Poisson distribution, or the binary logit distribution but in general there is no restriction on the form of  $p(\cdot)$ <sup>14, 15</sup>. We can further identify the predictors of the trajectories using a rich set of the baseline characteristics that are measured at the outset of the study period. The method handles missing data, including exposure time for count data. A generalization of  $P(Y_i | Z_i)$  specifies  $\pi_j$  as a multinomial logit function of baseline characteristics included in  $Z_i$ . This provides us with the capacity to examine whether and how such baseline characteristics influence the probability of an individual following the various baseline trajectories<sup>9</sup>.

### Data

NorthShore University HealthSystem maintains an electronic data warehouse (EDW) that captures information entered into the electronic health record. This data includes inpatient and outpatient health information that includes demographics, vitals, labs, microbiology, radiology, medications, comorbidities, surgeries and costs. For the current Wunderlich project, data were extracted for all patients who underwent 15 different elective surgeries (caesarian section, cardiac surgery, colectomy, craniectomy, hip replacement, hysterectomy, knee replacement, lap

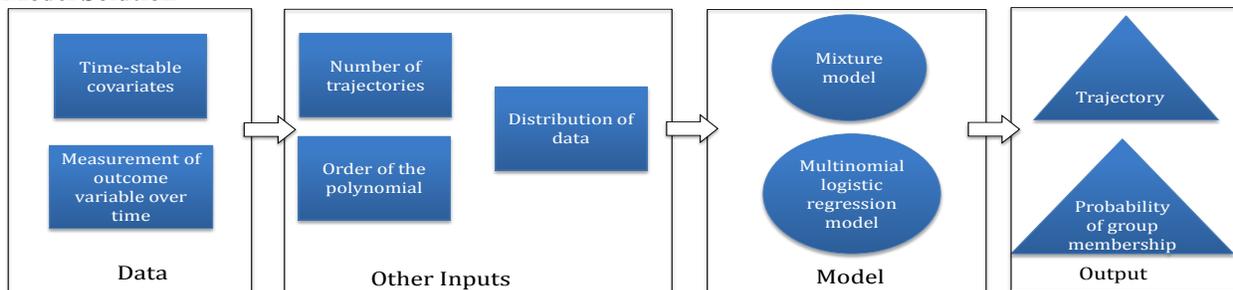
cholecystectomy, lap colectomy, mastectomy, multilevel spinal surgery, prostatectomy, thyroidectomy, transurethral prostatectomy, vascular surgery) between January 1, 2007 and December 31, 2013 to establish postoperative temperature patterns, but this study focuses only on the analysis of data on *elective knee replacement*. For each surgical episode, the following de-identified information regarding the index hospitalization and any subsequent patient encounters within 30 post-operative days were extracted from the EDW:

1. MRN of patients (Patients' names are not used)
2. Demographic data including age, sex, race, ethnicity, BMI
3. Comorbidities including ICD-9 codes, problem list, past medical history
4. All vital signs including temperature measurements and site of temperature measurement
5. All laboratory data which also includes microbiologic data
6. All blood transfusions administered
7. All medications received including antibiotics, antiemetics, antipyretics and the time administered
8. Length of surgery, start and end times of surgery and length of hospital stay
9. Whether patient was intubated and start and end times patient was on ventilator

The raw data included 6,392 unique patients with 136,996 temperature measurements across their inpatient stay, 125,370 medications administered, and 745 cases of blood transfusions. Data preprocessing resulted in almost 600 patients being excluded from the modeling, including patients with missing past medical history, those with missing or conflicting surgery start and stop times, and those patients with microbiologic cultures collected prior to the end of surgery during the index surgical hospitalization. Thus, 5,793 patients were identified and the maximum temperature at 4-hour intervals for each patient for each postoperative day up to 5 days was calculated, reducing the concern about overrepresentation of extreme values. This work was approved by the NorthShore University HealthSystem Institutional Review Board.

Descriptive summaries of key variables, trajectory analysis of temperature measurements using a censored normal distribution with important baseline characteristics such as age, gender, body mass index (BMI), whether the patient had diabetes or hypertension (HTN), and the actual procedure time (PTmins), and logit outcome analysis evaluating the proportion of patients who had a length of stay (LOS) greater than 3 days (18 time periods) was conducted.

### Model Solution



**Figure 1.** Model Solution using SAS<sup>13</sup>

Figure 1 depicts a summary of the solution procedure that embeds the trajectory modeling software in both its SAS and Stata versions with the data management and output analysis tasks. Once the data is prepared, the user specifies parameters of the model that allow evaluation of the best fit model for different counts of trajectories, polynomial functions and an appropriate data distribution (censored normal in this study), and executes the SAS procedure using SAS PROC Traj or the Stata plugin version. This procedure builds a mixture model for the trajectories, in which each trajectory has its own regression model<sup>9</sup>. The outputs of the model include trajectories that can be expressed as polynomial functions and probabilities of each individual being in each trajectory. The trajectory membership for each individual is decided by the individual's highest posterior probability of group membership.

One of the key decision points in group-based modeling is a determination of the number of groups that best represents the heterogeneity in the developmental trajectories of the population. For the  $T_{\max}$  measurement studied, models of minimum 3 trajectories and maximum 6 trajectories were built and their results compared. Model selection is based on the Bayesian Information Criterion (BIC) that determines the number of trajectory groups as well as the order of the polynomials for each trajectory group<sup>9</sup>. Normally, the model with maximum BIC is considered a better model. However, another practical criterion is to add a trajectory group only when it represents a distinct new segment of the population whose risk stratification is important from a clinical perspective.

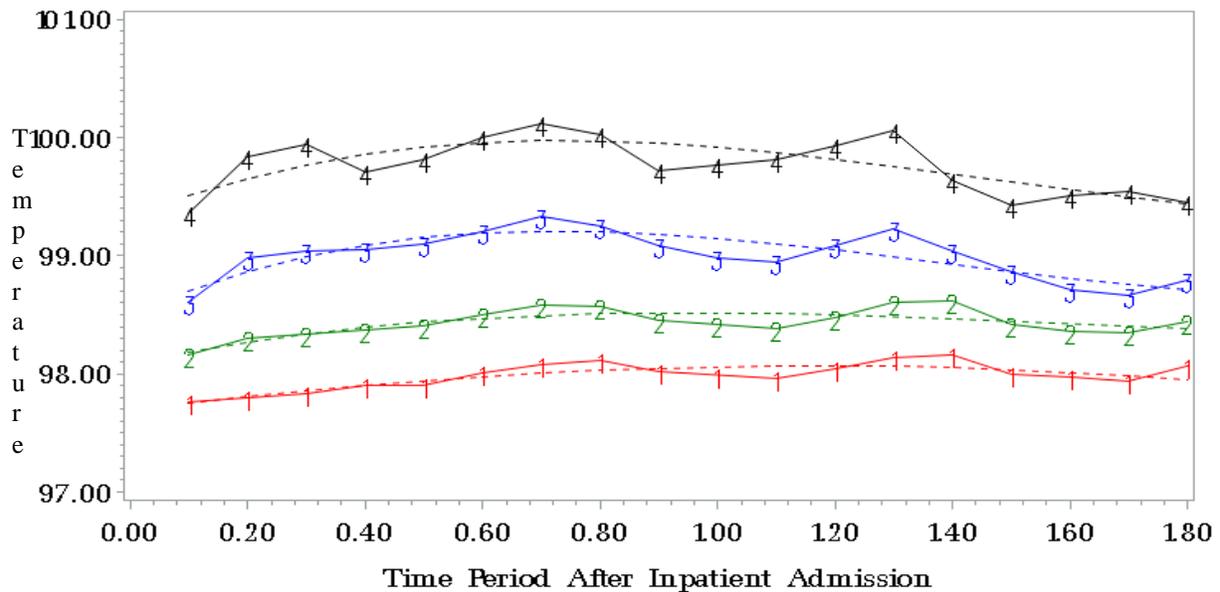
## Results

### Descriptive Summary

5,793 patients' postoperative temperature curves were evaluated. Three hundred sixty six (6.5%), 372 (6.7%), 225 (4.4%) and 30 (4.3%) patients had a  $T_{max} \geq 100.4^\circ\text{F}$  on postoperative days 0, 1, 2 and 3, respectively (patients censored on discharge), and 18.34 temperature readings, on average, in comparison to 16 readings for patients with  $T_{max} < 100.4^\circ\text{F}$  during the 72 hours after surgery. Two-thirds of the population was female (3,817), average age 68.5 years, average BMI 30.86, almost 14% had a diagnosis of diabetes, 54% had a diagnosis of hypertension (HTN), and average procedure time (length of surgery) was almost 2 hours.

### Trajectory Analysis

Using a censored normal distribution, we determined trajectories of patients' maximum temperatures (Figure 2). Four distinct trajectory clusters with increasing average  $T_{max}$  profiles were identified. 27.3% of the patients displayed group membership to trajectory 1 (red) and minimal increase in  $T_{max}$  over 3 days of inpatient stay. 53.9% were assigned to trajectory 2 (green), 16.3% belonged to trajectory 3 (blue), and 2.5% to trajectory 4 (black) with highest  $T_{max}$  elevation. Table 1 shows that less than 1% of Group 1 recorded  $T_{max} \geq 100.4^\circ\text{F}$ , and only 12.2% stayed more than 3 days, whereas 8.4% recorded  $T_{max} \geq 100.4^\circ\text{F}$  in Group 2, with more than 14% staying more than 3 days. Similarly, more than 51% recorded  $T_{max} \geq 100.4^\circ\text{F}$  in Group 3 and almost 20% stayed more than 3 days, while Group 4, though very small, had  $T_{max}$  for most of the patients at  $\geq 100.4^\circ\text{F}$  and a significant 40% were hospitalized for more than 3 days, likely due to elevated temperatures leading to potential complications. Table 2 displays the significant predictors of membership in each trajectory group - age (Group 2, likelihood estimate (LE) -0.030,  $p < 0.001$ , Group 3, LE -0.027,  $p < 0.001$  and Group 4, LE -0.034,  $p < 0.001$ ; reference Group 1) and BMI (Group 2, LE 0.034,  $p < 0.001$ , Group 3, LE 0.063,  $p < 0.001$  and Group 4, LE 0.71,  $p < 0.001$ ; reference Group 1). Group 1 displayed a quadratic functional form whereas Groups 2, 3 and 4 displayed cubic polynomial.



**Figure 2.** Maximum temperature trajectories of 5,793 postoperative knee arthroplasty patients in increments of 4 hours over 3 days

**Table 1.** Summary statistics by trajectory group

|   | Group1      | Group2      | Group3      | Group4      |
|---|-------------|-------------|-------------|-------------|
| # of Patients (Total = 5,793)                         | 1582        | 3122        | 944         | 145         |
| # of patients whose max temp $>100.4^\circ\text{F}$ * | 14 (0.9%)   | 261 (8.4%)  | 485 (51.4%) | 140 (96.5%) |
| # of patients who stayed $>3$ days (outcome =1)       | 193 (12.2%) | 450 (14.4%) | 186 (19.7%) | 58 (40%)    |

\*Note: Maximum temperature is calculated across all the temperature measurements recorded for a patient

## Outcome Analysis

Outcome analysis was performed using logit models to look at the proportion of patients with a LOS greater than 3 days. In this analysis, as you moved from the lower temperature Group 1 to the higher temperature Group 4, the proportion of the population staying past 3 days tended to increase.

**Table 2.** Maximum Likelihood estimates and logit model results of postoperative knee arthroplasty patients in increments of 4 hours (BMI: Body Mass Index; HTN: Hypertension; PTMins: Procedure Time in minutes)

| Maximum Likelihood Estimates   |           |           |           |             |           |
|--|-----------|-----------|-----------|-------------|-----------|
| Model: Censored Normal (CNORM)   |           |           |           |             |           |
|  |           | Standard  | T for H0: |             |           |
| Group  | Parameter | Estimate  | Error     | Parameter=0 | Prob >  T |
| 1  | Intercept | 97.68391  | 0.02493   | 3918.300    | 0.0000    |
|  | Linear    | 0.66251   | 0.05347   | 12.390      | 0.0000    |
|  | Quadratic | -0.28620  | 0.02962   | -9.664      | 0.0000    |
| 2  | Intercept | 98.08215  | 0.02351   | 4171.912    | 0.0000    |
|  | Linear    | 1.07023   | 0.09859   | 10.856      | 0.0000    |
|  | Quadratic | -0.80677  | 0.12635   | -6.385      | 0.0000    |
| 3  | Intercept | 98.49658  | 0.04872   | 2021.786    | 0.0000    |
|  | Linear    | 2.21936   | 0.17443   | 12.723      | 0.0000    |
|  | Quadratic | -2.06924  | 0.21279   | -9.724      | 0.0000    |
| 4  | Intercept | 99.32368  | 0.10764   | 922.757     | 0.0000    |
|  | Linear    | 1.98412   | 0.45091   | 4.400       | 0.0000    |
|  | Quadratic | -1.79307  | 0.53367   | -3.360      | 0.0008    |
|  | Cubic     | 0.40274   | 0.18722   | 2.151       | 0.0315    |
|  | Sigma     | 0.70668   | 0.00199   | 355.533     | 0.0000    |
| Group membership   |           |           |           |             |           |
| 1  | Constant  | (0.00000) | .         | .           | .         |
| 2  | Constant  | 2.08178   | 0.52437   | 3.970       | 0.0001    |
|  | Male      | -0.10346  | 0.09166   | -1.129      | 0.2590    |
|  | Age       | -0.03030  | 0.00507   | -5.975      | 0.0000    |
|  | BMI       | 0.03439   | 0.00822   | 4.183       | 0.0000    |
|  | Diabetes  | 0.31533   | 0.14049   | 2.245       | 0.0248    |
| Outcome analysis: LOGIT model  |           |           |           |             |           |
|  | GROUP1    | -1.96417  | 0.09422   | -20.847     | 0.0000    |
| 95% CI (Mean): (0.104, 0.123, 0.144)                                     |           |           |           |             |           |
|  | GROUP2    | -1.84030  | 0.06411   | -28.705     | 0.0000    |
| 95% CI (Mean): (0.123, 0.137, 0.153)                                     |           |           |           |             |           |
|  | GROUP3    | -1.29393  | 0.09329   | -13.871     | 0.0000    |
| 95% CI (Mean): (0.186, 0.215, 0.248)                                     |           |           |           |             |           |
|  | GROUP4    | -0.17795  | 0.21139   | -0.842      | 0.3999    |
| 95% CI (Mean): (0.356, 0.456, 0.559)                                     |           |           |           |             |           |
| BIC=-79348.11 (N=67869) BIC=-79297.67 (N=5793) AIC=-79161.04 L=-79120.04 |           |           |           |             |           |

## Discussion

Group based trajectory modeling was performed to characterize normal postoperative temperature over time. In this modeling strategy we found 4 distinct  $T_{max}$  temperature trajectories after elective knee replacement surgery. These temperature trajectories represent similar progression of maximum temperature mapping over four-hour time intervals. In these preliminary results, we found that the vast majority of patients belonged to the first 3 groups with only 2.5% of patients' displaying group membership to the fourth and highest  $T_{max}$  trajectory.

Age and BMI were significant covariates in all clusters with younger and high BMI patients tending to have membership in higher groups with overall higher  $T_{max}$  trajectories. The phenomenon of immune senescence or the gradual deterioration of the immune system and the inflammatory response is well described in older patients. Younger patients with their more robust immune systems and more significant inflammatory responses tend to have higher temperatures from the trauma of surgery resulting in membership in higher trajectories. Patients with higher BMI have been shown to have an increased incidence of postoperative fever and wound infection than non-obese patients undergoing hysterectomy<sup>16</sup>. It may be that the greater tissue damage in obese patients results in release of elevated levels of inflammatory mediators that generate a higher temperature and therefore result in a higher temperature trajectory.

As expected, groups with higher temperature trajectories tended to have greater proportion of patients that had LOS greater than 3 days. This may be because providers performed increased numbers of fever workups and because of a hesitation on the provider's part to discharge patients that may be perceived to have a higher complication risk. Ongoing work will evaluate the trajectories for patients' entire duration of inpatient stay, analyze complication rates for each of the four trajectories to determine whether patients with higher temperature trajectories or patients that deviate from a given trajectory are more likely to have a postoperative complication, and report performance metrics for group membership prediction using established metrics such as AUC and PPV. In addition, we are interested in evaluating how time dependent covariates such as antipyretic administration or blood transfusion affect probability of group membership and trajectory mapping.

## Conclusion

An isolated fever in the postoperative setting is a common finding, may be a normal physiologic response and therefore may offer no guidance as to whether a patient is at risk for an adverse event. Patients may display different temperature trajectories in the postoperative setting. This research serves as a necessary first step towards better determination of which patients should and should not be 'worked up' for complications of surgery. At present, 'misclassification' errors are common; most 'fever workups' are negative, but at the same time, we also commonly miss complications in patients who would have benefited from a prompt workup. A next step will be to correlate trajectory-based grouping with clinical outcomes.

## References

1. Mackowiak PA, Wasserman SS, Levine MM, Mackowiak PA, Wasserman SS, Levine MM. A critical appraisal of 98.6 degrees F, the upper limit of the normal body temperature, and other legacies of Carl Reinhold August Wunderlich.[see comment]. *JAMA*. Sep 23-30 1992;268(12):1578-1580.
2. Mackowiak PA, Mackowiak PA. Concepts of fever. *Archives of Internal Medicine*. Sep 28 1998;158(17):1870-1881.
3. Ishii Y, Noguchi H, Takeda M, Sato J, Takayama S, Toyabe S. Characteristics and significance of fever during 4 weeks after primary total knee arthroplasty. *Arch Orthop Trauma Surg*. 2014 May;134(5):707-12.
4. Athanassious C, Samad A, Avery A, Cohen J, Chalnick D. Evaluation of fever in the immediate postoperative period in patients who underwent total joint arthroplasty. *J Arthroplasty*. 2011 Dec;26(8):1404-8.
5. de la Torre SH, Mandel L, Goff BA. Evaluation of postoperative fever: usefulness and cost-effectiveness of routine workup. *Am J Obstet Gynecol*. 2003 Jun;188(6):1642-7.
6. Galicier C, Richet H. A prospective study of postoperative fever in a general surgery department. *Infect Control*. 1985 Dec;6(12):487-90.
7. Perlino CA. Postoperative fever. *Med Clin North Am*. 2001 Sep;85(5):1141-9.
8. Lesperance R, Lehman R, Lesperance K, Cronk D, Martin M. Early postoperative fever and the "routine" fever work-up: results of a prospective study. *J Surg Res*. 2011 Nov;171(1):245-50.
9. Nagin DS. 2005. *Group-Based Modeling of Development*. Harvard University Press: Cambridge, MA.
10. Goldstein H. 1995. *Multilevel Statistical Models*. 2<sup>nd</sup> ed. London: Edward Arnold.
11. Johnson MP, Zheng K, Padman R. Modeling the Longitudinality of User Acceptance of Technology with an evidence-adaptive Clinical Decision Support System. *Decision Support Systems*, 2014;57:444.
12. Nagin DS and Odgers CL. Group-based trajectory modeling in clinical research. *Annual Review of Clinical Psychology*. 2010; 6:109-138.
13. Padman R, Nagin DS, Xie Q. Disease Progression and Risk Prediction for Chronic Kidney Disease: Analysis of Electronic Health Record Data using Group-Based Trajectory Models. *Proceedings of Workshop on Information Systems and Technologies (WITS 2014)*, Auckland, New Zealand.
14. Jones BL, Nagin DS. Advances in Group-Based Trajectory Modeling and a SAS Procedure for Estimating Them. *Sociological Methods Research*, 2007, 35(4): 542-571.
15. Jones BL, Nagin DS. A Note on a STATA Plugin for Estimating Group-based Trajectory Models. (in press). *Sociological Methods and Research*, 2013, 42: 608-613.
16. Pitkin RM. Abdominal hysterectomy in obese women. *Surg Gynecol Obstet* 1974; 142:532-536.