Predicting Post-Operative Complications with Wearables

Jingwen Zhang

10/03/2023





Pancreatic Surgery

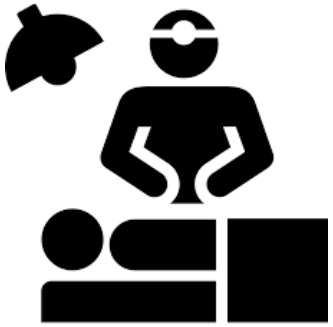
CPSL
Cyber-Physical Systems Laboratory

Pancreatic cancer is one of the most dangerous diseases:

- Low 5-year survival rate [American Cancer Society's Cancer Facts and Figures 2022]
- The **only** cure is the **surgery**



"**complications**": composite outcomes of readmission or severe complications



High rate of **readmission** and **severe complications** after the pancreatic surgery (40% - 60%) [Lubrano et al., 2018]

Surgical Outcome Prediction

CPSL
Cyber-Physical Systems Laboratory

Goal: reliable prediction of the post-operative complications

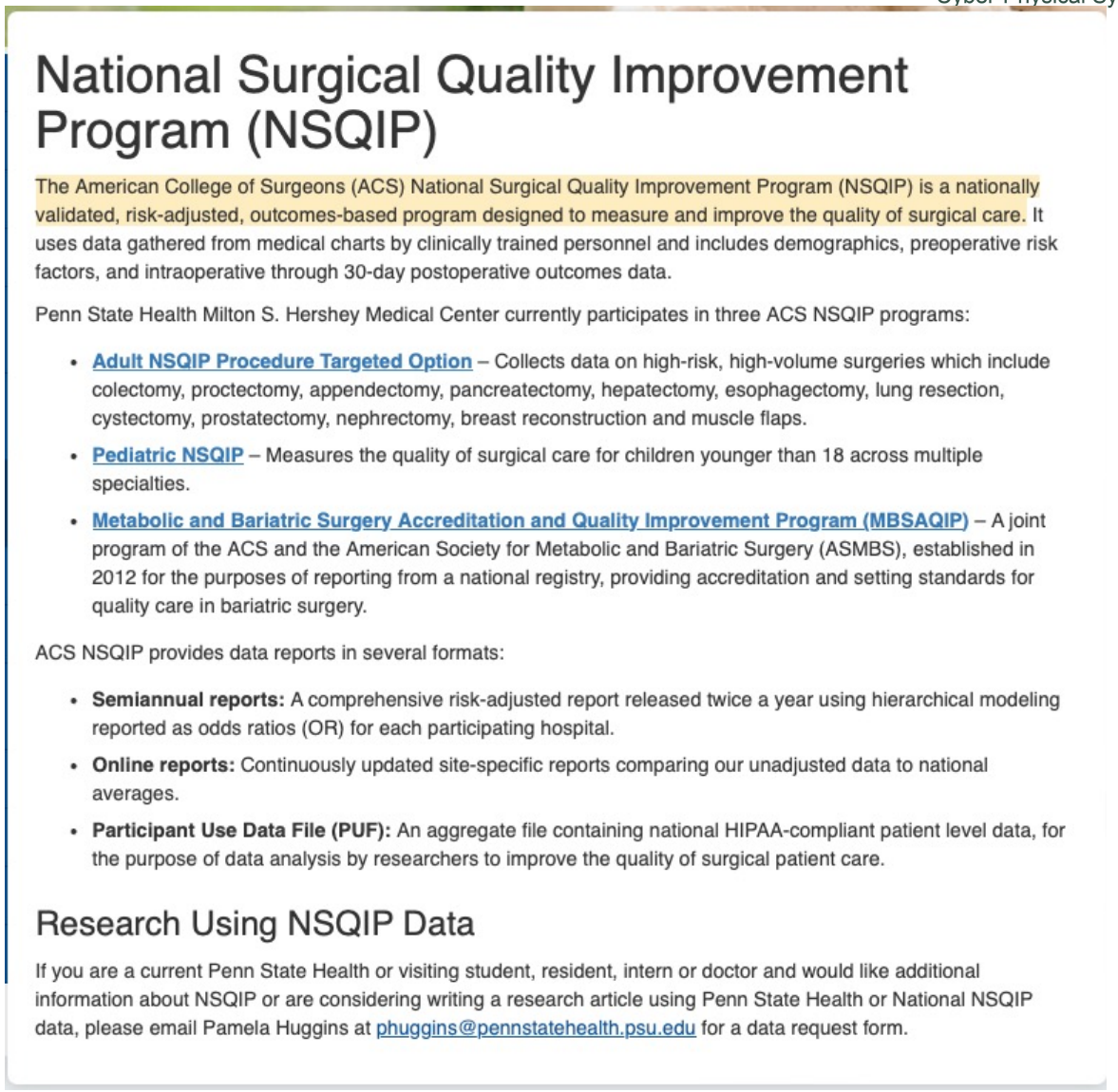
- Decision support: suitability for surgery
- Intervention: pre-habilitation





ACS-NSQIP

- Rule-based model used in practice leverages static clinical features (**ACS-NSQIP**)
- Predictive performance is not accurate



National Surgical Quality Improvement Program (NSQIP)

The American College of Surgeons (ACS) National Surgical Quality Improvement Program (NSQIP) is a nationally validated, risk-adjusted, outcomes-based program designed to measure and improve the quality of surgical care. It uses data gathered from medical charts by clinically trained personnel and includes demographics, preoperative risk factors, and intraoperative through 30-day postoperative outcomes data.

Penn State Health Milton S. Hershey Medical Center currently participates in three ACS NSQIP programs:

- Adult NSQIP Procedure Targeted Option – Collects data on high-risk, high-volume surgeries which include colectomy, proctectomy, appendectomy, pancreatectomy, hepatectomy, esophagectomy, lung resection, cystectomy, prostatectomy, nephrectomy, breast reconstruction and muscle flaps.
- Pediatric NSQIP – Measures the quality of surgical care for children younger than 18 across multiple specialties.
- Metabolic and Bariatric Surgery Accreditation and Quality Improvement Program (MBSAQIP) – A joint program of the ACS and the American Society for Metabolic and Bariatric Surgery (ASMBS), established in 2012 for the purposes of reporting from a national registry, providing accreditation and setting standards for quality care in bariatric surgery.

ACS NSQIP provides data reports in several formats:

- **Semiannual reports:** A comprehensive risk-adjusted report released twice a year using hierarchical modeling reported as odds ratios (OR) for each participating hospital.
- **Online reports:** Continuously updated site-specific reports comparing our unadjusted data to national averages.
- **Participant Use Data File (PUF):** An aggregate file containing national HIPAA-compliant patient level data, for the purpose of data analysis by researchers to improve the quality of surgical patient care.

Research Using NSQIP Data

If you are a current Penn State Health or visiting student, resident, intern or doctor and would like additional information about NSQIP or are considering writing a research article using Penn State Health or National NSQIP data, please email Pamela Huggins at phuggins@pennstatehealth.psu.edu for a data request form.

Static Clinical Features

CPSL
Cyber-Physical Systems Laboratory

- Demographics
  - e.g. age, gender, ethnicity, race
- Comorbidity information
  - e.g. liver disease, chronic kidney disease
- Clinical presentations
  - e.g. Glucose, sepsis, albumin



Static Clinical Features
From Medical Records
Measured in Hospital

Clinical Prediction with Wearables

CPSL
Cyber-Physical Systems Laboratory

Long-term, non-obtrusive monitoring



Static Clinical Features
From Medical Records





Time Series Wearable Data
(1 month before surgery)
From Fitbit

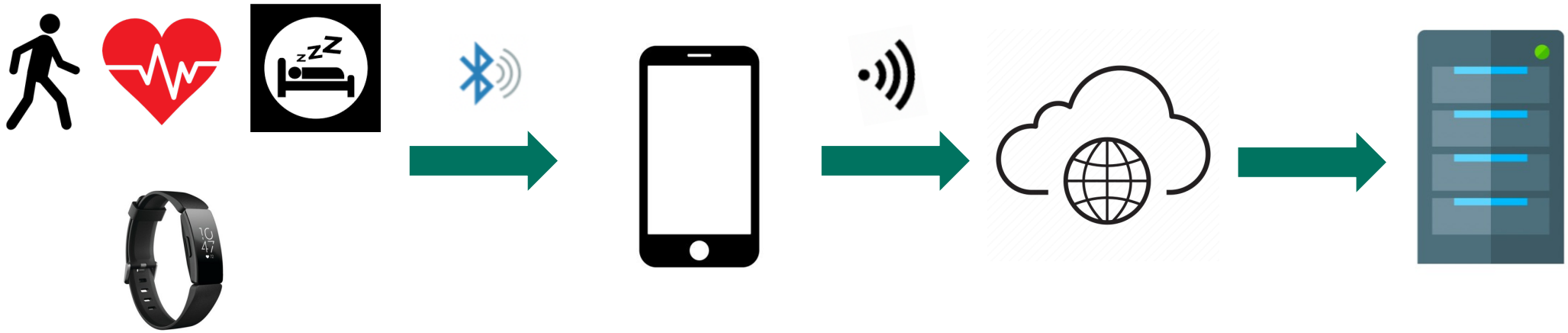


Machine Learning

Post-operative Complications
(binary classification)

Data Collection Pipeline

CPSL
Cyber-Physical Systems Laboratory



Fitbit API

CPSL
Cyber-Physical Systems Laboratory

Reference

- Device API
- Companion API
- Settings API
- **Web API**
  - Developer Guide
  - Active Zone Minutes Time Series
  - Activity
  - Activity Time Series
  - Authorization
  - Body
  - Body Time Series
  - Breathing Rate
  - Cardio Fitness Score (VO2 Max)
  - Devices
  - Electrocardiogram
  - Friends
  - Heart Rate Time Series
  - Heart Rate Variability
  - Intraday
  - Nutrition
  - Nutrition Time Series
  - Sleep

Web API Reference

Fitbit provides a set of public Web APIs that developers may use to retrieve Fitbit user data collected by the Fitbit trackers & smartwatches, Aria & Aria 2 scales, and manually entered log data. Anyone may use the Web APIs to build integrations with the Fitbit data services, so long as their application complies with the Fitbit Platform Terms of Service, the Fitbit User Data and Developer Policy, and the Fitbit user consents to share their data with the developer's application.

Here are some resources to help you get started.

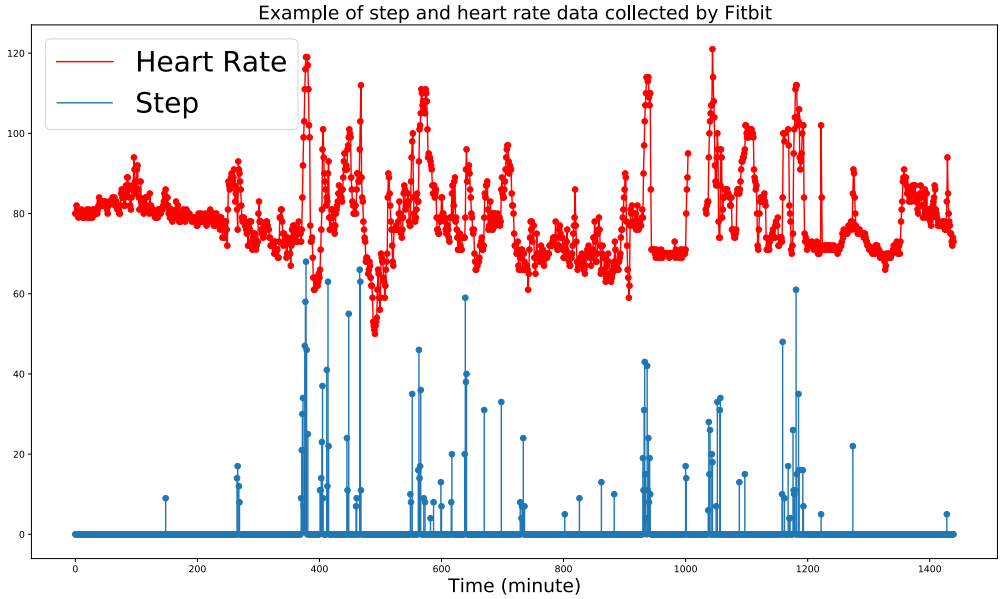
- Developer Guide contains information such as how to get started using the Web APIs, tips for designing your application, best practices and sample code.
- Troubleshooting Guide helps you debug common problems and errors, and tells you where to find some tools to help test and troubleshoot the endpoints.
- The API endpoint pages show you the proper syntax, the expected response and meaning, and any nuances that you need to be aware. The endpoints are grouped by the following,
  - Active Zone Minutes Time Series provides a user's heart-pumping activity throughout the day.
  - Activity and Activity Time Series provides information about a user's activity and their activity goals. The time series endpoints can be used to observe trends.
  - Authorization contains the endpoints needed to manage consent to Fitbit user data.
  - Body and Body Time Series provides data on a user's weight and body fat percentage. The time series endpoints can be used to observe trends.
  - Breathing Rate provides a user's average breaths per minute at night.
  - Cardio Fitness Score (VO2 Max) returns the maximum or optimum rate at which the user's heart, lungs, and muscles can effectively use oxygen during exercise.
  - Devices contains information about the devices paired to a user's account and when their last sync time.
  - Electrocardiogram contains information about the user's on-device electrocardiogram readings.
  - Friends contains information about a user's friends and their leaderboard.
  - Heart Rate Time Series returns a user's heart rate and resting heart rate values. The time series endpoints can be used to observe trends.
  - Heart Rate Variability provides the room mean square of successive differences values recorded during a user's period of sleep.
  - Intraday provides a daily, granular-level of detail to a user's active zone minutes, activity, breathing rate, heart rate, heart rate variability and SpO2 metrics.
  - Nutrition and Nutrition Time Series allows a user to record the foods they consumed and include the food's nutritional metadata. The time series endpoints can be used to observe trends.
  - Sleep returns information about a user's sleep patterns.
  - SpO2 provides a user's blood oxygen levels.
  - Subscription notifies your application via webhooks when new data is available.
  - Temperature returns a user's core and skin temperature.
  - User returns basic information about the user who shared their data with your application.

< Settings API

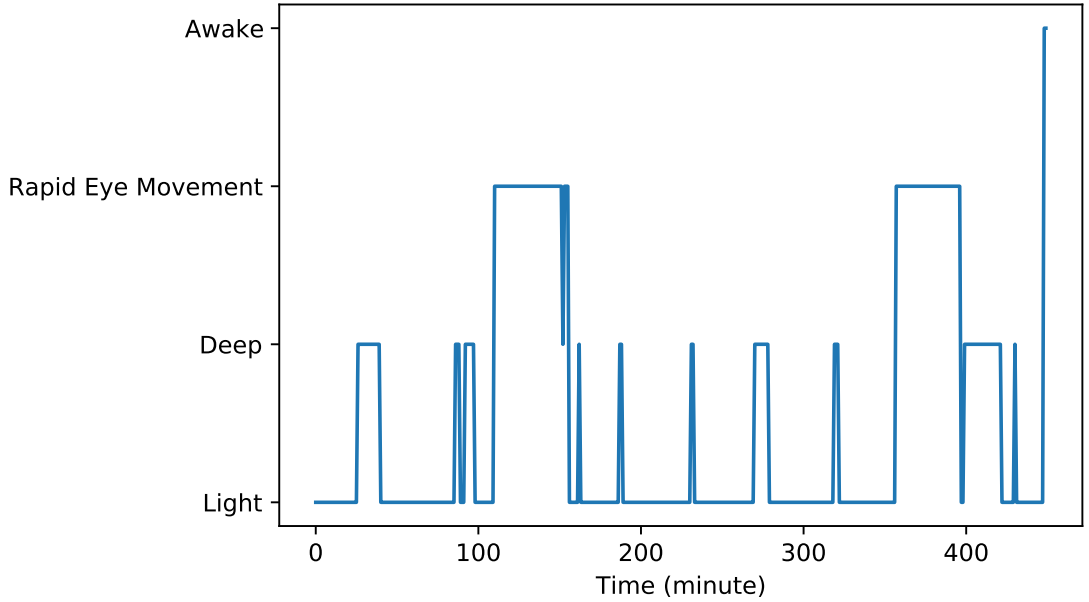
Challenges

CPSL
Cyber-Physical Systems Laboratory

- A small study cohort (61 patients)
- Fine-grained time series wearable data at minute granularity



Step and heart rate

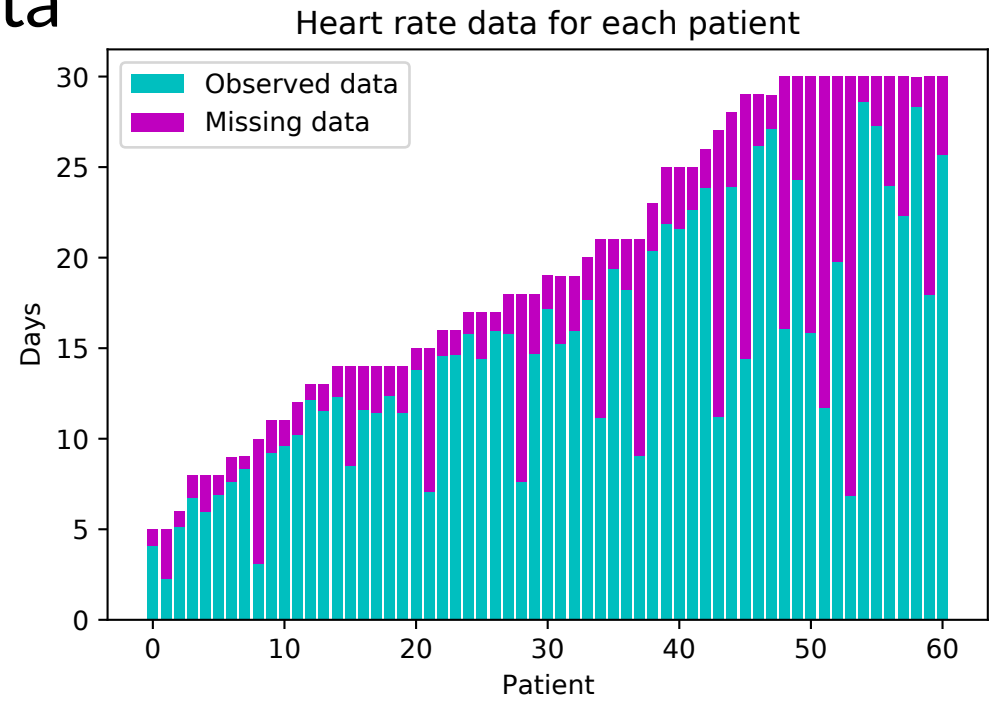


Sleep stages

Challenges

CPSL
Cyber-Physical Systems Laboratory

- A small study cohort (61 patients)
- Fine-grained time series wearable data at minute granularity
- Noisy and incomplete wearable data



Visualization of the missingness of heart rate for each patient

Previous Methods

CPSL
Cyber-Physical Systems Laboratory

- A small study cohort
- Fine-grained time series wearable data
- Noisy and incomplete wearable data

→ "Simple" machine learning models

→ Simple statistics of daily features [Bae et al., 2016, Low et al., 2019, Xu et al., 2019]

→ Imputation [Li et al., 2020]

Simple daily features robust to missing data [Wang et al., 2018]

Biorhythm features [Doryab et al., 2019]

- Lose temporal information
- May introduce extra bias
- Not good enough to describe the time series data

Feature Engineering Pipeline

CPSL
Cyber-Physical Systems Laboratory

I. Daily Feature Extraction

II. High-level Feature Extraction

Step Count

Heart Rate

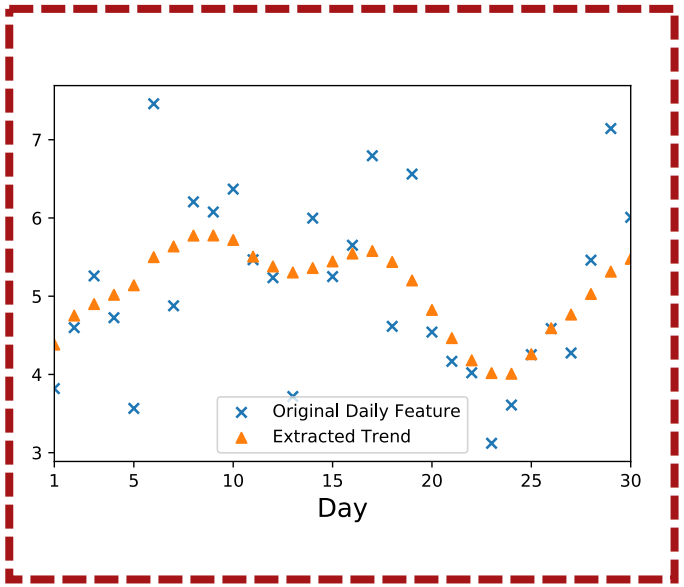
Sleep Stage

Wearable Time Series Data

N

D

D: Number of Daily Features
N: Number of Days



Static Clinical Features

+

Inputs to predictive model

Missing Wearable Data

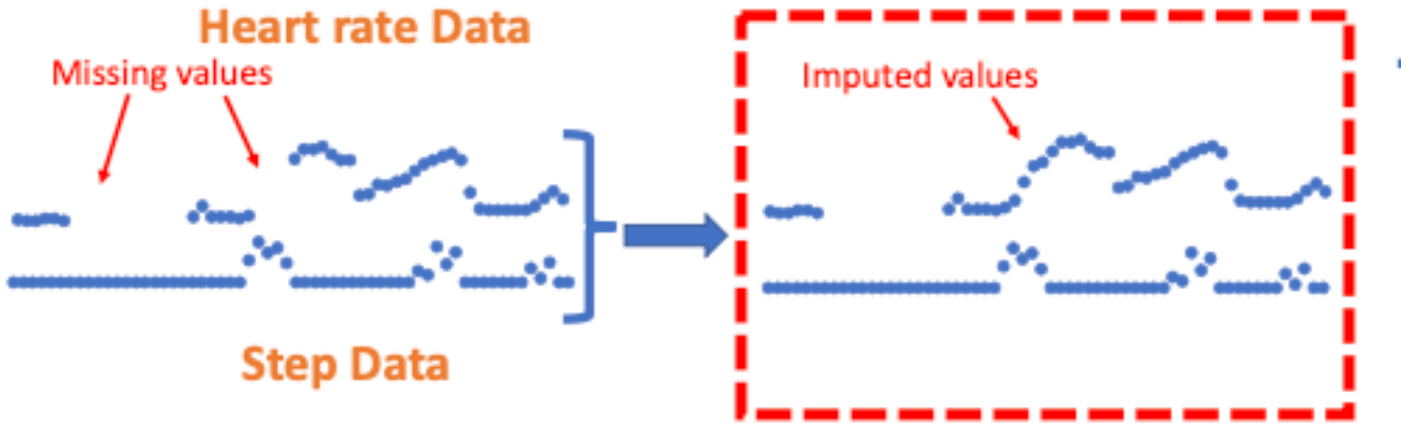
CPSL
Cyber-Physical Systems Laboratory

- **Step count**: always have values because the Fitbit device records steps as long as it is turned on
- **Heart rate**: may be complete, partially missing, or completely missing
  - Impute short missing segments
- **Sleep stage**: can be complete or completely missing for one night
  - No imputation for missing data
    - Not wearing the device
    - No heart rate
    - Zero step values

Time Series Data Imputation

CPSL
Cyber-Physical Systems Laboratory

- **Imputation threshold**: Impute short missing HR segments when the segment length <= imputation threshold (via k-nearest neighbors algorithm)
- **Extraction threshold:** Daily features are extracted when the daily yield >= extraction threshold

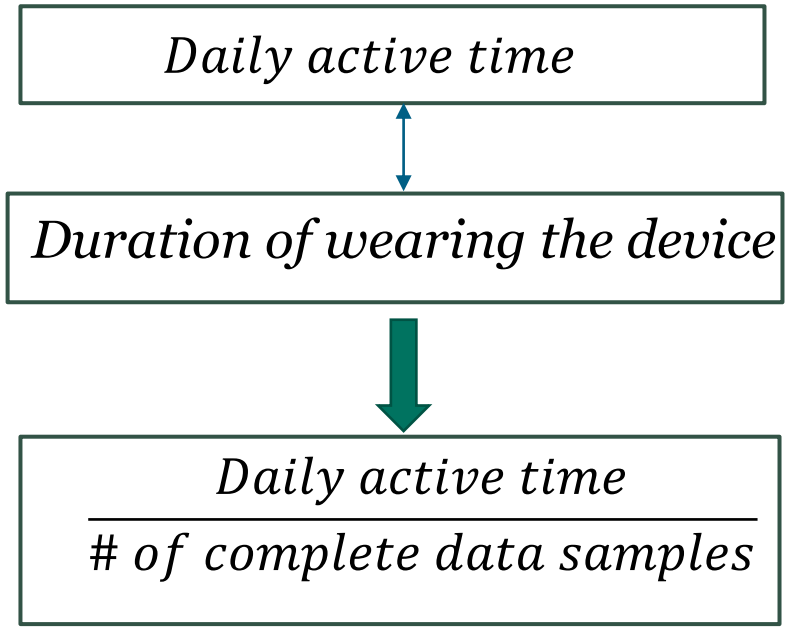


I. Daily Feature Extraction

CPSL
Cyber-Physical Systems Laboratory

- Semantic features
  - Active/sedentary information, e.g., sedentary bout counts
  - Sleep status information, e.g., time of sleep,
- Statistical features
  - First-order features, e.g., skewness
  - Second-order features, e.g., inertia
- Detrended fluctuation analysis features
  - Capture local fluctuation

Different missing rate?



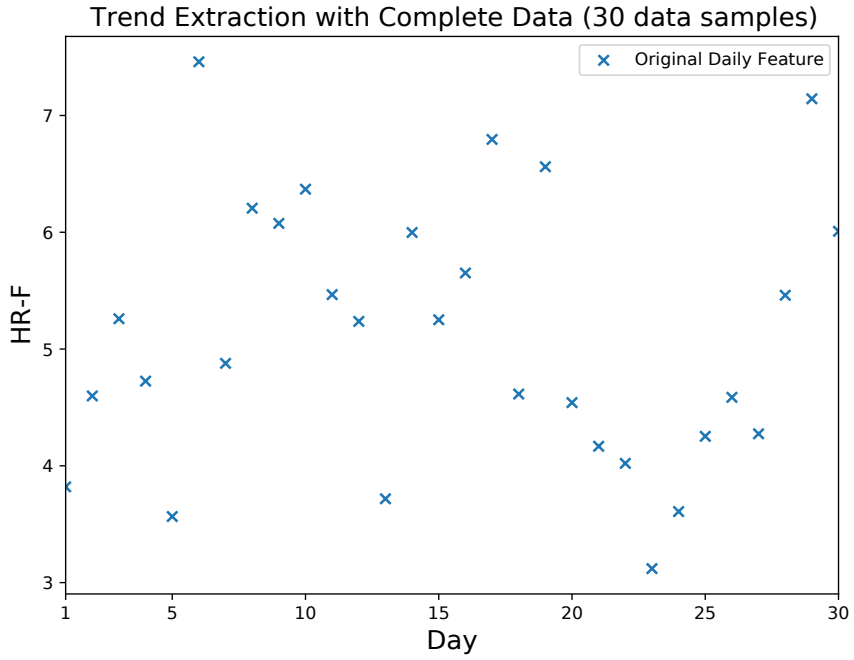
Standardization is used to accommodate missing data

II. Denoise with SSA

CPSL
Cyber-Physical Systems Laboratory

Singular spectrum analysis (SSA)

- **Denoises** time series of daily features

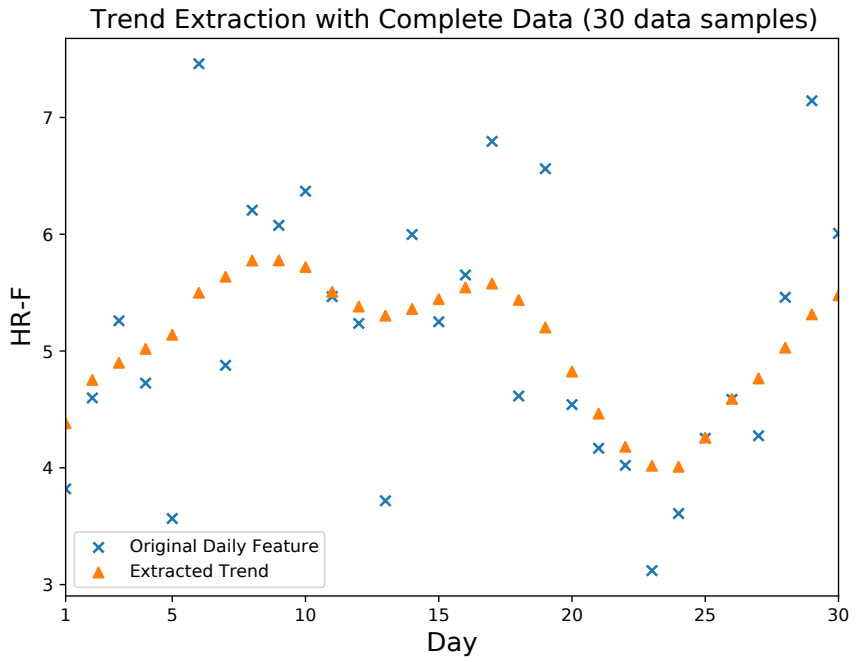


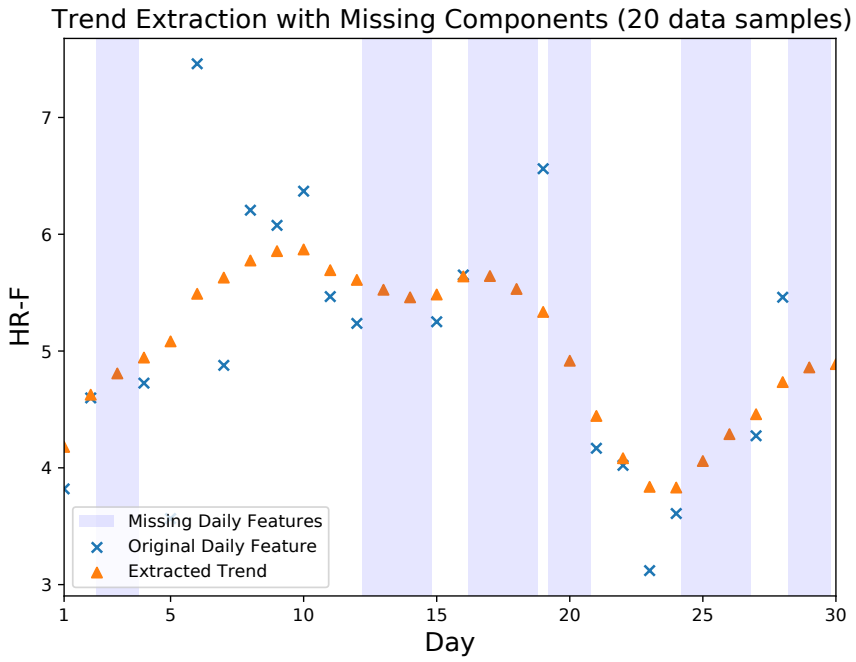
II. Denoise with SSA

CPSL
Cyber-Physical Systems Laboratory

Singular spectrum analysis (SSA)

- **Denoises** time series of daily features
- Is **robust** to missing data





- mean
- standard deviation
- slope

Clinical Study

CPSL
Cyber-Physical Systems Laboratory

Clinical study: 61 patients undergoing pancreatic surgery

Label	Number
“1” (Complications)	25
“0” (Without Complications)	36





Static Clinical Features



Wearable Time Series Data

Different Data Source

CPSL
Cyber-Physical Systems Laboratory

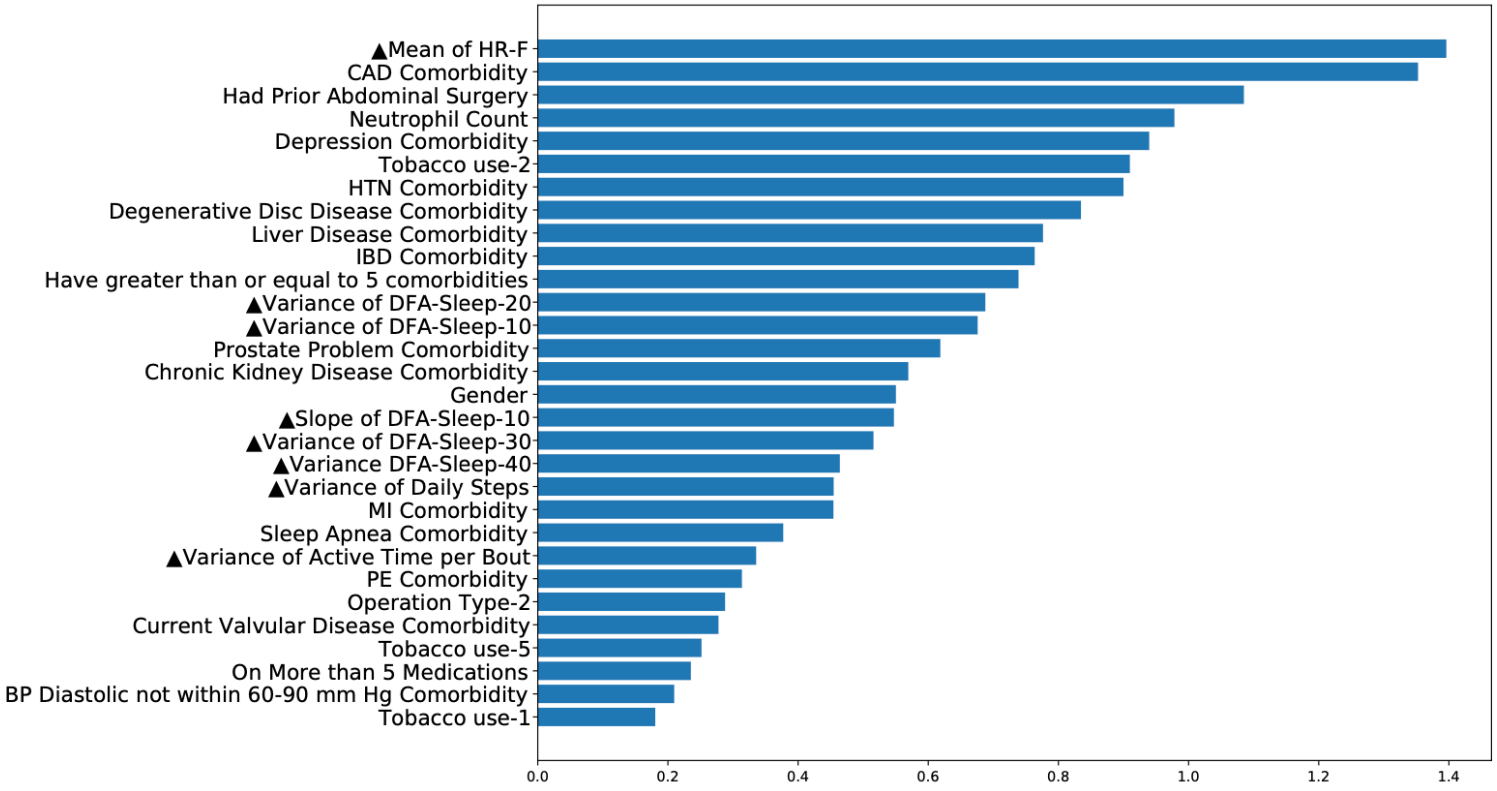
Data Source	AUROC	AUPRC	Sensitivity	Specificity
Random weighted classifier	0.5097 (0.0585)	0.4322 (0.0469)	0.1520 (0.0854)	0.8583 (0.0504)
NSQIP with Clinical Characteristics	0.6114 (0.0000)	0.4075 (0.0000)	0.2800 (0.0000)	0.8571 (0.0000)
ML with Clinical Characteristics	0.7632 (0.0085)	0.7374 (0.0206)	0.5800 (0.0699)	0.8583 (0.0083)
Wearable Data	0.7326 (0.0074)	0.7192 (0.0154)	0.5480 (0.0440)	0.8583 (0.0083)
Clinical Characteristics + Wearable Data	**0.8802 (0.0050)**	**0.8871 (0.0087)**	**0.8320 (0.0160)**	0.8583 (0.0083)

- Machine learning models outperform standard surgical risk scores
  - x2 AUPRC
  - x3 sensitivity at the same specificity
- Wearable data + clinical characteristics → best predictive performance

- NSQIP: American College of Surgeons National Surgical Quality Improvement Program
- AUROC: Area Under the Receiver Operating Characteristic Curve
- AUPRC: Area Under the Precision-Recall Curve

Feature Importance

CPSL
Cyber-Physical Systems Laboratory



Top 30 Important Features

Patients without complications:

- The heart rate can rebound more rapidly following a significant change
- Sleep status is more stable
- More active

Evaluation: Pre-operative Duration

CPSL
Cyber-Physical Systems Laboratory

Pre-operative Duration	AUROC	AUPRC	Sensitivity	Specificity
Original Dataset	0.8802 (0.0050)	**0.8871 (0.0087)**	**0.8320 (0.0160)**	0.8583 (0.0083)
5 Days for All the Patients	**0.8916 (0.0066)**	**0.8866 (0.0087)**	0.7760 (0.0265)	0.8583 (0.0083)

Predictive Performance with the Same Pre-operative Duration

Comparison to Baselines

CPSL
Cyber-Physical Systems Laboratory

Method	AUROC	AUPRC	Sensitivity	Specificity
Biorhythm Features [Doryab et al. 2019]	0.6267 (0.0000)	0.4835 (0.0000)	0.4000 (0.0000)	0.8611 (0.0000)
Min, Max, Mean of Daily Features [Li et al. 2020]	0.6844 (0.0101)	0.6211 (0.0318)	0.4520 (0.0360)	0.8528 (0.0127)
Our Feature Engineering Approach	**0.7326 (0.0074)**	**0.7192 (0.0154)**	**0.5480 (0.0440)**	0.8583 (0.0083)

Predictive Performance with Different Feature Extraction Methods on Wearable Data

Better at handling missing and noisy time series data

- Biorhythm features: cosinor analysis to describe the biorhythm
- Min, Max, Mean of daily features: forward imputation + statistics of daily features

Handling Missing and Noisy Data

CPSL
Cyber-Physical Systems Laboratory

- Impact of standardization when generating daily features (for raw time series data)

Daily Features	AUROC	AUPRC	Sensitivity	Specificity
Non-standardized	0.6919 (0.0116)	0.6332 (0.0289)	0.4320 (0.0560)	0.8611 (0.0000)
Standardized	**0.7326 (0.0074)**	**0.7192 (0.0154)**	**0.5480 (0.0440)**	0.8583 (0.0083)

→ Handle missing data

- Influence of SSA to denoise daily features (for daily features)

Feature Extraction	AUROC	AUPRC	Sensitivity	Specificity
Without SSA Denoising	0.6831 (0.0086)	0.5923 (0.0224)	0.3720 (0.0402)	0.8583 (0.0083)
With SSA	**0.7326 (0.0074)**	**0.7192 (0.0154)**	**0.5480 (0.0440)**	0.8583 (0.0083)

→ Handle noisy and missing data

Conclusion

CPSL
Cyber-Physical Systems Laboratory

- Two-level feature engineering approach for wearable data
- Explore its feasibility in a clinical study with 61 patients (AUROC = 0.8802)
- Combine wearable data and clinical characteristics
- Extract expressive high-level features from noisy and incomplete wearable data

Reference

CPSL
Cyber-Physical Systems Laboratory

[1] American Cancer Society's (ACS) publication, Cancer Facts & Figures 2022, the ACS website, the International Agency for Research on Cancer website, and the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) Program.

[2] Lubrano, Jean, et al. "Severe postoperative complications decrease overall and disease free survival in pancreatic ductal adenocarcinoma after pancreaticoduodenectomy." European Journal of Surgical Oncology 44.7 (2018): 1078-1082.

[3] Sangwon Bae, Anind K Dey, and Carissa A Low. 2016. Using passively collected sedentary behavior to predict hospital readmission. In Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing. 616–621.

[4] Carissa A Low, Dana H Bovbjerg, Steven Ahrendt, M Haroon Choudry, Matthew Holtzman, Heather L Jones, James F Pingpank Jr, Lekshmi Ramalingam, Herbert J Zeh III, Amer H Zureikat, et al. 2018. Fitbit step counts during inpatient recovery from cancer surgery as a predictor of readmission. Annals of Behavioral Medicine 52, 1 (2018), 88–92.

[5] Xuhai Xu, Prerna Chikersal, Afsaneh Doryab, Daniella K Villalba, Janine M Dutcher, Michael J Tumminia, Tim Althoff, Sheldon Cohen, Kasey G Creswell, J David Creswell, et al. 2019. Leveraging routine behavior and contextually-filtered features for depression detection among college students. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 3, 3 (2019), 1–33.

[6] Dingwen Li, Jay Vaidya, Michael Wang, Ben Bush, Chenyang Lu, Marin Kollef, and Thomas Bailey. 2020. Feasibility Study of Monitoring Deterioration of Outpatients Using Multimodal Data Collected by Wearables. ACM Transactions on Computing for Healthcare 1, 1 (2020), 1–22.

[7] Rui Wang, Weichen Wang, Alex DaSilva, Jeremy F Huckins, William M Kelley, Todd F Heatherton, and Andrew T Campbell. 2018. Tracking depression dynamics in college students using mobile phone and wearable sensing. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 2, 1 (2018), 1–26.

[8] Afsaneh Doryab, Anind K Dey, Grace Kao, and Carissa Low. 2019. Modeling Biobehavioral Rhythms with Passive Sensing in the Wild: A Case Study to Predict Readmission Risk after Pancreatic Surgery. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 3, 1 (2019), 1–21