

Modeling the Temporal Evolution of Postoperative Complications

Shara I. Feld, PhD¹, Alexander G. Cobian, MS^{2,3}, Sarah E. Tevis, MD¹,
Gregory D. Kennedy, MD PhD¹, Mark W. Craven, PhD^{3,2}

¹Department of Surgery, ²Department of Computer Sciences,
³Department of Biostatistics and Medical Informatics
University of Wisconsin-Madison

Abstract

Post-operative complications have a significant impact on patient morbidity and mortality; these impacts are exacerbated when patients experience multiple complications. However, the task of modeling the temporal sequencing of complications has not been previously addressed. We present an approach based on Markov chain models for characterizing the temporal evolution of post-operative complications represented in the American College of Surgeons National Surgery Quality Improvement Program database. Our work demonstrates that the models have significant predictive value. In particular, an inhomogenous Markov chain model effectively predicts the development of serious complications (coma longer than a day, cardiac arrest, myocardial infarction, septic shock, renal failure, pneumonia) and interventional complications (unplanned re-intubation, longer than 2 days on a ventilator and bleeding transfusion).

Introduction

Patients who suffer from post-operative complications have higher rates of post-operative morbidity and mortality, resulting in longer hospital stays, higher rates of readmission and higher cost of care [1-8]. While many studies have evaluated pre-operative predictors of complication development [1, 9-16], recent work has shown that we can extend our analysis to post-operative predictors. Tevis et al. [17] found that there are relationships in the occurrence of post-operative complications. Wakeam et al. [18] found that there are associations between complication timing and mortality. However, prior research has not investigated how sequences of multiple complications develop given post-operative events. Tevis et al. [17] employed a Bayesian network approach that revealed dependencies among multiple complications, but did not account for the ordering or timing of post-operative complications. The analysis of Wakeam et al. [18] used post-operative complications to predict mortality, but did not extend to analyzing cascades of complications.

The goal of this study was to assess whether we could accurately model the temporal evolution of postoperative complications. Specifically, we evaluated how well the likelihood and timing of complication development given prior post-operative complications could be represented using Markov chain models, a well established class of models for characterizing temporal and sequential data [19].

Methods

Complication data

The American College of Surgeons National Surgical Quality Improvement Program (ACS NSQIP) database includes 30-day postoperative outcomes for patients who underwent major inpatient and outpatient surgical procedures [20]. ACS NSQIP data are collected by a trained surgical clinical reviewer at each site, and subsequently audited for reliability by the NSQIP program. This study includes inpatient and outpatient surgical cases from over 435 institutions during the years 2005 to 2013.

The complications analyzed were the 21 reported ACS NSQIP complications occurring within 30 days after the operation and post-operative mortality (Table 1). Each recorded complication was annotated with the number of days after the operation that the complication was first diagnosed. If a complication was diagnosed multiple times post-operatively, only the first date of diagnosis was recorded. Complications included infectious types (superficial surgical site infection (SSI), deep SSI, organ space SSI, wound disruption, urinary tract infection, sepsis and septic shock), physiologic complications (peripheral nerve injury, pneumonia, deep vein thrombosis or thrombophlebitis, pulmonary embolism, renal insufficiency, renal failure, stroke or cardiovascular incident, myocardial infarction,

cardiac arrest, or coma longer than 24 hours), and interventional complications (unplanned intubation, cumulative ventilator-assisted respiration greater than 48 hours, bleeding transfusion up to 72 hours post-operatively, and graft failure requiring intervention).

Given the de-identified nature of the ACS NSQIP data, work with this dataset has been deemed exempt by the University of Wisconsin Health Sciences IRB.

Table 1. Frequency of individual complications, among all patients diagnosed with at least one post-operative complication.

Complication	Patients with Complication N = 390,646 n (%)
Infection	
Superficial SSI	65711 (17)
Deep SSI	19952 (5)
Organ SSI	34890 (9)
Wound Disruption	14914 (4)
UTI	46231 (12)
Sepsis	50047 (13)
Septic shock	27365 (7)
Physiologic	
Nerve injury	1077 (<1)
Pneumonia	40463 (10)
DVT	19299 (5)
Pulmonary Embolism	9667 (3)
Renal insufficiency	9338 (2)
Acute renal failure	11723 (3)
Stroke/CVA	6806 (2)
MI	9942 (3)
Cardiac Arrest	10690 (3)
Coma>24 hrs	1414 (<1)
Intervention	
Unplanned Intubation	33496 (9)
On ventilator>48 hrs	50335 (13)
Bleeding transfusion	140413 (37)
Graft failure	3741 (1)
Death	39575 (10)

Markov models

Markov models allow us to represent the temporal evolution of complications as patients progress through series of health states [19]. We constructed discrete-time Markov chain models to characterize the development of subsequent complications given knowledge of the complications a patient has already experienced. To develop our Markov models, we defined states that represent the patients' health condition and estimated the transition probabilities of changing from each state to another. The outcomes that we predicted were the occurrence of each of the 21 post-operative complications and mortality. The time scales of the data and model are one day steps, from the day of surgery to 30 days post-operation.

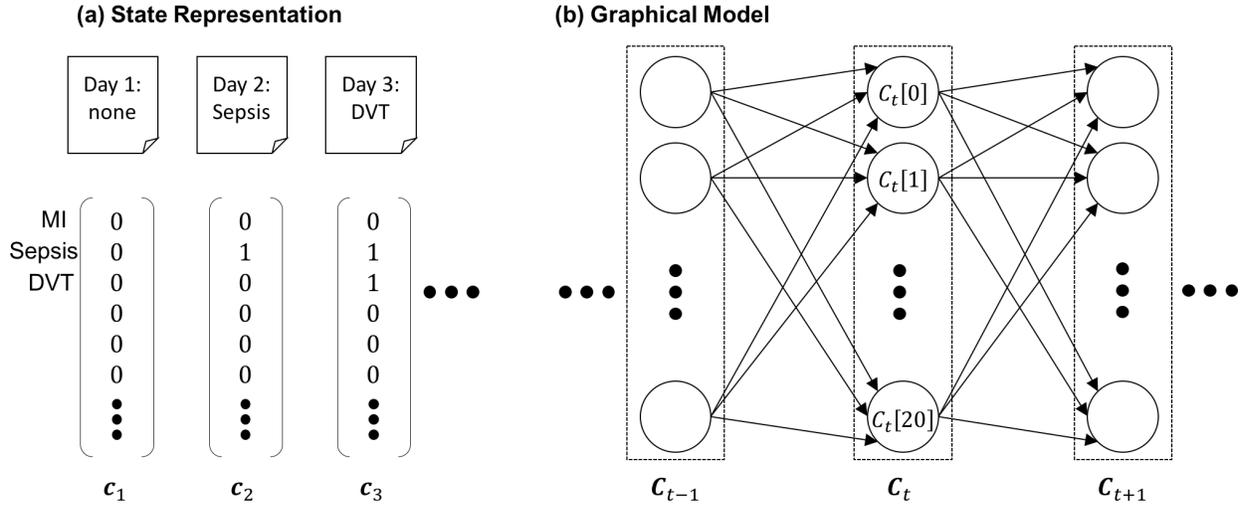


Figure 1. (a) The state representation used by our Markov model. The top part of the panel shows the new complications that are recorded on each of three days in a given patient’s record. The bottom part of the panel shows the vector of binary variables representing the patient’s state on each day. (b) The graphical model representation of our Markov model. Each node corresponds to a random variable representing the status of a given complication on a specific day. Each dashed box delineates the set of variables that represent the state for a given day. The edges impinging on each node represent the potential dependencies of the corresponding variable on the complications from the previous day.

We define a patient’s health state at a given point in time by the set of complications they have experienced prior to this point in time. With this definition, we represent a state using a vector of binary variables, one per complication, where a given variable is set to 1 if the corresponding complication has already occurred, and is set to 0 otherwise. An example using this representation is shown in Figure 1(a). The NSQIP data does not specify the durations of complications or multiple occurrences of a given complication in a patient’s record. Therefore, we designed our state representation to indicate, for a given day, all of the complications that have occurred in the patient up to that day.

Since there are 21 complications (excluding death), there are 2^{21} possible states. We use a factored state and transition representation for our models, as shown in Figure 1(b), and thus our models can also be considered dynamic Bayesian networks. Let \mathbf{C}_t represent the vector of random variables characterizing the complication state at time t , $C_t[i]$ indicate the variable for the i^{th} complication at time t , and \mathbf{c}_t and $c_t[i]$ represent the assignment of values to the vector and to the i^{th} variable, respectively.¹

We consider two different representations for state transitions. In both of them, the probability of a complication occurring at time t is conditionally independent of other complications occurring at time t , given the previous state, \mathbf{c}_{t-1} :

$$\Pr(\mathbf{c}_t | \mathbf{c}_{t-1}) = \prod_i \Pr(c_t[i] | \mathbf{c}_{t-1}).$$

A patient’s history can be viewed as a sequence of state transitions. Since the NSQIP data records events at the time granularity of days, the state transitions in our model occur on a daily basis. In our first representation for state transitions, we use a set of logistic regression models as follows. The probability of the i^{th} complication variable being 1 at time t is given by:

¹ We use bold symbols to indicate vectors, uppercase letters to indicate random variables, and lowercase letters to indicate values of random variables.

$$\Pr(C_t[i] = 1 | \mathbf{c}_{t-1}) = \begin{cases} 1 & \text{if } C_{t-1}[i] = 1 \\ p^{(i)}(\mathbf{c}_{t-1}) & \text{otherwise} \end{cases}$$

where

$$p^{(i)}(\mathbf{c}_{t-1}) = \frac{1}{1 + e^{-(\mathbf{w}^{(i)} \cdot \mathbf{c}_{t-1} + b^{(i)})}}.$$

That is, if the variable $C_{t-1}[i]$ is set to 1, indicating that the i^{th} complication had occurred by the previous time step, then it must remain in this state. Otherwise a logistic regression model is used to predict the probability of the complication occurring at time t . The logistic regression model uses the state from the previous day, \mathbf{c}_{t-1} , for its predictor variables. Note that the model is specific for the i^{th} complication and hence we use the superscript (i) for the parameters of the model, $\mathbf{w}^{(i)}$ and $b^{(i)}$. The parameters in the logistic regression models were estimated using the maximum likelihood estimation algorithm implemented in the Matlab 2015a library *glmfit* [21].

A second representation for state transitions we consider is based on the Noisy-Or function using the method of Onisko et al. [22]:

$$\Pr(C_t[i] = 1 | \mathbf{c}_{t-1}) = \begin{cases} 1 & \text{if } C_{t-1}[i] = 1 \\ 1 - (1 - p^{(i,0)}) \prod_{j: j \neq i, C_{t-1}[j]=1} \frac{1 - p^{(i,j)}}{1 - p^{(i,0)}} & \text{otherwise} \end{cases}$$

where the product in the second case ranges over the complications that have been set to 1 on the previous day, and $p^{(i,0)}$ and $p^{(i,j)}$ are defined as follows:

$$p^{(i,0)} = \Pr(C_t[i] = 1 | C_{t-1}[0] = 0, \dots, C_{t-1}[20] = 0)$$

$$p^{(i,j)} = \Pr(C_t[i] = 1 | C_{t-1}[0] = 0, \dots, C_{t-1}[j] = 1, \dots, C_{t-1}[20] = 0).$$

In other words, $p^{(i,0)}$ represents the probability of the i^{th} complication occurring on day t , given that no other complications have yet occurred, and $p^{(i,j)}$ represents the probability of the i^{th} complication occurring on day t given that only the j^{th} complication has previously occurred.

In both the logistic regression and Noisy-Or linking functions, the time scales of the data and model are one-day steps, from the day of surgery to 30 days post-operatively. The status of complications that a patient experienced through time is represented by linking the transition functions for the likelihood of complication development over the 30-day post-operative period [19].

For both the logistic regression and Noisy-Or approaches, we consider homogenous and inhomogenous variants of the models. In the homogenous case, the data from all days is pooled and then used to infer model parameters that are invariant throughout the 30-day post-operative time period. For the inhomogenous models, we partition the data into time-based subsets and then learn model parameters that are specific to given post-operative time periods. The partitioning is done such that each period contains approximately 20% of the complications that occurred across the population. These time periods were the day of surgery, 1-3 days after surgery, 3-8 days after surgery, 8-15 days after surgery and 15-30 days after surgery.

Results

Model evaluation methodology

To evaluate the extent to which our models have captured the temporal evolution and dependencies among complications, we assessed how well they are able to predict a patient's health state one day in advance given the post-operative complications the patient has experienced up to that point. We did this using a 10-fold cross validation methodology. The data were stratified by patient (i.e., a given patient's health history was either entirely in the training set or entirely in the test set) and randomly partitioned. The predictions for each test set were made using models learned from the corresponding training set. The test set predictions were then pooled for a cumulative evaluation using ROC curves.

Model predictive value

The ROC curves in Figure 2 display predictive accuracy for the homogenous and the inhomogenous models that use logistic regression to represent transition probabilities. Each curve represents the true positive rate and false positive rate of the model predicting all cases (day 0 predicting day 1, day 1 predicting day 2, up to day 29 predicting day 30) for a specific complication.

The inhomogenous models had significant predictive value for almost all complications. The models were best at predicting 30-day mortality, serious complications (coma longer than a day, cardiac arrest, MI, septic shock, renal failure, pneumonia) and interventional complications (unplanned re-intubation, longer than 2 days on a ventilator and bleeding transfusion). The complications in this set were all predicted with an area under the curve greater than 0.85. The other complications are more difficult to predict because they are likely to occur in isolation, and are not as dependent on prior complications. For example, surgical site infections are dependent on wound closure and post-operative wound management, which are not characterized in the NSQIP dataset, rather than other complications.

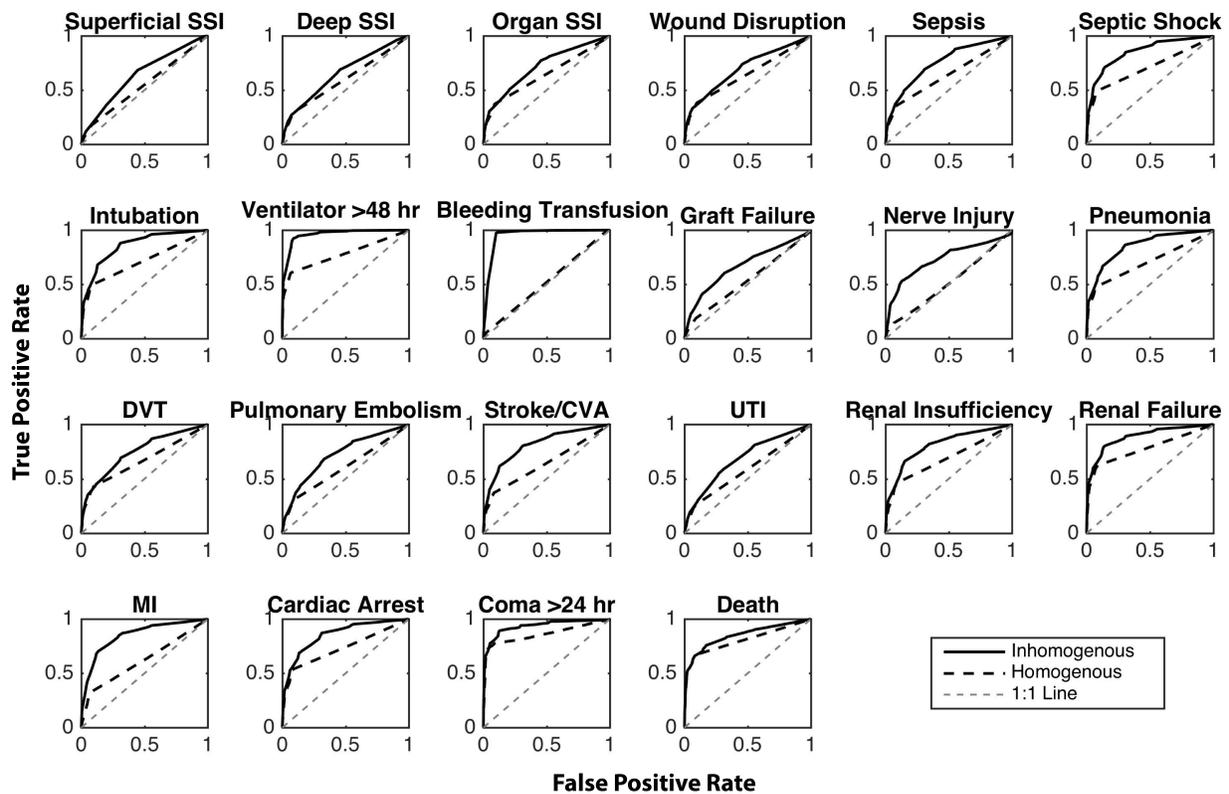


Figure 2. ROC curves for the logistic regression-based models showing the true positive rate (TPR) against the false positive rate (FPR). Solid lines show predictive accuracy for the inhomogenous models, heavy dashed lines represent the homogenous models, and the light dashed lines represent the expected curves for random guessing (i.e., no predictive value).

Comparison of homogenous and inhomogenous models

As indicated in Figure 2, the inhomogenous models had superior predictive accuracy to the homogenous models. For all complications, there was improved predictive accuracy in using the homogenous models, and for some, the improvement was dramatic. This result was true for both the models that used logistic regression and those that used Noisy-Or to represent transition probabilities. This result indicates that the transition probabilities and

dependencies vary according to the time elapsed since surgery. Not only does sequencing matter when predicting whether a complication is likely to occur, but how far out the patient is from surgery.

The improvement in predictive accuracy that comes from using inhomogenous models indicates that modeling sequences of post-operative complications requires incorporating multiple factors that influence complication development. Some complications depend on specific timing of occurrence – for example, bleeding transfusions are by design only recorded within the first 72 hours. Other complications may occur in isolation or not in a predictive fashion, such as superficial SSI. In other cases, the inhomogenous model better characterizes transition probabilities that vary throughout the post-operative period.

Comparison of logistic regression and Noisy-or models

The predictive accuracy of the logistic regression and Noisy-Or models was similar in both the inhomogenous and homogenous cases. Figure 3 shows a comparison of these two different approaches for representing transition probabilities. Each point in the figure represents a specific complication, and the coordinates of the point are given by the area under the ROC curves for the Noisy-Or and logistic-regression models. The tight clustering of the points along the diagonal indicates that the two approaches have similar predictive accuracies for every complication.

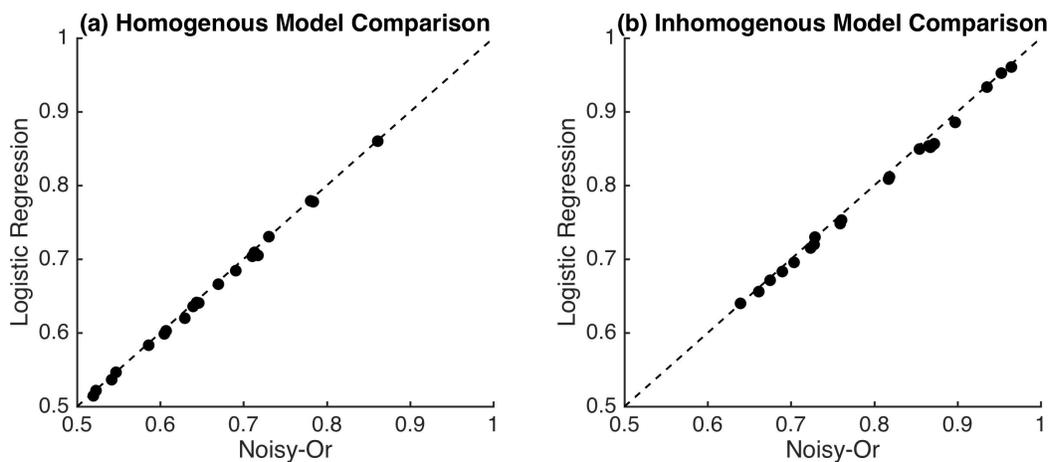


Figure 3. Scatterplots comparing area under the ROC curves for the logistic regression and Noisy-Or models. Each point represents the area under the ROC curves for a specific complication. Shown are the (a) homogenous comparison and (b) inhomogenous comparison.

How relationships between complications are highlighted

The models can be used to highlight the strength of relationships among the complications that may be anecdotally understood but not formally measured. For example, the inhomogenous logistic regression model, as depicted in Figure 4, highlights that stroke/cardiovascular event, cardiac arrest, and prolonged period on the ventilator are most predictive for developing a coma. The inhomogenous model is able to parse out the temporal variation in the extent to which predictor complications contribute to the development of the dependent complications. The inhomogenous model further highlights how complications such as surgical site infections (SSI), that may have appeared non-contributory in the homogenous model, can serve as indicators for more serious future complications at different stages of the post-operative course.

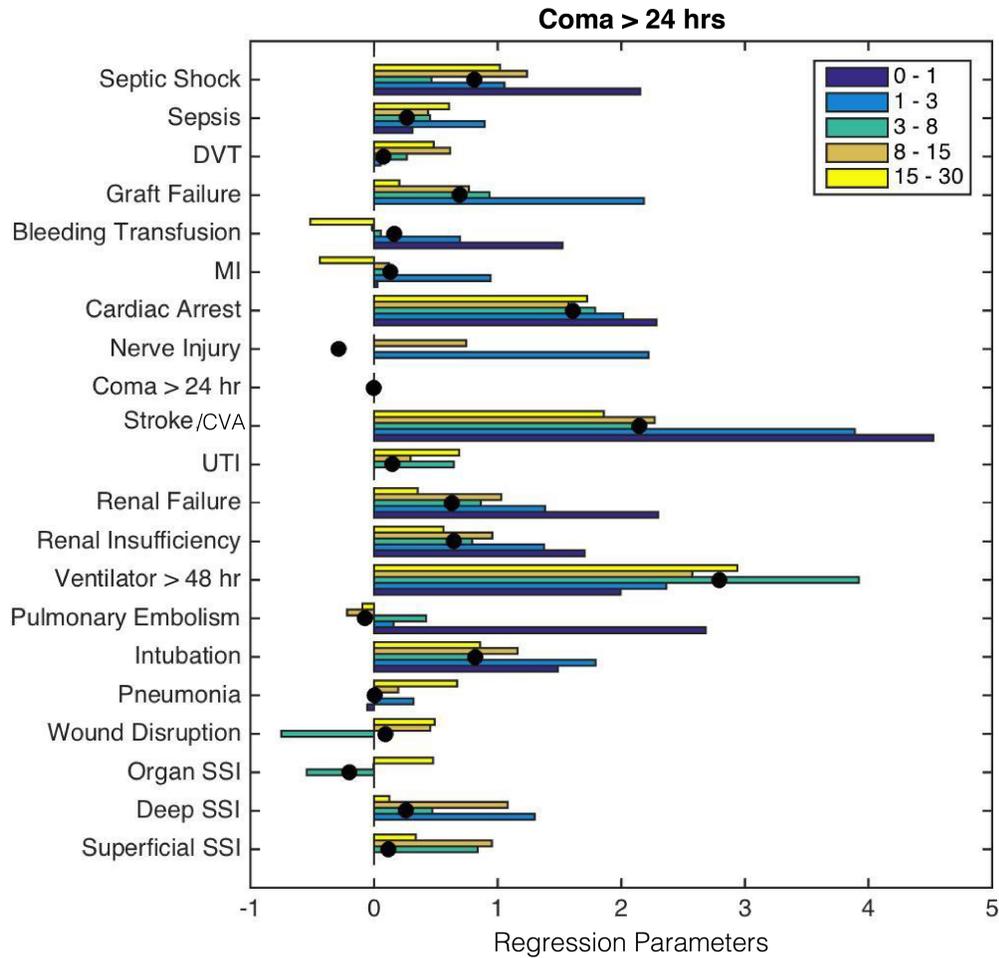


Figure 4. Parameters for the logistic regression models that predict the occurrence of the complication, coma > 24 hours. Shown are the parameters for both the homogenous model (•) and the inhomogenous model (bars). In the inhomogenous model, there are separate parameters for each time period: days 0-1 post operation, days 1-3, days 3-8, days 8-15 and days 15-30.

Conclusion

Our empirical results tell us that Markov models can effectively represent risk for later complications depending on the complications a patient has already experienced. We found that complications that are strongly dependent on a given timing (such as bleeding transfusion), dependent on prior complications (such as mortality), or both are modeled well. Complications that occur in isolation and in a wide range of times (such as superficial surgical site infection) are modeled poorly. While Tevis et al. [17] demonstrated that there are relationships among which post-operative complications occur, the present work demonstrates that those relationships are often sequential and thus have prognostic power for assessing patient outcome. The models further indicate the extent to which complications are prognostic for the subsequent occurrence of dependent complications.

Developing the model required considering timing after surgery in addition to which predictive complications have occurred. Prior work indicates that Markov models work well in applications in which the timing of events is important, which is beneficial in medical decision making [19]. We found that an inhomogeneous model, in which dependencies are conditioned upon when the events occurred over the 30-day post-operative period, outperformed a homogenous model in terms of predictive accuracy. This supports prior work noting that complication occurrence after surgery varies depending on time after surgery [24].

The standardization and size of the ACS NSQIP database also strengthens our study. The data is collected in a prospective manner by trained surgical clinical reviewers, complications are strictly defined, and the national database has a large patient population which allows us to assess the temporal dependencies in the <5% of surgical patients included who developed multiple complications.

This study has limitations inherent to the dataset and to our Markov models. We were limited by the prevalence of complications. The dataset does not indicate the durations of complications, and it only includes the first date on which a complication is diagnosed even if the complication is experienced multiple times. Furthermore, the complication diagnosis data is recorded at a temporal resolution of one day, excluding modeling dependencies between same day complications.

The limitations of our models include their first-order nature and their inability to explicitly represent dependencies between complications co-occurring on the same day. We do not view the first limitation as significant given that our state representation encodes the history of complications experienced up to the given day. Our logistic regression versus Noisy-Or comparison suggests that the inability to represent dependencies between same-day complications does not significantly hinder predictive accuracy either. Although the logistic regression models cannot explicitly represent such dependencies, they are able to compensate for their effects somewhat. Our experiments demonstrated that the logistic regression models did not offer any predictive advantages over the Noisy-Or models which assume that there are no such dependencies. Our models could be extended in a number of ways in order to capture more information about the temporal evolution of complications. Some of the more promising extensions would be to use explicit duration modeling, and to incorporate other NSQIP variables (demographics, surgical procedures, comorbidities, etc.) in the models.

Acknowledgments

This work was supported by NIH/NIDDK grant T35 DK062709, NIH/NLM grant T15 LM07359, NIH/NCATS grant UL1 TR000427, BD2K program NIH/NIAID grant U54 AI117924, and the University of Wisconsin Department of Surgery.

References

1. Longo W, Virgo K, Johnson F, Oprian C, Vernava A, Wade T, et al. Risk factors for morbidity and mortality after colectomy for colon cancer. *Diseases of the Colon & Rectum*. 2000;43(1):83-91.
2. Kassin M, Owen R, Perez S, Leeds I, Cox J, Schnier K, et al. Risk factors for 30-Day hospital readmission among general surgery patients. *Journal of the American College of Surgeons*. 2012;215(3):322-30.
3. Morris A, Baldivin L, Matthews B, Dominitz J, Barlow W, Dobie S, et al. Reoperation as a quality indicator in colorectal surgery - A population-based analysis. *Annals of Surgery*. 2007;245(1):73-9.
4. Legner V, Massarweh N, Symons R, McCormick W, Flum D. The significance of discharge to skilled care after abdominopelvic surgery in older adults. *Annals of Surgery*. 2009;249(2):250-5.
5. Kohlnhofer B, Tevis S, Weber S, Kennedy G. Multiple complications and short length of stay are associated with postoperative readmissions. *American Journal of Surgery*. 2014;207(4):449-56.
6. Lawson EH, Hall BL, Louie R, Ettner SL, Zingmond DS, Han L, et al. Association between occurrence of a postoperative complication and readmission: Implications for quality improvement and cost savings. *Annals of Surgery*. 2013;258(1):10-8.
7. Dimick JB, Chen SL, Taheri PA, Henderson WG, Khuri SF, Campbell Jr DA. Hospital costs associated with surgical complications: A report from the private-sector National Surgical Quality Improvement Program. *Journal of the American College of Surgeons*. 2004;199(4):531-7.
8. McAleese P, Odling-Smee W. The effect of complications on length of stay. *Annals of Surgery*. 1994;220(6):740-4.
9. Makary M, Segev D, Pronovost P, Syin D, Bandeen-Roche K, Patel P, et al. Frailty as a predictor of surgical outcomes in older patients. *Journal of the American College of Surgeons*. 2010;210(6):901-8.
10. de Silva S, Ma C, Proulx M, Crespin M, Kaplan B, Hubbard J, et al. Postoperative complications and mortality following colectomy for ulcerative colitis. *Clinical Gastroenterology and Hepatology*. 2011;9(11):972-80.
11. Saxton A, Velanovich V. Preoperative frailty and quality of life as predictors of postoperative complications. *Annals of Surgery*. 2011;253(6):1223-9.

12. Hawn M, Houston T, Campagna E, Graham L, Singh J, Bishop M, et al. The attributable risk of smoking on surgical complications. *Annals of Surgery*. 2011;254(6):914-20.
13. Fischer J, Wink J, Nelson J, Kovach S. Among 1,706 cases of abdominal wall reconstruction, what factors influence the occurrence of major operative complications? *Surgery*. 2014;155(2):311-9.
14. Turner P, Saager L, Dalton J, Abd-Elsayed A, Roberman D, Melara P, et al. A nomogram for predicting surgical complications in bariatric surgery patients. *Obesity Surgery*. 2011;21(5):655-62.
15. Schoenfeld AJ, Ochoa LM, Bader JO, Belmont PJ. Risk factors for immediate postoperative complications and mortality following spine surgery: A study of 3475 patients from the National Surgical Quality Improvement Program. *The Journal of Bone & Joint Surgery*. 2011; 93(17):1577-1582.
16. Cohen M, Bilimoria K, Ko C, Hall B. Development of an American College of Surgeons National Surgery Quality Improvement Program: Morbidity and mortality risk calculator for colorectal Surgery. *Journal of the American College of Surgeons*. 2009;208(6):1009-16.
17. Tevis S, Cobian A, Truong H, Craven M, Kennedy G. Implications of multiple complications on the postoperative recovery of general surgery patients. *Annals of Surgery*. 2016; 263(6):1213-18.
18. Wakeam E, Hyder JA, Tsai TC, Lipsitz SR, Orgill DP, Finlayson SR. Complication timing and association with mortality in the American College of Surgeons' National Surgical Quality Improvement Program database. *Journal of Surgical Research*. 2015;193(1):77-87.
19. Sonnenberg FA, Beck JR. Markov models in medical decision making: a practical guide. *Medical Decision Making*. 1993;13(4):322-38.
20. NSQIP. About ACS-NSQIP: participating sites; ACS NSQIP data; participant use data field. Available from: <http://www.acsnsqip.org/main/puf/PufRequestHomepage.jsp>.
21. Collett D. *Modeling Binary Data*. New York: Chapman & Hall; 2002.
22. Oniśko A, Druzdzel M, Wasyluk H. Learning Bayesian network parameters from small data sets: application of Noisy-OR gates. *International Journal of Approximate Reasoning*. 2001;27(2):165-82.
23. Fawcett T. ROC graphs: Notes and practical considerations for researchers. *Machine Learning*. 2004;31: 1-38.
24. Thompson JS, Baxter B, Allison JG, Johnson FE, Lee KK, Park W. Temporal patterns of postoperative complications. *Archives of Surgery*. 2003;138(6):596-603.